

MÉTODOS Y COMPOSICIONES PARA EDITAR ÁCIDOS RIBONUCLEICOS

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

5 Esta solicitud reclama prioridad a una solicitud internacional con la Solicitud Internacional No. PCT/CN2018/110105 presentada el 12 de octubre de 2018 y una solicitud internacional con la Solicitud Internacional No. CN2019/082713 presentada el 15 de abril de 2019, y cuyo contenido se incorporan aquí como referencia en su totalidad.

ENVÍO DE LISTADO DE SECUENCIAS EN ARCHIVO DE TEXTO ASCII

10 El contenido de la siguiente presentación en archivo de texto ASCII es incorporado en la presente por referencia en su totalidad: una forma legible por computadora (CRF) del Listado de Secuencias (nombre del archivo: FC00158PCT-Nuevo-listado de secuencias-YZG-zhc.TXT, fecha de registro: 11 de octubre de 2019, tamaño: 104 KB).

CAMPO DE LA INVENCION

20 La presente invención está relacionada con métodos y composiciones para editar ARN usando un ARN diseñado capaz de reclutar una adenosina desaminasa para desaminar una o más adenosinas en ARN objetivo.

ANTECEDENTES DE LA INVENCION

25 La edición del genoma es una herramienta poderosa para la investigación biomédica y el desarrollo de terapias para enfermedades. Hasta ahora, la tecnología de edición de genes más popular es el sistema de repeticiones palindrómicas cortas, agrupadas y regularmente interespaciadas (CRISPR)-Cas, que se desarrolló a partir del sistema inmunológico adaptativo de bacterias y arqueas. CRISPR-Cas puede dirigir y escindir con precisión el ADN del genoma, generando rupturas de ADN de

30 doble cadena (DSB). El DSB se puede reparar a través de vías de unión de extremos no homólogos (NHEJ), lo que da como resultado una inserción o delección (Indel), que, en la mayoría de los casos, inactiva el gen. Alternativamente, la vía de reparación dirigida por homología (HDR) puede reparar el DSB usando plantillas homólogas ADNds o ADNss, y así lograr una edición precisa del genoma.

35 Recientemente, aprovechando las proteínas desaminasas, como la Adenosina Desaminasa que Actúa sobre el ARN (ADAR), se desarrollaron herramientas novedosas para la edición

de ARN. En las células de mamíferos, hay tres tipos de proteínas ADAR, Adar1 (dos isoformas, p110 y p150), Adar2 y Adar3 (catalíticamente inactivas). El sustrato catalítico de la proteína ADAR es ARN de cadena doble y puede eliminar el grupo -NH₂ de una nucleobase de adenosina (A), cambiando A a inosina (I), que se reconoce como guanosina (G) y se combina con citidina (C) durante los posteriores procesos de transcripción y traducción celular. Los investigadores fusionaron el péptido λN con el dominio de desaminasa Adar1 o Adar2 humano para construir el sistema λN-ADARDD, que podría guiarse para unirse a objetivos específicos de ARN mediante un ARN de fusión que consta de un bucle de tallo BoxB y ARN antisentido. Este método puede editar el objetivo A a I mediante la introducción de un desajuste A-C en la base A objetivo, lo que da como resultado la edición de la base del ARN A a G. Otros métodos para la edición de ARN incluyen fusionar ARN antisentido con el motivo R/G (andamio de ARN de reclutamiento de ADAR) para editar el ARN objetivo sobreexpresando las proteínas Adar1 o Adar2 en células de mamíferos, y usando dCas13-ADAR para dirigir y editar con precisión el ARN.

Las descripciones de todas las publicaciones, patentes, solicitudes de patente y solicitudes de patente publicadas a las que se hace referencia en este documento se incorporan aquí como referencia en su totalidad.

BREVE DESCRIPCIÓN DE LA INVENCION

La edición de ácidos nucleicos tiene un enorme potencial para la investigación biológica y el desarrollo de terapias. Las herramientas actuales para la edición de ADN o ARN se basan en la introducción de proteínas exógenas en organismos vivos, que están sujetos a riesgos potenciales o barreras técnicas debido a la posible actividad efectora aberrante, límites de entrega e inmunogenicidad. En algunos aspectos, la presente solicitud proporciona un enfoque programable que emplea un ARN corto para aprovechar las proteínas ADAR (Adenosina Desaminasa que Actúan sobre el ARN) endógenas para la edición de ARN dirigida. En algunos aspectos, la presente solicitud proporciona un ARN diseñado que es parcialmente complementario al transcripto objetivo para reclutar ADAR1 o ADAR2 nativo para cambiar adenosina en inosina en un sitio específico en un ARN objetivo. Los métodos descritos en este documento se denominan colectivamente "LEAPER" (Aprovechamiento de ADAR Endógeno para la Edición Programable en ARN) y los ARN de reclutamiento de ADAR se denominan indistintamente "ARNd" o "ARNar".

En un aspecto, la presente solicitud proporciona un método para editar un ARN objetivo en una célula hospedera, que comprende introducir un ARN de reclutamiento de desaminasas (ARNd) o una construcción que codifica el ARN de reclutamiento de desaminasas en la célula hospedera, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y en donde el ARNd es capaz de reclutar una adenosina desaminasa que actúa sobre el

ARN (ADAR) para desaminar una adenosina objetivo (A) en el ARN objetivo.

En determinadas modalidades, la célula hospedera es una célula eucariota. En algunas modalidades, la célula hospedera es una célula de mamífero. En algunas modalidades, la célula hospedera es una célula humana. En algunas modalidades, la célula hospedera es una célula de ratón.

5 En algunas modalidades, la célula hospedera es una célula procariota.

En determinadas modalidades, la ADAR está presente de forma natural o endógena en la célula hospedera, por ejemplo, está presente de forma natural o endógena en la célula eucariota. En algunas modalidades, la ADAR se expresa de forma endógena por la célula hospedera. En determinadas modalidades, la ADAR es exógena a la célula hospedera. En algunas modalidades, ADAR está codificada por un ácido nucleico (por ejemplo, ADN o ARN). En algunas modalidades, el método comprende introducir ADAR o una construcción que codifica ADAR en la célula hospedera. En algunas modalidades, el método no comprende la introducción de ninguna proteína en la célula hospedera. En determinadas modalidades, ADAR es ADAR1 y/o ADAR 2. En algunas modalidades, ADAR es una o más ADAR seleccionadas del grupo que consiste en hADAR1, hADAR2, ADAR1 murina y ADAR2 murina.

15 En ciertas modalidades, el ARNd no es reconocido por un Cas. En algunas modalidades, el ARNd no comprende ARNcr, ARNtracr o ARNg usados en un sistema CRISPR/Cas. En algunas modalidades, el método no comprende la introducción de una Cas o proteína de fusión Cas en la célula hospedera.

En determinadas modalidades, la desaminación del objetivo A en el ARN objetivo da como resultado una mutación sin sentido, un codón de terminación temprano, empalme aberrante o empalme alternativo en el ARN objetivo. En algunas modalidades, el ARN objetivo codifica una proteína, y la desaminación del objetivo A en el ARN objetivo da como resultado una mutación puntual, truncamiento, alargamiento y/o plegamiento incorrecto de la proteína. En algunas modalidades, la desaminación del objetivo A en el ARN objetivo da como resultado la inversión de una mutación sin sentido, un codón de terminación temprano, empalme aberrante o empalme alternativo en el ARN objetivo. En algunas modalidades, en las que el ARN objetivo codifica una proteína truncada, alargada, mutada o mal plegada, la desaminación del objetivo A en el ARN objetivo da como resultado una proteína funcional, de longitud completa, correctamente plegada y/o de tipo salvaje por reversión de una mutación sin sentido, un codón de terminación temprano, empalme aberrante o empalme alternativo en el ARN objetivo. En algunas modalidades, el ARN objetivo es un ARN regulador, y la desaminación del objetivo A da como resultado un cambio en la expresión corriente abajo de una molécula regulada por el ARN objetivo. En ciertas modalidades, el método es para aprovechar una adenosina desaminasa endógena para editar en un ARN objetivo para generar una mutación puntual y/o plegamiento incorrecto de la proteína codificada por el ARN objetivo, y/o generar un codón de terminación temprano, un sitio de empalme aberrante, y/o un sitio de empalme alternativo en el ARN

objetivo.

En determinadas modalidades, se proporciona un método para editar una pluralidad de ARN objetivo en células hospederas, en el que el método comprende introducir una pluralidad de ARNd o construcciones que codifican una pluralidad de ARNd en las células hospederas, en donde cada uno de la pluralidad de ARN que recluta desaminasas comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo correspondiente en la pluralidad de ARN objetivo, y donde cada ARNd es capaz de reclutar una adenosina desaminasa que actúa sobre el ARN (ADAR) para desaminar una adenosina objetivo (A) en el ARN objetivo correspondiente.

En algunas modalidades, se proporciona un ARN editado o una célula hospedera que tiene un ARN editado producido por cualquiera de los métodos de edición de ARN descritos anteriormente.

En un aspecto, la presente solicitud proporciona un método para tratar o prevenir una enfermedad o afección en un individuo, que comprende editar un ARN objetivo asociado con la enfermedad o afección en una célula del individuo de acuerdo con cualquiera de los métodos para la edición de ARN como se describe arriba. En algunas modalidades, el método comprende editar el ARN objetivo en la célula *ex vivo*. En algunas modalidades, el método comprende administrar la célula editada al individuo. En algunas modalidades, el método comprende administrar al individuo una cantidad eficaz del ARNd o la construcción que codifica el ARNd. En algunas modalidades, el método comprende además introducir en la célula la ADAR o una construcción (por ejemplo, un vector viral) que codifica la ADAR. En algunas modalidades, el método comprende además administrar al individuo la ADAR o una construcción (por ejemplo, un vector viral) que codifica la ADAR. En algunas modalidades, la enfermedad o afección es una enfermedad genética hereditaria. En algunas modalidades, la enfermedad o afección está asociada con una o más mutaciones genéticas adquiridas, por ejemplo, resistencia a fármacos.

Un aspecto de la presente solicitud proporciona un ARNd para la desaminación de una adenosina objetivo en un ARN objetivo mediante el reclutamiento de una Adenosina Desaminasa que Actúa sobre el ARN (ADAR), que comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo. En algunas modalidades, el ARNd comprende una secuencia de ácido nucleico de cualquiera de las SEQ ID NOs: 25-44, 142-205, 341-342.

En algunas modalidades de acuerdo con cualquiera de los métodos o ARNd descritos en este documento, el ARNd comprende una secuencia de ARN que comprende una citidina (C), adenosina (A) o uridina (U) directamente opuesta a la adenosina objetivo que se va a editar en el ARN objetivo. En determinadas modalidades, la secuencia de ARN comprende además una o más guanosinas, cada una directamente opuesta a una adenosina no objetivo en el ARN objetivo. En ciertas modalidades, el vecino más cercano 5' del objetivo A en la secuencia de ARN objetivo es un nucleótido

seleccionado entre U, C, A y G con la preferencia $U > C \approx A > G$ y el vecino más cercano 3' del objetivo A en la secuencia de ARN objetivo es un nucleótido seleccionado de G, C, A y U con la preferencia $G > C > A \approx U$. En ciertas modalidades, el objetivo A está en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN objetivo. En determinadas modalidades, en las que el motivo de tres bases es UAG, el ARNd comprende una A directamente opuesta a la U en el motivo de tres bases, una C directamente opuesta a la A objetivo y una C, G o U directamente opuesta a la G en el motivo de tres bases. En determinadas modalidades, en donde el motivo de tres bases es UAG en el ARN objetivo, el ARNd comprende ACC, ACG o ACU opuesto a la UAG del ARN objetivo.

En algunas modalidades de acuerdo con cualquiera de los métodos o ARNd descritos en este documento, el ARN de reclutamiento de desaminasas comprende más de 40, 45, 50, 55, 60, 65, 70, 75 u 80 nucleótidos. En ciertas modalidades, el ARN de reclutamiento de desaminasas tiene 40-260, 45-250, 50-240, 60-230, 65-220, 70-210, 70-200, 70-190, 70-180, 70-170, 70-160, 70-150, 70-140, 70-130, 70-120, 70-110, 70-100, 70-90, 70-80, 75-200, 80-190, 85-180, 90-170, 95-160, 100-150 o 105-140 nucleótidos de longitud. En ciertas modalidades, el ARN objetivo es un ARN seleccionado del grupo que consiste en un ARN pre-mensajero, un ARN mensajero, un ARN ribosómico, un ARN de transferencia, un ARN largo no codificante y un ARN pequeño (por ejemplo, ARNmi).

En algunas modalidades de acuerdo con cualquiera de los métodos o ARNd descritos en este documento, el ARNd es un ARN de cadena sencilla. En algunas modalidades, la secuencia de ARN complementaria es de cadena sencilla y en la que el ARNd comprende además una o más regiones de cadena doble. En algunas modalidades, el ARNd comprende una o más modificaciones, tales como modificación con 2'-O-metilo y/o modificación con fosforotioato.

En algunas modalidades, se proporciona una construcción (por ejemplo, vector viral o plásmido) que codifica cualquiera de los ARNd descritos anteriormente. En algunas modalidades, la construcción comprende un promotor operativamente unido a una secuencia que codifica el ARNd. En algunas modalidades, la construcción es una construcción de ADN. En algunas modalidades, la construcción es un vector viral, como un AAV, como AAV8.

En algunas modalidades, se proporciona una biblioteca que comprende una pluralidad de ARNd de conformidad con cualquiera de los ARNd descritos anteriormente o una pluralidad de construcciones de conformidad con cualquiera de las construcciones descritas anteriormente.

También se proporcionan composiciones, células hospederas, kits y artículos de fabricación que comprenden cualquiera de los ARNd descritos en este documento, cualquiera de las construcciones descritas en este documento o cualquiera de las bibliotecas descritas en este documento.

Las figuras 1A a 1H muestran la edición de ARN con un solo ARNd que utiliza la proteína Adar1 endógena. La figura 1A y 1B muestran representaciones esquemáticas de edición de ARN con la proteína Adar1 endógena. La figura 1C muestra la edición de ARNm reportero con ARNd utilizando la proteína Adar1 endógena. La figura 1D muestra el análisis estadístico de los resultados de la figura 1B. La figura 1E muestra la desactivación de Adar1 y los resultados de rescate de Adar1 (p110), Adar1 (p150) y Adar2. La figura 1F muestra el análisis estadístico de los resultados en la figura 1D. La figura 1G muestra el efecto de la sobreexpresión de Adar1(p110), Adar1(p150) o Adar2 en la edición de ARN mediada por ARNd en células 293T-WT. La figura 1H muestra que los resultados de la secuenciación profunda (es decir, secuenciación de próxima generación, NGS) confirmaron la edición A a G en el sitio de direccionamiento.

Las figuras 2A a 2H muestran la optimización de ARNd. La figura 2A muestra una representación esquemática de cuatro tipos de bases (A, U, C y G) que se identifican frente a la adenosina dirigida. La figura 2B muestra los efectos de la identificación de bases opuestas a la adenosina de direccionamiento en la eficiencia de edición del ARN mediante el ARNd. La figura 2C muestra una representación esquemática de ARNd con una, dos o tres bases que no coinciden con el sitio de direccionamiento de UAG. La figura 2D muestra los efectos de una, dos o tres bases que no coinciden con el sitio de direccionamiento de UAG en la edición del ARN reportero mediante ARNd. El ARNd prefirió el desajuste A-C en la adenosina de direccionamiento. La figura 2E muestra una representación esquemática de ARNd con longitud variable. La figura 2F muestra el efecto de la longitud del ARNd sobre la eficiencia de edición del ARN basado en el reportero-2 de fluorescencia dual. La figura 2G muestra una representación esquemática de diferentes posiciones de desajuste A-C. La figura 2H muestra el efecto de la posición de desajuste A-C sobre la eficiencia de edición de ARN.

Las figuras 3A a 3B muestran flexibilidad de edición para la edición de ARN endógeno a través del método de edición de ARN ejemplar de la presente solicitud. La figura 3A muestra la cuantificación porcentual de la eficacia de edición del ARN endógeno en los 16 motivos de 3 bases diferentes. La figura 3B muestra el mapa de calor de las preferencias de base 5' y 3' de la edición de ARN endógeno para 16 motivos de 3 bases diferentes.

Las figuras 4A a 4H muestran la edición del ARNm de genes endógenos con ARNd en células 293T. La figura 4A muestra una representación esquemática de ARNm objetivo y ARNd KRAS con longitud variable. La figura 4B muestra la edición del ARNm del gen KRAS endógeno con ARNd en células 293T. El vector vacío, los plásmidos ARNd-91nt se transfectaron en células 293T-WT, respectivamente. 60 horas más tarde, se aisló el ARN para RT-PCR y, a continuación, se amplificó y secuenció el ADNc en Illumina NextSeq. La figura 4C muestra la representación esquemática del objetivo de ARNm PPIB (sitio1, sitio2 y sitio3) y el diseño de ARNd correspondiente. Las figuras 4D, 4E

y 4F muestran la edición del ARNm del gen PPIB endógeno con ARNd en células 293T. La figura 4G muestra una representación esquemática del objetivo de ARNm de β -actina y el ARNd (71-nt y 131-nt). La figura 4H muestra la edición del ARNm del gen de la β -actina endógena con ARNd en células 293T.

Las figuras 5A a 5G muestran el análisis fuera del objetivo. La figura 5A muestra la
 5 representación esquemática de la ventana de secuencia en la que se analizaron las ediciones A a I para el objetivo de ARNm PPIB (sitio 1 de PPIB). La flecha negra indica la adenosina dirigida. La figura 5B muestra la cuantificación de secuenciación profunda de la edición de ARN de A a I mediante ARNd 151-nt dirigido al objetivo de ARNm PPIB (sitio 1 de PPIB). La figura 5C muestra la representación esquemática de la ventana de secuencia en la que se analizaron las ediciones A a I para el objetivo de
 10 ARNm KRAS. La flecha negra indica la adenosina dirigida. La figura 5D muestra la cuantificación de secuenciación profunda de la edición de ARN de A a I mediante ARNd 91-nt y 111-nt dirigido al objetivo de ARNm KRAS. La figura 5E muestra una representación esquemática de cuatro tipos diseñados de variantes de ARNd 91-nt o 111-nt que contienen diferentes combinaciones de desajustes A-G. El desajuste A-G se diseñó basándose en los resultados estadísticos de la figura 5D y
 15 conocimiento existente sobre códigos génicos para diferentes aminoácidos. La figura 5F muestra los resultados de la edición de A56 dirigida por ARNd y diferentes tipos de variantes de ARNd en la figura 5E. La figura 5G muestra la cuantificación de secuenciación profunda de la edición de ARN de A a I mediante ARNd 111-nt y cuatro tipos de variantes de ARNd 111-nt dirigidas al objetivo de ARNm KRAS.

Las figuras 6A a 6H muestran la edición de ARN con un solo ARNd que utiliza la
 20 proteína Adar1 endógena. La figura 6A muestra una representación esquemática de la edición de ARN mediante proteínas de fusión dLbuCas13-ADARDD. El dLbuCas13 catalíticamente inactivo se fusionó con los dominios de ARN desaminasa de ADAR1 o ADAR2. La figura 6B muestra la representación esquemática del diseño de ARN guía y objetivo de ARNm reportero de fluorescencia dual. La figura 6C muestra el análisis estadístico de los resultados en las figuras 6A y 6B. La figura 6D muestra el nivel de
 25 ARNm de Adar1 y Adar2 en células 293T-WT. La figura 6E muestra los resultados de la genotipificación del gen ADAR1 en líneas celulares 293T-ADAR1-KO mediante PCR del genoma. La figura 6F muestra el nivel de expresión de Adar1(p110) y Adar1(p150) en líneas celulares 293T-WT y 293T-ADAR1-KO mediante Western blotting. La figura 6G muestra los efectos de la sobreexpresión de Adar1(p110), Adar1(p150) o Adar2 en la edición de ARN mediada por ARNd en células 293T-WT a través de FACS. La
 30 figura 6H muestra que los resultados de la secuenciación de Sanger mostraron una edición de A a G en el sitio de adenosina dirigida.

Las figuras 7A a 7C muestran la optimización de los ARNd. La figura 7A muestra la representación esquemática de ARNd con longitud variante y los resultados de la edición de ARNm dirigido por ARNd con longitud variante basada en el reportero-1 de fluorescencia dual. La figura 7B
 35 muestra la representación esquemática de diferentes posiciones de desajuste A-C y el efecto de la

posición de desajuste A-C sobre la eficiencia de edición de ARN basada en el reportero-1 de fluorescencia dual. La figura 7C muestra una representación esquemática de diferentes posiciones de desajuste A-C y el efecto de la posición de desajuste A-C sobre la eficiencia de edición de ARN basada en el reportero-3 de fluorescencia dual.

5 Las figuras 8A a 8B muestran la edición del ARNm de genes endógenos con ARNd en células 293T. La figura 8A muestra la edición del ARNm del gen de β -actina endógena (sitio 2) con ARNd en células 293T. La figura 8B muestra la edición del ARNm del gen GAPDH endógeno con ARNd en células 293T.

10 La figura 9 muestra la edición de ARN por ARNd en diferentes líneas celulares. La figura 9 muestra que los plásmidos reporteros y los plásmidos de ARNd se cotransfectaron en diferentes líneas celulares, y los resultados mostraron que el ARNd podría funcionar bien en múltiples líneas celulares, lo que indica la universalidad de la aplicación del ARNd.

Las figuras 10A a 10D muestran la exploración de una plataforma de edición de ARN ejemplar eficiente. La figura 10A, esquema de la proteína de fusión dLbuCas13a-ADAR1_{DD} (E1008Q) y el ARNcr correspondiente. La LbuCas13a catalítica inactiva se fusionó con el dominio desaminasa de ADAR1 (variante E1008Q hiperactiva) usando un enlazador 3× GGGGS. El ARNcr (ARNcr^{Cas13a}) constaba de un andamio Lbu-ARNcr y un espaciador, que era complementario al ARN de direccionamiento con un desajuste A-C como se indica. La figura 10B, esquema del sistema reportero fluorescente dual y el Lbu-ARNcr con varias longitudes de espaciadores como se indica. La figura 10C, cuantificación de las células EGFP positivas (EGFP⁺). Se transfectaron células HEK293T que expresaban de forma estable reportero-1 con las longitudes indicadas de ARNcr^{Cas13a}, con o sin coexpresión de dLbuCas13a-ADAR1_{DD} (E1008Q), seguido de análisis FACS. Los datos se presentan como la media $\pm$ s.e.m. (n = 3). La figura 10D, FACS representativo resultado del experimento realizado con el control (ARNcr₇₀ Ctrl) o el espaciador de direccionamiento (ARNcr₇₀).

25 Las figuras 11A a 11G muestran métodos ejemplares para aprovechar la proteína ADAR1 endógena para la edición de ARN dirigida. La figura 11A, esquema del reportero-1 y el ARNar 70-nt. La figura 11B, análisis FACS representativo de la expresión de EGFP inducida por ARNar en células de tipo salvaje (HEK293T, superior) o desactivadas por ADAR1 (HEK293T ADAR1^{-/-}, inferior) que expresan de forma estable el reportero-1. La figura 11C, análisis Western blot que muestra niveles de expresión de proteínas ADAR1 en células de tipo salvaje y HEK293T ADAR1^{-/-}, así como en células HEK293T ADAR1^{-/-} transfectadas con isoformas de ADAR1 (p110 y p150). La figura 11D, análisis Western blot que muestra niveles de expresión de proteínas ADAR2 en células de tipo salvaje y HEK293T ADAR1^{-/-}, así como en células HEK293T ADAR1^{-/-} transfectadas con ADAR2. La figura 11E, cuantificación de las células EGFP positivas (EGFP⁺). El reportero-1 y las construcciones indicadas que expresan ADAR se cotransfectaron en células HEK293T ADAR1^{-/-}, junto con ARN₇₀ Ctrl o con el ARNar₇₀

de direccionamiento, seguido de análisis FACS. Los porcentajes de EGFP⁺ se normalizaron mediante la eficiencia de transfección, que fue determinada por mCherry⁺. Los datos son valores medios $\pm$ s.e.m. (n = 4). La figura 11F, los electroferogramas que muestran los resultados de secuenciación de Sanger en la región objetivo ARN₇₀ Ctrl (superior) o ARN₇₀ (inferior). La figura 11G, cuantificación de la tasa de conversión de A a I en el sitio de dirigido mediante secuenciación profunda.

Las figuras 12A a 12B muestran el nivel de expresión de ARNm de ADAR1/ADAR2 y la edición de ARN mediada por ARN_{ar}. La figura 12A, PCR cuantitativa que muestra los niveles de ARNm de ADAR1 y ADAR2 en células HEK293T. Los datos se presentan como la media $\pm$ s.e.m. (n = 3). La figura 12B, FACS representativo resulta de la figura 1E.

La figura 13 muestra resultados de PCR cuantitativa que demuestran los efectos de un método LEAPER ejemplar sobre los niveles de expresión de transcritos de reportero-1 dirigidos por ARN_{ar} 111-nt o ARN de control en células HEK293T. Los datos se presentan como la media $\pm$ s.e.m. (n = 3); prueba t de Student de dos lados no apareados, ns, no significativa.

Las figuras 14A a 14D muestran la edición de ARN dirigido con un método LEAPER ejemplar en múltiples líneas celulares. La figura 14A, resultados del Western blot que muestran los niveles de expresión de ADAR1, ADAR2 y ADAR3 en las líneas celulares humanas indicadas. Se utilizó β -tubulina como control de carga. Los datos mostrados son representativos de tres experimentos independientes. ADAR1^{-/-}/ADAR2 representa células HEK293T inactivas para ADAR1 que sobreexpresan ADAR2. La figura 14B, niveles de expresión de la proteína ADAR relativa normalizados por la expresión de β -tubulina. La figura 14C, se transfectaron células humanas indicadas con reportero-1, junto con el ARN_{ar} de control 71-nt (ARN₇₁ Ctrl) o con el ARN_{ar} objetivo 71-nt (ARN_{ar71}) seguido de análisis FACS. La figura 14D, las líneas celulares de ratón indicadas se analizaron como se describe en la figura 14C. Los porcentajes de EGFP⁺ se normalizaron mediante la eficiencia de transfección, que fue determinada por mCherry⁺. Las barras de error en las figuras 14B, 14C y 14D indican la media $\pm$ s.e.m. (n = 3); prueba t de Student de dos lados no apareados, *P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001; ns, no significativo.

Las figuras 15A a 15C muestran esquemas de reportero-1 (a), -2 (b) y -3 (c), así como sus correspondientes ARN_{ar}.

Las figuras 16A a 16G muestran la caracterización y optimización de métodos LEAPER ejemplares. La figura 16A, arriba, esquema del diseño de ARN_{ar} con triplete cambiado (5'-CNA, N denota A, U, C o G) opuesto al UAG objetivo. Abajo, porcentaje de EGFP⁺ que muestra los efectos de bases variables opuestas a la adenosina objetivo sobre la eficiencia de edición de ARN. La figura 16B, arriba, el diseño de ARN_{ar} con bases vecinas cambiadas que flanquean la citidina en el desajuste A-C (5'-N¹CN²). Abajo, los efectos de 16 combinaciones diferentes de N¹CN² sobre la eficiencia de la edición de ARN. La figura 16C, resumen de la preferencia de los sitios vecinos más cercanos 5' y 3' de la

citidina en el desajuste A-C. La figura 16D, superior, el diseño de ARNar de longitud variable. Abajo, el efecto de la longitud del ARNar sobre la eficiencia de edición del ARN basado en reportero-1 y reportero-2. La figura 16E, arriba, el diseño de ARNar con posición variable de desajuste A-C. Abajo, el efecto de la posición de desajuste A-C en la eficiencia de edición de ARN basado en el reportero 1 y reportero-2. La figura 16F, superior, el diseño de los motivos de triplete en el reportero-3 con bases vecinas más cercanas variables que rodean la adenosina de direccionamiento (5'-N¹CN²) y el motivo opuesto (5'-NN²CN¹) en el ARNar 111-nt (ARNar₁₁₁). Abajo, resultados de secuenciación profunda que muestran la tasa de edición en la adenosina objetivo en el motivo 5'-N²CN¹. La figura 16G, resumen de las preferencias básicas de 5' y 3' de la edición mediada por LEAPER en reportero-3. Las barras de error en las figuras 16A, 16B, 16D, 16E y 16F indican valores medios $\pm$ s.e.m. (n = 3).

Las figuras 17A a 17I muestran la edición de transcriptos endógenos con métodos LEAPER ejemplares. La figura 17A, esquema de los transcriptos endógenos de direccionamiento de cuatro genes relacionados con la enfermedad (PPIB, KRAS, SMAD4 y FANCC) y los ARNar correspondientes. La figura 17B, resultados de secuenciación profunda que muestran la tasa de edición en la adenosina dirigida de los transcriptos de PPIB, KRAS, SMAD4 y FANCC mediante la introducción de las longitudes indicadas de ARNar. La figura 17C, resultados de secuenciación profunda que muestran la tasa de edición en sitios no UAN de transcriptos endógenos de PPIB, FANCC e IDUA. La figura 17D, velocidad de edición multiplexada por dos ARNar 111-nt. Los ARNar indicados se transfectaron solos o se cotransfectaron en las células HEK293T. La edición dirigida en los dos sitios se midió a partir de células cotransfectadas. La figura 17E, esquema de la secuencia de transcripción de PPIB cubierta por el ARNar 151-nt. La flecha negra indica la adenosina dirigida. Todas las adenosinas se marcaron en rojo. La figura 17F, mapa de calor de la tasa de edición en adenosinas cubiertas por las longitudes indicadas de ARNar que se dirigen al gen PPIB (marcado en azul en negrita). Para la región cubierta de ARNar 111-nt o ARNar₁₅₁-PPIB, las tasas de edición de A22, A30, A33 y A34 se determinaron mediante ARN-sec debido a la falta de cebadores de PCR efectivos para amplificar esta región. De lo contrario, la tasa de edición se determinó mediante un análisis de secuenciación profunda dirigida. La figura 17G, arriba, el diseño de los motivos del triplete en el reportero-3 con bases vecinas más cercanas variables que rodean la adenosina de direccionamiento (5'-N¹AN²) y el motivo opuesto (5'-N²GN¹) en el ARNar 111-nt (ARNar₁₁₁). Abajo, resultados de secuenciación profunda que muestran la tasa de edición. La figura 17H, arriba, el diseño de ARNar con dos desajustes consecutivos en el motivo 5'-N¹GN² opuesto al motivo 5'-UAG o los motivos 5'-AAG. Resultados de la secuenciación profunda que muestran la velocidad de edición de un ARNar₁₁₁ con dos desajustes consecutivos en el motivo 5'-N¹GN² opuesto al motivo 5'-UAG (parte inferior izquierda) o el motivo 5'-AAG (parte inferior derecha). La figura 17I, mapa de calor de la tasa de edición en adenosinas cubiertas por variantes de ARNar₁₁₁ diseñadas que se dirigen al gen KRAS. Los datos de las figuras 17B, 17C, 17D, 17G y 17H se

presentan como la media $\pm$ s.e.m. ($n = 3$); prueba t de Student de dos lados no apareados, $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$; NS, no significativo. Los datos en (f e i) se presentan como la media ($n = 3$).

Las figuras 18A a 18B muestran efectos de métodos LEAPER ejemplares sobre los niveles de expresión de productos proteicos y transcritos dirigidos. La figura 18A, PCR cuantitativa que muestra los niveles de expresión de transcritos dirigidos de PPIB, KRAS, SMAD4 y FANCC mediante el ARNar 151-nt correspondiente o el ARN de control en células HEK293T. Los datos se presentan como la media $\pm$ s.e.m. ($n = 3$); prueba t de Student de dos lados no apareados, $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$; ns, no significativo. La figura 18B, resultados de Western blot que muestran los efectos sobre los productos proteicos del gen KRAS dirigido por ARNar 151-nt en células HEK293T. Se utilizó β -tubulina como un control de carga.

Las figuras 19A a 19F muestran la edición de transcritos endógenos con métodos LEAPER ejemplares. La figura 19A, esquema de la secuencia del transcrito de KARS cubierto por el ARNar 151-nt. La flecha indica la adenosina dirigida. Todas las adenosinas se marcaron en rojo. La figura 19B, mapa de calor de la velocidad de edición en adenosinas cubiertas por ARNar indicados en el transcrito de KARS (marcado en azul en negrita). La figura 19C, esquema del transcrito de SMAD4 cubierto por el ARNar 151-nt. La figura 19D, mapa de calor de la tasa de edición en adenosinas cubiertas por ARNar indicados en el transcrito de SMAD4. La figura 19E, esquema del transcrito de FANCC cubierto por el ARNar 151-nt. La figura 19F, mapa de calor de la tasa de edición en adenosinas cubiertas por ARNar indicados en el transcrito de FANCC. Para cada ARNar, la región del ARN dúplex se resalta con un marco en negrita en azul. Datos (figuras 19B, 19D y 19F) se presentan como la media ($n = 3$).

Las figuras 20A a 20D muestran la especificidad de todo el transcriptoma de la edición de ARN por LEAPER. Las figuras 20A y 20B, análisis fuera del objetivo de todo el transcriptoma de ARN₁₅₁ Ctrl y ARNar₁₅₁-PPIB. El sitio de direccionamiento (PPIB) está resaltado en rojo. Los posibles sitios fuera del objetivo identificados tanto en ARN Ctrl como en los grupos de ARN dirigidos a PPIB están marcados en azul. La figura 20C, la afinidad de hibridación predicha entre sitios fuera del objetivo y el correspondiente ARN₁₅₁ Ctrl o ARNar₁₅₁-

PPIB. La energía libre mínima (ΔG) del ARN de doble cadena formado por sitios fuera del objetivo (150-nt corriente arriba y corriente abajo de los sitios de edición) y el correspondiente ARN₁₅₁ Ctrl o ARNar₁₅₁-PPIB se predijo con ARNhíbrido, una herramienta de sitio web en línea. La figura 20D, superior, esquema de la región altamente complementaria entre ARNar₁₅₁-PPIB y los sitios potenciales fuera del objetivo indicados, que se predijeron mediante la búsqueda de secuencias homólogas a través de NCBI-BLAST. Abajo, secuenciación profunda que muestra la tasa de edición en el sitio objetivo y todos los sitios fuera del objetivo previstos de ARNar₁₅₁-PPIB. Los datos se presentan

como la media $\pm$ s.e.m. ($n = 3$).

Las figuras 21A a 21B muestran la evaluación de posibles fuera de objetivos. La figura 21A, esquema de la región altamente complementaria de ARNar₁₁₁-FANCC y la secuencia potencial indicada fuera del objetivo, que se predijeron mediante la búsqueda de secuencias homólogas a través de NCBI-BLAST. La figura 21B, secuenciación profunda que muestra la tasa de edición en el sitio objetivo y todos los sitios fuera del objetivo predichos de ARNar₁₁₁-FANCC. Todos los datos se presentan como la media $\pm$ s.e.m. ($n = 3$).

Las figuras 22A a 22F muestran la evaluación de seguridad de la aplicación de métodos LEAPER ejemplares en células de mamífero. Las figuras 22A y 22B, análisis de todo el transcriptoma de los efectos de ARN₁₅₁ Ctrl (a) ARNar₁₅₁-PPIB (b) en sitios de edición nativos mediante la secuenciación de ARN de todo el transcriptoma. El análisis del coeficiente de correlación de Pearson se utilizó para evaluar la tasa de edición de ARN diferencial en sitios de edición nativos. Las figuras 22C y 22D, análisis de expresión génica diferencial de los efectos de ARN₁₅₁ Ctrl (c) ARNar₁₅₁-PPIB (d) con datos de ARN-sec a nivel del transcriptoma. Se utilizó el análisis del coeficiente de correlación de Pearson para evaluar la expresión genética diferencial. Las figuras 22E y 22F, efecto de la transfección de ARNar sobre la respuesta inmune innata. Los ARNar indicados o el poli(I:C) se transfectaron en células HEK293T. Luego se analizó el ARN total usando PCR cuantitativa para determinar los niveles de expresión de IFN- β (e) e IL-6 (f). Los datos (e y f) se presentan como la media $\pm$ s.e.m. ($n = 3$).

Las figuras 23A a 23D muestran la recuperación de la actividad reguladora transcripcional del mutante TP53W53X por LEAPER. La figura 23A, arriba, esquema de la secuencia del transcripto de TP53 cubierto por el ARNar 111-nt que contiene c.158G>A mutación sin sentido clínicamente relevante (Trp53Ter). La flecha negra indica la adenosina dirigida. Todas las adenosinas se marcaron en rojo. Abajo, el diseño de dos ARNar optimizados dirigidos a transcriptos TP53^{W53X} con desajuste A-G en A⁴⁶⁰ para ARNar₁₁₁-AG1, y en A¹⁶⁰, A⁴⁶⁰, A⁹¹⁰ y A⁹⁴⁰ juntos para ARNar₁₁₁-AG4 para minimizar los posibles fuera de objetivos en motivos "propensos a la edición". La figura 23B, resultados de secuenciación profunda que muestran la edición dirigida en transcriptos TP53^{W53X} por ARNar₁₁₁, ARNar₁₁₁-AG1 y ARNar₁₁₁-AG4. La figura 23C, Western blot que muestra la producción recuperada de la proteína p53 de longitud completa a partir de los transcriptos TP53^{W53X} en las células HEK293T TP53^{-/-}. La figura 23D, detección de la actividad reguladora de la transcripción de la proteína p53 restaurada usando un sistema indicador de luciferasa p53-luciérnaga, normalizado por el vector de luciferasa Renilla cotransfectado. Los datos (b, c y d) se presentan como la media $\pm$ s.e.m. ($n = 3$); prueba t de Student de dos lados no apareados, * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$; ns, no significativo.

La figura 24 muestran la edición de transcriptos de TP53W53X mutantes mediante un método LEAPER ejemplar. Arriba, esquema de la secuencia del transcripto de TP53 cubierto por los

ARNar 111-nt. La flecha indica la adenosina dirigida. Todas las adenosinas se marcaron en rojo. Abajo, un mapa de calor de la velocidad de edición de las adenosinas cubiertas por los ARNar indicados en el transcripto de TP53.

La figura 25 muestra una representación esquemática del ADNc relevante para la enfermedad seleccionada que contiene la mutación G a A de los datos de ClinVar y el ARNar 111-nt correspondiente.

La figura 26 muestra la corrección de mutaciones patógenas mediante un método LEAPER ejemplar. Corrección de A a I de la mutación G>A relevante para la enfermedad de los datos de ClinVar mediante el ARNar 111-nt correspondiente, dirigido a mutaciones relacionadas con la clínica de seis genes patógenos como se indica (figura 25 y los cuadros de las secuencias de ARNar y ARN de control y ADNc relacionado con la enfermedad a continuación). Los datos se presentan como la media $\pm$ s.e.m. (n = 3); prueba t de Student de dos lados no apareados, * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001; ns, no significativo.

Las figuras 27A a 27C muestran la edición de ARN en múltiples células primarias humanas mediante métodos LEAPER ejemplares. La figura 27A, cuantificación de las células EGFP positivas (EGFP⁺) inducidas por edición de ARN mediada por LEAPER. Se transfectaron fibroblastos pulmonares primarios humanos y células epiteliales bronquiales primarias humanas con reportero-1, junto con el ARN de control 151-nt (ARN₁₅₁ Ctrl) o el ARNar de direccionamiento 151-nt (ARNar₁₅₁) seguido de análisis FACS. Las figuras 27B y 27C, resultados de secuenciación profunda que muestran la tasa de edición en los transcriptos de PPIB en fibroblastos pulmonares primarios humanos, células epiteliales bronquiales primarias humanas (b), y células T primarias humanas (c). Los datos en a, b y grupo no tratado (c) se presentan como la media $\pm$ s.e.m. (n = 3); los datos de ARN₁₅₁ Ctrl y ARNar₁₅₁ (c) se presentan como la media $\pm$ s.e.m. (n = 2).

Las figuras 28A a 28D muestran la edición dirigida mediante transducción lentiviral de ARNar y electroporación de oligonucleótidos de ARNar sintetizados. La figura 28A, cuantificación de las células EGFP⁺. Las células HEK293T que expresan de manera estable reportero-1 se infectaron con lentivirus que expresaban 151-nt de ARN Ctrl o el ARNar objetivo. Los análisis FACS se realizaron 2 días y 8 días después de la infección. Las proporciones de células EGFP⁺ se normalizaron mediante la eficiencia de transducción lentiviral (proporciones BFP⁺). La figura 28B, resultados de secuenciación profunda que muestran la tasa de edición en los transcriptos de PPIB tras la transducción lentiviral de ARNa 151-nt en células HEK293T. La figura 28C, esquema de la secuencia PPIB y el ARNar de direccionamiento 111-nt correspondiente. *(en rojo) representa un nucleótido con modificaciones de enlace 2'-O-metilo y fosforotioato. La figura 28D, resultados de secuenciación profunda que muestran la velocidad de edición en los transcriptos de PPIB tras la electroporación de oligonucleótidos de ARNar sintéticos 111-nt en células T primarias humanas.

Las figuras 29A a 29E muestran la restauración de la actividad de α -L-iduronidasa en fibroblasto primario derivado del paciente con síndrome de Hurler mediante un método LEAPER ejemplar. La figura 29A, superior, información genética de la mutación patógena en fibroblasto GM06214 derivado del paciente; medio, esquema de la secuencia de ARNm maduro de IDUA de células GM06214 (negro) que contiene una mutación homocigótica TGG>TAG en el exón 9 del gen IDUA (Trp402Ter), y el correspondiente ARNar₁₁₁-IDUA-V1 111-nt dirigido (azul); abajo, esquema de la secuencia de pre-ARNm de IDUA de células GM06214 (negro) y el ARNar₁₁₁-IDUA-V2 de direccionamiento 111-nt correspondiente (azul). *(en rojo) representa nucleótidos con modificaciones de enlace 2'-O-metilo y fosforotioato. La figura 29B, medición de la actividad catalítica de α -L-iduronidasa con sustrato de 4-metilumbeliferil α -L-iduronidasa en diferentes momentos. Los datos se presentan como la media $\pm$ s.e.m. (n = 2). La figura 29C, resultados de secuenciación profunda que muestran la tasa de edición dirigida en transcritos de IDUA en células GM06214, 48 horas después de la electroporación. La figura 29D, arriba, esquema de la secuencia del transcrito de IDUA cubierta por los ARNar 111-nt. La flecha indica la adenosina dirigida. Todas las adenosinas se marcaron en rojo. Abajo, un mapa de calor de la tasa de edición de las adenosinas cubiertas por los ARNar indicados en el transcrito de IDUA (marcado en azul en negrita). e, PCR cuantitativa que muestra las expresiones de interferón de tipo I, genes estimulados por interferón y genes proinflamatorios en la electroporación de ARNar o poli(I:C). Los datos se presentan como la media (n = 3).

Las figuras 30A a 30C muestran tres versiones de reporteros de fluorescencia dual (reportero-1, -2 y -3), mCherry y EGFP. La figura 30A, estructura del reportero-1, figura 30B, estructura del reportero-2 y figura 30C, estructura del reportero-3.

La figura 31 muestra la información del plásmido de AAV8-IDUA-ARNad151-KD2 y AAV8-Aleatorio151-KD2.

Las figuras 32A a 32B muestran la actividad enzimática relativa de α -L-iduronidasa de ratones MPSI inyectados con AAV8-IDUA-ARNad151-KD2 y AAV8-Aleatorio151-KD2. Células GM01323 como un control.

Las figuras 33A a 33B muestran la eficiencia de edición de ARN de ratones MPSI inyectados con AAV8-IDUA-ARNad151-KD2 y AAV8-Aleatorio151-KD2.

30 **DESCRIPCIÓN DETALLADA DE LA INVENCION**

La presente solicitud proporciona métodos de edición de ARN (denominados aquí métodos "LEAPER") y ARN especialmente diseñados, denominados aquí ARN de reclutamiento de desaminasas ("ARNd") o ARN de reclutamiento de ADAR ("ARNar"), para editar ARN objetivo en una célula hospedera. Sin estar ligado a ninguna teoría o hipótesis, el ARNd actúa hibridando con su ARN

objetivo de una manera específica de secuencia para formar un ARN de cadena doble, que recluta una adenosina desaminasa que actúa sobre el ARN (ADAR) para desaminar una adenosina objetivo en el ARN objetivo. Como tal, se puede lograr una edición de ARN eficiente en algunas modalidades sin sobreexpresión o ectópica de las proteínas ADAR en la célula hospedera. También se proporcionan
 5 métodos y composiciones para tratar o prevenir una enfermedad o afección en un individuo usando los métodos de edición de ARN.

Los métodos de edición de ARN descritos en este documento no usan proteínas de fusión que comprenden una ADAR y una proteína que se une específicamente a un ácido nucleico guía, como Cas. Los ARN de reclutamiento de desaminasas ("ARNd") descritos en el presente documento no
 10 comprenden ARNcr, ARNtracr o ARNg utilizados en el sistema CRISPR/Cas. En algunas modalidades, el ARNd no comprende un dominio de reclutamiento de ADAR o modificaciones químicas. En algunas modalidades, el ARNd puede expresarse a partir de un plásmido o un vector viral, o sintetizarse como un oligonucleótido, lo que podría lograr una eficiencia de edición deseable.

Los métodos LEAPER descritos en este documento tienen tasas de fuera del objetivo
 15 manejables en los transcritos objetivo y raras fuera de objetivos globales. Los inventores han utilizado el método LEAPER para restaurar la función de p53 mediante la reparación de una mutación puntual específica relevante para el cáncer. Los métodos LEAPER descritos en el presente documento también se pueden aplicar a un amplio espectro de tipos de células que incluyen múltiples células primarias humanas, y se pueden usar para restaurar la actividad catalítica de α -L-iduronidasa en fibroblastos
 20 primarios derivados de pacientes con síndrome de Hurler sin provocar respuestas inmunes innatas. En algunas modalidades, el método LEAPER implica un sistema de una sola molécula (es decir, ARNd). Los métodos LEAPER descritos en este documento permiten una edición de ARN precisa y eficiente, que ofrece un potencial transformador para la investigación básica y terapéutica.

Definiciones

La presente invención se describirá con respecto a modalidades particulares y con referencia a ciertos dibujos, pero la invención no se limita a las mismas, sino únicamente a las reivindicaciones. Cualquier signo de referencia en las reivindicaciones no se interpretará como una limitación del alcance. Cuando el término "que comprende" se usa en la presente descripción y
 25 reivindicaciones, no excluye otros elementos o pasos. Donde se usa un artículo indefinido o definido cuando se hace referencia a un sustantivo singular, por ejemplo "un" o "una", "el/la", esto incluye un plural de ese sustantivo a menos que se indique específicamente algo más. Para la enumeración de intervalos numéricos de nucleótidos en el presente documento, se contempla explícitamente cada número intermedio entre ellos. Por ejemplo, para el intervalo de 40-260 nucleótidos, se contempla cualquier número entero de nucleótidos entre 40 y 260 nucleótidos además de los números de 40
 35 nucleótidos y 260 nucleótidos.

Los siguientes términos o definiciones se proporcionan únicamente para ayudar en la comprensión de la invención. A menos que se defina específicamente en el presente documento, todos los términos utilizados en el presente documento tienen el mismo significado que tendrían para un experto en la técnica de la presente invención. Los profesionales están especialmente dirigidos a

5 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2ª ed., Cold Spring Harbor Press, Plainville, Nueva York (1989); y Ausubel et al., *Current Protocols in Molecular Biology* (Suplemento 47), John Wiley & Sons, Nueva York (1999), para definiciones y términos de la técnica. Las definiciones proporcionadas en el presente documento no deben interpretarse como que tienen un alcance menor que el comprendido por una persona con conocimientos ordinarios en la técnica.

10 Los términos "ARN de reclutamiento de desaminasa", "ARNd", "ARN de reclutamiento de ADAR" y "ARNar" se usan aquí indistintamente para referirse a un ARN diseñado capaz de reclutar una ADAR para desaminar una adenosina objetivo en un ARN.

Los términos "polinucleótido", "secuencia de nucleótidos" y "ácido nucleico" se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sea

15 desoxiribonucleótidos o ribonucleótidos, o análogos de los mismos.

Los términos "adenina", "guanina", "citocina", "timina", "uracilo" e "hipoxantina", como se usan en el presente documento, se refieren a las nucleobases como tales. Los términos "adenosina", "guanosina", "citidina", "timidina", "uridina" e "inosina" se refieren a las nucleobases unidas al radical de azúcar ribosa o desoxirribosa. El término "nucleósido" se refiere a la nucleobase

20 unida a la ribosa o desoxirribosa. El término "nucleótido" se refiere a la respectiva nucleobase-ribosil-fosfato o nucleobase-desoxirribosil-fosfato. A veces, los términos adenosina y adenina (con la abreviatura "A"), guanosina y guanina (con la abreviatura "G"), citosina y citidina (con la abreviatura "C"), uracilo y uridina (con la abreviatura, "U"), timina y timidina (con la abreviatura "T"), inosina e hipoxantina (con la abreviatura "I"), se usan indistintamente para referirse a la nucleobase,

25 nucleósido o nucleótido correspondiente. A veces, los términos nucleobase, nucleósido y nucleótido se usan indistintamente, a menos que el contexto claramente requiera lo contrario.

En el contexto de la presente solicitud, "ARN objetivo" se refiere a una secuencia de ARN a la que se diseña una secuencia de ARN de reclutamiento de desaminasas para que tenga una complementariedad perfecta o una complementariedad sustancial, y la hibridación entre la secuencia

30 objetivo y el ARNd forma un ARN de cadena doble (ARNds) que contiene una adenosina objetivo, que recluta una adenosina desaminasa que actúa sobre el ARN (ADAR) que desamina la adenosina objetivo. En algunas modalidades, la ADAR está presente de forma natural en una célula hospedera, tal como una célula eucariota (preferiblemente, una célula de mamífero, más preferiblemente, una célula humana). En algunas modalidades, la ADAR se introduce en la célula hospedera.

35 Como se usa en el presente documento, "complementariedad" se refiere a la

capacidad de un ácido nucleico para formar enlaces de hidrógeno con otro ácido nucleico mediante el acoplamiento de bases tradicional de Watson-Crick. Un porcentaje de complementariedad indica el porcentaje de residuos en una molécula de ácido nucleico que puede formar enlaces de hidrógeno (es decir, acoplamiento de bases Watson-Crick) con un segundo ácido nucleico (por ejemplo, 5 aproximadamente 5, 6, 7, 8, 9, 10 de cada 10, siendo aproximadamente un 50%, 60%, 70%, 80%, 90% y 100% complementarios respectivamente). "Perfectamente complementario" significa que todos los residuos contiguos de una secuencia de ácido nucleico forman enlaces de hidrógeno con el mismo número de residuos contiguos en una segunda secuencia de ácido nucleico. "Sustancialmente complementario", como se usa en este documento, se refiere a un grado de complementariedad que 10 es al menos aproximadamente cualquiera de 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% o 100 % sobre una región de aproximadamente 40, 50, 60, 70, 80, 100, 150, 200, 250 o más nucleótidos, o se refiere a dos ácidos nucleicos que se hibridan en condiciones rigurosas.

Como se usa en el presente documento, las "condiciones rigurosas" para la hibridación se refieren a condiciones en las que un ácido nucleico que tiene complementariedad con una secuencia 15 objetivo se hibrida predominantemente con la secuencia objetivo y sustancialmente no hibrida con secuencias no objetivo. Las condiciones rigurosas generalmente dependen de la secuencia y varían según una serie de factores. En general, cuanto más larga es la secuencia, mayor es la temperatura a la que la secuencia se hibrida específicamente con su secuencia objetivo. Los ejemplos no limitantes de condiciones rigurosas se describen en detalle en Tijssen (1993), Laboratory Techniques In Biochemistry And Molecular Biology- Hybridization With Nucleic Acid Probes Part I, segundo capítulo 20 "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y.

"Hibridación" se refiere a una reacción en la que uno o más polinucleótidos reaccionan para formar un complejo que se estabiliza mediante enlaces de hidrógeno entre las bases de los residuos de nucleótidos. El enlace de hidrógeno puede producirse mediante acoplamiento de bases de 25 Watson Crick, unión de Hoogstein o de cualquier otra manera específica de secuencia. Una secuencia capaz de hibridar con una secuencia dada se denomina "complemento" de la secuencia dada.

Como se usa en este documento, los términos "célula", "línea celular" y "cultivo celular" se usan indistintamente y todas estas designaciones incluyen progenie. Se entiende que toda la progenie puede no ser exactamente idéntica en contenido de ADN, debido a mutaciones deliberadas 30 o inadvertidas. Se incluye la progenie variante que tiene la misma función o actividad biológica que las células originales.

Métodos de edición de ARN

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula hospedera (por ejemplo, célula eucariota), que comprende introducir un ARN de 35 reclutamiento de desaminasas (ARNd) o una construcción que codifica el ARNd en la célula hospedera,

donde el ARNd comprende una secuencia de ARN complementaria que se hibrida al ARN objetivo, y en donde el ARNd es capaz de reclutar una adenosina desaminasa que actúa sobre el ARN (ADAR) para desaminar una adenosina objetivo (A) en el ARN objetivo.

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula hospedera (por ejemplo, célula eucariota), que comprende introducir un ARNd o una construcción que codifica el ARNd en la célula hospedera, donde el ARNd comprende una secuencia de ARN complementaria que hibrida al ARN objetivo, y en el que el ARNd recluta una ADAR expresada endógenamente de la célula hospedera para desaminar una objetivo A en el ARN objetivo. En algunas modalidades, el método no comprende la introducción de ninguna proteína o construcción que codifique una proteína (por ejemplo, Cas, ADAR o una proteína de fusión de ADAR y Cas) en la célula hospedera.

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir: (a) un ARNd o una construcción que codifica el ARNd, y (b) una ADAR o una construcción que codifica la ADAR en la célula hospedera, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd recluta a la ADAR para desaminar un objetivo A en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR codificada endógenamente de la célula hospedera, en donde la introducción de la ADAR comprende sobreexpresar la ADAR en la célula hospedera. En algunas modalidades, la ADAR es exógena a la célula hospedera. En algunas modalidades, la construcción que codifica ADAR es un vector, como un plásmido, o un vector viral (por ejemplo, un vector lentiviral).

En algunas modalidades, se proporciona un método para editar una pluralidad (por ejemplo, al menos aproximadamente 2, 3, 4, 5, 10, 20, 50, 100 o más) de ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir una pluralidad de ARNd o construcciones que codifican la pluralidad de ARNd en la célula hospedera, donde cada ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo correspondiente en la pluralidad de ARN objetivo, y donde cada ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo correspondiente.

En algunas modalidades, se proporciona un método para editar una pluralidad (por ejemplo, al menos aproximadamente 2, 3, 4, 5, 10, 20, 50, 100 o más) de ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir una pluralidad de ARNd o construcciones que codifican la pluralidad de ARNd en la célula hospedera, donde cada ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo correspondiente en la pluralidad de ARN objetivo, y donde cada ARNd recluta una ADAR expresada endógenamente para desaminar una objetivo A en el ARN objetivo correspondiente.

En algunas modalidades, se proporciona un método para editar una pluralidad (por ejemplo, al menos aproximadamente 2, 3, 4, 5, 10, 20, 50, 100, 1000 o más) de ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir: (a) una pluralidad de ARNd o construcciones que codifican la pluralidad de ARNd, y (b) una ADAR o una construcción que codifica ADAR en la célula hospedera, donde cada ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo correspondiente en la pluralidad de ARN objetivo, y en el que cada ARNd recluta la ADAR para desaminar un objetivo A en el ARN objetivo correspondiente.

En un aspecto, la presente solicitud proporciona un método para editar una pluralidad de ARN en células hospederas mediante la introducción de una pluralidad de ARN de reclutamiento de desaminasas, una o más construcciones que codifican los ARN de reclutamiento de desaminasas, o una biblioteca descrita en este documento, en las células hospederas.

En ciertas modalidades, el método para editar en un ARN objetivo comprende introducir múltiples ARN de reclutamiento de desaminasas o una o más construcciones que comprenden los ARN de reclutamiento de desaminasas múltiples en las células hospederas para reclutar adenosina desaminasa que actúa sobre ARN (ADAR) para realizar una reacción de desaminación en una o más adenosinas objetivo en uno o más ARN objetivo, en donde cada ARN de reclutamiento de desaminasas comprende una secuencia de ARN complementaria a un ARN objetivo correspondiente.

En un aspecto, la presente solicitud proporciona un método para generar una o más modificaciones en un ARN objetivo y/o la proteína codificada por un ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir un ARNd o una construcción que codifica el ARNd en la célula hospedera, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde una o más modificaciones se seleccionan del grupo que consiste en una mutación puntual de la proteína codificada por el ARN objetivo, plegamiento incorrecto de la proteína codificada por el ARN objetivo, un codón de terminación temprana en el ARN objetivo, un sitio de empalme aberrante en el ARN objetivo y un sitio de empalme alternativo en el ARN objetivo.

En algunas modalidades, la presente solicitud proporciona un método para restaurar la expresión o actividad de un ARN objetivo o la proteína codificada por un ARN objetivo en una célula hospedera (por ejemplo, una célula eucariota), que comprende introducir un ARNd o una construcción que codifica el ARNd en la célula hospedera, en donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, en donde el ARN objetivo tiene una mutación de G a A que da como resultado un codón de terminación temprana, un sitio de empalme aberrante o un sitio de empalme alternativo en el ARN objetivo, o plegamiento incorrecto de la proteína, y en donde el

ARNd recluta una ADAR para editar la mutación A, restaurando así la expresión o actividad del ARN objetivo o la proteína codificada por el ARN objetivo.

En determinadas modalidades, el método para generar una o más modificaciones en una pluralidad de ARN objetivo y/o las proteínas codificadas por los ARN objetivo en una célula hospedera (por ejemplo, célula eucariota), comprende introducir una pluralidad de ARN de reclutamiento de desaminasas o construcciones que codifican la pluralidad de ARN de reclutamiento de desaminasas en la célula hospedera, donde cada ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo correspondiente en la pluralidad de ARN objetivo, y donde cada ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo correspondiente.

En un aspecto, la presente solicitud proporciona el uso de un ARN de reclutamiento de desaminasas de acuerdo con cualquiera de los ARNd descritos en el presente documento para editar un ARN objetivo en una célula hospedera. En determinadas modalidades, el ARN de reclutamiento de desaminasas comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo que se va a editar.

En un aspecto, la presente solicitud proporciona el uso de un ARN de reclutamiento de desaminasas de acuerdo con cualquiera de los ARNd descritos en este documento para generar una o más modificaciones en un ARN objetivo y/o la proteína codificada por un ARN objetivo, en donde las modificaciones, una o más, se seleccionan de un grupo que consiste en una mutación puntual de la proteína codificada por el ARN objetivo, plegamiento incorrecto de la proteína codificada por el ARN objetivo, un codón de terminación temprana en el ARN objetivo, un sitio de empalme aberrante en el ARN objetivo y un sitio de empalme alternativo en el ARN objetivo. En determinadas modalidades, el ARN de reclutamiento de desaminasas comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo que se va a editar.

La invención también se refiere a un método para aprovechar una adenosina desaminasa endógena para editar un ARN objetivo en una célula eucariota, que comprende introducir un ARNd o una construcción que codifica el ARNd, como se describe en este documento, en la célula eucariota para reclutar adenosina desaminasa endógena natural que actúa sobre ARN (ADAR) para realizar una reacción de desaminación en una adenosina objetivo en la secuencia de ARN objetivo.

En ciertas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, el ARNd comprende más de aproximadamente cualquiera de 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240 o 250 nucleótidos. En ciertas modalidades, el ARNd es aproximadamente cualquiera de 40-260, 45-250, 50-240, 60-230, 65-220, 70-220, 70-210, 70-200, 70-190, 70-180, 70-170, 70-160, 70-150, 70-140,

70-130, 70-120, 70-110, 70-100, 70-90, 70-80, 75-200, 80-190, 85-180, 90-170, 95-160, 100-200, 100-150, 100-175, 110-200, 110-175, 110-150 o 105-140 nucleótidos de longitud. En algunas modalidades, el ARNd tiene una longitud de aproximadamente 71 nucleótidos. En algunas modalidades, el ARNd tiene una longitud de aproximadamente 111 nucleótidos.

5 En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, el ARNd no comprende un dominio de reclutamiento de ADAR. El "dominio de reclutamiento de ADAR" puede ser una secuencia o estructura de nucleótidos que se une con alta afinidad a ADAR, o una secuencia de nucleótidos que se une a un compañero de unión fusionado a ADAR en una construcción ADAR diseñada. Los dominios de reclutamiento de ADAR
10 ejemplares incluyen, pero no se limitan a, GluR-2, GluR-B (R/G), GluR-B (Q/R), GluR-6 (R/G), 5HT2C y el dominio FlnA (Q/R); ver, por ejemplo, Wahlstedt, Helene and Marie, "Site-selective versus promiscuous A-to-I editing." Wiley Interdisciplinary Reviews: ARN 2.6 (2011): 761-771, que se incorpora en la presente para referencia en su totalidad. En algunas modalidades, el ARNd no comprende una porción de cadena doble. En algunas modalidades, el ARNd no comprende una
15 horquilla, como el tallo bucle de MS2. En algunas modalidades, el ARNd es de cadena sencilla. En algunas modalidades, la ADAR no comprende un dominio de unión a DSB. En algunas modalidades, el ARNd consiste en (o consiste esencialmente en) la secuencia de ARN complementaria.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, el ARNd no comprende modificaciones químicas. En algunas
20 modalidades, el ARNd no comprende un nucleótido modificado químicamente, tal como 2'-O-metilnucleótido o un nucleótido que tenga un enlace fosforotioato. En algunas modalidades, el ARNd comprende modificaciones de enlace de 2'-O-metilo y fosforotioato solo en los primeros tres y últimos tres residuos. En algunas modalidades, el ARNd no es un oligonucleótido antisentido (ASO).

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, la célula hospedera es una célula procariota. En algunas
25 modalidades, la célula hospedera es una célula eucariota. En algunas modalidades, la célula hospedera es una célula de mamífero. En algunas modalidades, la célula hospedera es una célula humana. En algunas modalidades, la célula hospedera es una célula murina. En algunas modalidades, la célula hospedera es una célula vegetal o una célula fúngica.

30 En algunas modalidades según cualquiera de los métodos o usos descritos en el presente documento, la célula hospedera es una línea celular, como HEK293T, HT29, A549, HepG2, RD, SF268, SW13 y célula HeLa. En algunas modalidades, la célula hospedera es una célula primaria, como fibroblasto, célula epitelial o inmunitaria. En algunas modalidades, la célula hospedera es una célula T. En algunas modalidades, la célula hospedera es una célula posmitosis. En algunas modalidades, la
35 célula hospedera es una célula del sistema nervioso central (CNS), como una célula cerebral, por

ejemplo, una célula del cerebelo.

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula hospedera primaria (por ejemplo, una célula T o una célula del CNS) que comprende la introducción de un ARNd o una construcción que codifica el ARNd en la célula hospedera, donde el

5 ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y en donde el ARNd recluta una ADAR expresada endógenamente de la célula hospedera para desaminar un objetivo A en el ARN objetivo.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, la ADAR es endógena a la célula hospedera. En algunas

10 modalidades, la adenosina desaminasa que actúa sobre el ARN (ADAR) está presente de forma natural o endógena en la célula hospedera, por ejemplo, está presente de forma natural o endógena en la célula eucariota. En algunas modalidades, la ADAR se expresa de forma endógena por la célula hospedera. En determinadas modalidades, la ADAR se introduce de forma exógena en la célula hospedera. En algunas modalidades, ADAR es ADAR1 y/o ADAR2. En determinadas modalidades, la

15 ADAR es una o más ADAR seleccionadas del grupo que consiste en hADAR1, hADAR2, ADAR1 de ratón y ADAR2. En algunas modalidades, ADAR es ADAR1, como la isoforma p110 de ADAR1 ("ADAR1^{p110}") y/o la isoforma p150 de ADAR1 ("ADAR1^{p150}"). En algunas modalidades, ADAR es ADAR2. En algunas modalidades, la ADAR es una ADAR2 expresado por la célula hospedera, por ejemplo, ADAR2 expresado por las células del cerebelo.

En algunas modalidades, la ADAR es una ADAR exógena a la célula hospedera. En algunas modalidades, la ADAR es un mutante hiperactivo de una ADAR de origen natural. En algunas modalidades, ADAR es ADAR1 que comprende una mutación E1008Q. En algunas modalidades, ADAR no es una proteína de fusión que comprende un dominio de unión. En algunas modalidades, la ADAR no comprende un dominio de unión de ácido nucleico de cadena doble modificado por ingeniería

25 genética. En algunas modalidades, la ADAR no comprende un dominio MCP que se una a la horquilla de MS2 que se fusiona con la secuencia de ARN complementaria en el ARNd. En algunas modalidades, ADAR no comprende un dominio de unión a DSB.

En algunas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, la célula hospedera tiene un alto nivel de expresión de ADAR1 (como ADAR1^{p110} y/o

30 ADAR1^{p150}), por ejemplo, al menos aproximadamente cualquiera de 10%, 20%, 50%, 100%, 2x, 3x, 5x o más en relación con el nivel de expresión de proteína de β -tubulina. En algunas modalidades, la célula hospedera tiene un alto nivel de expresión de ADAR2, por ejemplo, al menos aproximadamente cualquiera de 10%, 20%, 50%, 100%, 2x, 3x, 5x o más en relación con el nivel de expresión de proteínas de β -tubulina. En algunas modalidades, la célula hospedera tiene un nivel de expresión bajo

35 de ADAR3, por ejemplo, no más de aproximadamente uno cualquiera de 5x, 3x, 2x, 100%, 50%, 20%

o menos con respecto al nivel de expresión de proteína de β -tubulina.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, la secuencia de ARN complementaria comprende una citidina, adenosina o uridina directamente opuesta al objetivo A en el ARN objetivo. En algunas modalidades, la secuencia de ARN complementario comprende un desajuste de citidina directamente opuesto al objetivo A en el ARN objetivo. En algunas modalidades, el desajuste de citidina se localiza al menos 5 nucleótidos, por ejemplo, al menos 10, 15, 20, 25, 30 o más nucleótidos, lejos del extremo 5' de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina se localiza al menos a 20 nucleótidos, por ejemplo, al menos a 25, 30, 35 o más nucleótidos, lejos del extremo 3' de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina no se localiza en 20 (por ejemplo, 15, 10, 5 o menos) nucleótidos del extremo 3' de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina se localiza al menos 20 nucleótidos (por ejemplo, al menos 25, 30, 35 o más nucleótidos) lejos del extremo 3' y al menos 5 nucleótidos (por ejemplo, al menos 10, 15, 20, 25, 30 o más nucleótidos) del extremo 5' de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina se localiza en el centro de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina se localiza dentro de los 20 nucleótidos (por ejemplo, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 nucleótidos) del centro de la secuencia complementaria en el ARNd.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, la secuencia de ARN complementaria comprende además una o más guanosinas, tales como 1, 2, 3, 4, 5, 6 o más G, es decir, cada una directamente opuesta a una adenosina no objetivo en el ARN objetivo. En algunas modalidades, la secuencia de ARN complementario comprende dos o más nucleótidos con desajuste consecutivos (por ejemplo, 2, 3, 4, 5 o más nucleótidos con desajuste) opuestos a una adenosina no objetivo en el ARN objetivo. En algunas modalidades, el ARN objetivo comprende no más de aproximadamente 20 As no objetivo, tal como no más de aproximadamente cualquiera de 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 A no objetivo. La Gs y los nucleótidos de desajuste consecutivos opuestos a As no objetivo pueden reducir los efectos de edición fuera del objetivo por ADAR.

En ciertas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, el vecino más cercano 5' del objetivo A es un nucleótido seleccionado entre U, C, A y G con la preferencia $U > C \approx A > G$ y el vecino más cercano 3' del objetivo A es un nucleótido seleccionado entre G, C, A y U con la preferencia $G > C > A \approx U$. En ciertas modalidades, el objetivo A está en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN objetivo. En ciertas modalidades, el motivo de tres bases es UAG, y el ARNd comprende una A directamente opuesta a la U

en el motivo de tres bases, una C directamente opuesta al objetivo A y una C, G o U directamente opuesta a la G en el motivo de tres bases. En ciertas modalidades, el motivo de tres bases es UAG en el ARN objetivo, y el ARNd comprende ACC, ACG o ACU que es opuesto a la UAG del ARN objetivo. En determinadas modalidades, el motivo de tres bases es UAG en el ARN objetivo, y el ARNd comprende ACC que está opuesta a la UAG del ARN objetivo.

En ciertas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, el ARN objetivo es cualquiera seleccionado del grupo que consiste en un ARN pre-mensajero, un ARN mensajero, un ARN ribosómico, un ARN de transferencia, un ARN largo no codificante y un ARN pequeño (por ejemplo, ARNm_i). En algunas modalidades, el ARN objetivo es un ARN pre-mensajero. En algunas modalidades, el ARN objetivo es un ARN mensajero.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, el método comprende además introducir un inhibidor de ADAR3 en la célula hospedera. En algunas modalidades, el inhibidor de ADAR3 es un ARNi contra ADAR3, tal como un ARNsh contra ADAR3 o un ARNsi contra ADAR3. En algunas modalidades, el método comprende además introducir un estimulador de interferón en la célula hospedera. En algunas modalidades, ADAR es inducible por interferón, por ejemplo, ADAR es ADAR^{p150}. En algunas modalidades, el estimulador de interferón es IFN α . En algunas modalidades, el inhibidor de ADAR3 y/o el estimulador de interferón están codificados por la misma construcción (por ejemplo, vector) que codifica el ARNd.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, la eficiencia de edición del ARN objetivo es de al menos aproximadamente el 5%, tal como al menos aproximadamente cualquiera de 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65% o más. En algunas modalidades, la eficacia de edición del ARN objetivo es de al menos aproximadamente un 5% (por ejemplo, al menos aproximadamente un 7%) *in vivo*, por ejemplo, en un animal. En algunas modalidades, la eficacia de la edición del ARN objetivo es de al menos aproximadamente el 10% (por ejemplo, al menos aproximadamente el 15%) en una célula *in vitro* o *ex vivo*. En algunas modalidades, la eficiencia de la edición está determinada por la secuenciación de Sanger. En algunas modalidades, la eficacia de la edición se determina mediante la secuenciación de próxima generación.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, el método tiene una baja tasa de edición fuera del objetivo. En algunas modalidades, el método tiene menos de aproximadamente 1% (por ejemplo, no más de aproximadamente cualquiera de 0.5%, 0.1%, 0.05%, 0.01%, 0.001% o menos) de eficiencia de edición en As no objetivo en el ARN objetivo. En algunas modalidades, el método no edita el As no objetivo en el ARN objetivo. En algunas modalidades, el método tiene menos de aproximadamente

0.1% (por ejemplo, no más de aproximadamente cualquiera de 0.05%, 0.01%, 0.005%, 0.001%, 0.0001% o menos) de eficiencia de edición en As en ARN no objetivo.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en el presente documento, el método no induce una respuesta inmune, tal como una respuesta inmune innata. En algunas modalidades, el método no induce la expresión de interferón y/o interleucina en la célula hospedera. En algunas modalidades, el método no induce la expresión de IFN- β y/o IL-6 en la célula hospedera.

También se proporcionan ARN editado o células hospederas que tienen un ARN editado producido por cualquiera de los métodos descritos en el presente documento. En algunas modalidades, el ARN editado comprende una inosina. En algunas modalidades, la célula hospedera comprende un ARN que tiene una mutación sin sentido, un codón de terminación temprano, un sitio de empalme alternativo o un sitio de empalme aberrante. En algunas modalidades, la célula hospedera comprende una proteína mutante, truncada o mal plegada.

"Célula hospedera", como se describe en el presente documento, se refiere a cualquier tipo de célula que pueda usarse como célula hospedera siempre que pueda modificarse como se describe en el presente documento. Por ejemplo, la célula hospedera puede ser una célula hospedera con adenosina desaminasa expresada de forma endógena que actúa sobre el ARN (ADAR), o puede ser una célula hospedera en la que se introduce una adenosina desaminasa que actúa sobre el ARN (ADAR) mediante un método conocido en la técnica. Por ejemplo, la célula hospedera puede ser una célula procariota, una célula eucariota o una célula vegetal. En algunas modalidades, la célula hospedera se deriva de una línea celular preestablecida, como líneas celulares de mamíferos que incluyen líneas celulares humanas o líneas celulares no humanas. En algunas modalidades, la célula hospedera se deriva de un individuo, como un individuo humano.

"Que introduce" o "introducción" utilizado en el presente documento significa entregar uno o más polinucleótidos, tales como ARNd o una o más construcciones que incluyen vectores como se describe en el presente documento, uno o más transcritos de los mismos, a una célula hospedera. La invención sirve como una plataforma básica para permitir la edición dirigida de ARN, por ejemplo, ARN pre-mensajero, ARN mensajero, ARN ribosómico, ARN de transferencia, ARN largo no codificante y ARN pequeño (como ARNm). Los métodos de la presente solicitud pueden emplear muchos sistemas de administración, que incluyen, pero no se limitan a, virus, liposomas, electroporación, microinyección y conjugación, para lograr la introducción del ARNd o construcción como se describe en este documento en una célula hospedera. Pueden usarse métodos convencionales de transferencia de genes basados en virus y no virus para introducir ácidos nucleicos en células de mamíferos o tejidos objetivo. Dichos métodos se pueden usar para administrar ácidos nucleicos que codifican el ARNd de la presente solicitud a células en cultivo o en un organismo hospedero. Los sistemas de administración de

vectores no virales incluyen plásmidos de ADN, ARN (por ejemplo, un transcripto de una construcción descrita en este documento), ácido nucleico desnudo y ácido nucleico complejado con un vehículo de administración, como un liposoma. Los sistemas de administración de vectores virales incluyen virus de ADN y ARN, que tienen genomas episomales o integrados para su administración a la célula

5 hospedera.

Los métodos de suministro no viral de ácidos nucleicos incluyen lipofección, nucleofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, polimerización o lípido: conjugados de ácido nucleico, electroporación, nanopartículas, exosomas, microvesículas o pistola de genes, ADN desnudo y viriones artificiales.

10 El uso de sistemas basados en virus de ARN o ADN para la administración de ácidos nucleicos tiene una alta eficiencia para dirigir un virus a células específicas y transportar la carga útil viral a los núcleos celulares.

En determinadas modalidades de acuerdo con cualquiera de los métodos o usos descritos en este documento, el método comprende introducir un vector viral (tal como un vector

15 lentivírico) que codifica el ARNd en la célula hospedera. En algunas modalidades, el vector viral es un AAV, por ejemplo, AAV8. En algunas modalidades, el método comprende introducir un plásmido que codifica el ARNd en la célula hospedera. En algunas modalidades, el método comprende introducir (por ejemplo, mediante electroporación) el ARNd (por ejemplo, ARNd sintético) en la célula hospedera. En algunas modalidades, el método comprende la transfección del ARNd en la célula hospedera.

20 Después de la desaminación, se puede determinar la modificación del ARN objetivo y/o la proteína codificada por el ARN objetivo usando diferentes métodos dependiendo de las posiciones de las adenosinas objetivo en el ARN objetivo. Por ejemplo, para determinar si "A" se ha editado en "I" en el ARN objetivo, se pueden usar métodos de secuenciación de ARN conocidos en la técnica para detectar la modificación de la secuencia de ARN. Cuando la adenosina objetivo se

25 encuentra en la región codificante de un ARNm, la edición del ARN puede provocar cambios en la secuencia de aminoácidos codificada por el ARNm. Por ejemplo, pueden introducirse mutaciones puntuales en el ARNm de una mutación puntual innata o adquirida en el ARNm que pueden revertirse para producir productos génicos de tipo salvaje debido a la conversión de "A" en "I". La secuenciación de aminoácidos por métodos conocidos en la técnica puede usarse para encontrar cualquier cambio de

30 residuos de aminoácidos en la proteína codificada. Las modificaciones de un codón de terminación se pueden determinar evaluando la presencia de una proteína funcional, alargada, trunca, de longitud completa y/o de tipo salvaje. Por ejemplo, cuando la adenosina objetivo se encuentra en un codón de terminación UGA, UAG o UAA, la modificación del objetivo A (UGA o UAG) o As (UAA) puede crear una mutación de lectura completa y/o una proteína alargada, o una proteína trunca codificada por el

35 ARN objetivo puede invertirse para crear una proteína funcional, de longitud completa y/o de tipo

salvaje. La edición de un ARN objetivo también puede generar un sitio de empalme aberrante, y/o un sitio de empalme alternativo en el ARN objetivo, lo que conduce a una proteína alargada, truncada o mal plegada, o un empalme aberrante o un sitio de empalme alternativo codificado en el ARN objetivo puede invertirse para crear una proteína funcional, de plegado correcto, de longitud completa y/o de tipo salvaje. En algunas modalidades, la presente solicitud contempla la edición de cambios genéticos tanto innatos como adquiridos, por ejemplo, mutación sin sentido, codón de terminación temprano, empalme aberrante o sitio de empalme alternativo codificado por un ARN objetivo. El uso de métodos conocidos para evaluar la función de la proteína codificada por el ARN objetivo puede averiguar si la edición del ARN logra los efectos deseados. Debido a que la desaminación de la adenosina (A) a una inosina (I) puede corregir una A mutada en la posición objetivo en un ARN mutante que codifica una proteína, la identificación de la desaminación en inosina puede proporcionar una evaluación de si una proteína funcional está presente, o si un ARN asociado a una enfermedad o resistencia a fármacos causado por la presencia de una adenosina mutada se invierte o se invierte parcialmente. De manera similar, debido a que la desaminación de la adenosina (A) a una inosina (I) puede introducir una mutación puntual en la proteína resultante, la identificación de la desaminación en inosina puede proporcionar una indicación funcional para identificar una causa de enfermedad o un factor relevante de una enfermedad.

Cuando la presencia de una adenosina objetivo provoca un empalme aberrante, la lectura puede ser la evaluación de la aparición y la frecuencia de empalmes aberrantes. Por otro lado, cuando es deseable la desaminación de una adenosina objetivo para introducir un sitio de empalme, se pueden utilizar enfoques similares para comprobar si se produce el tipo de empalme requerido. Un método adecuado ejemplar para identificar la presencia de una inosina después de la desaminación de la adenosina objetivo es la RT-PCR y la secuenciación, usando métodos que son bien conocidos por los expertos en la técnica.

Los efectos de la desaminación de la o las adenosinas objetivo incluyen, por ejemplo, mutación puntual, codón de terminación temprano, sitio de empalme aberrante, sitio de empalme alternativo y plegamiento incorrecto de la proteína resultante. Estos efectos pueden inducir cambios estructurales y funcionales de ARN y/o proteínas asociadas con enfermedades, ya sean heredadas genéticamente o causadas por mutaciones genéticas adquiridas, o pueden inducir cambios estructurales y funcionales de ARN y/o proteínas asociados con la aparición de resistencia a fármacos. Por lo tanto, los ARNd, las construcciones que codifican los ARNd y los métodos de edición de ARN de la presente solicitud pueden usarse en la prevención o el tratamiento de enfermedades o afecciones genéticas hereditarias, o enfermedades o afecciones asociadas con mutaciones genéticas adquiridas cambiando la estructura y/o función de los ARN y/o proteínas asociadas a la enfermedad.

En algunas modalidades, el ARN objetivo es un ARN regulador. En algunas

modalidades, el ARN objetivo que se va a editar es un ARN ribosómico, un ARN de transferencia, un ARN largo no codificante o un ARN pequeño (por ejemplo, ARNmi, ARNpre-mi, ARNpi, ARNsi, ARNsno, ARNsn, ARNex o ARNsca). Los efectos de la desaminación de las adenosinas objetivo incluyen, por ejemplo, cambios estructurales y funcionales del ARN ribosómico, ARN de transferencia, ARN largo no codificante o ARN pequeño (por ejemplo, ARNmi), incluidos cambios de estructura tridimensional y/o pérdida de la función o ganancia de la función del ARN objetivo. En algunas modalidades, la desaminación del As objetivo en el ARN objetivo cambia el nivel de expresión de una o más moléculas corriente abajo (por ejemplo, proteína, ARN y/o metabolitos) del ARN objetivo. Los cambios del nivel de expresión de las moléculas corriente abajo pueden aumentar o disminuir en el nivel de expresión.

Algunas modalidades de la presente solicitud implican la edición multiplex del ARN objetivo en células hospederas, que son útiles para seleccionar diferentes variantes de un gen objetivo o diferentes genes en las células hospederas. En algunas modalidades, en las que el método comprende introducir una pluralidad de ARNd en las células hospederas, al menos dos de los ARNd de la pluralidad de ARNd tienen diferentes secuencias y/o tienen diferentes ARN objetivo. En algunas modalidades, cada ARNd tiene una secuencia diferente y/o un ARN objetivo diferente. En algunas modalidades, el método genera una pluralidad (por ejemplo, al menos 2, 3, 5, 10, 50, 100, 1000 o más) de modificaciones en un único ARN objetivo en las células hospederas. En algunas modalidades, el método genera una modificación en una pluralidad (por ejemplo, al menos 2, 3, 5, 10, 50, 100, 1000 o más) de ARN objetivo en las células hospederas. En algunas modalidades, el método comprende editar una pluralidad de ARN objetivo en una pluralidad de poblaciones de células hospederas. En algunas modalidades, cada población de células hospederas recibe un ARNd diferente o un ARNd que tiene un ARN objetivo diferente de las otras poblaciones de células hospederas.

ARN, construcción y biblioteca de reclutamiento de desaminasas

En un aspecto, la presente solicitud proporciona un ARN de reclutamiento de desaminasas útil para cualquiera de los métodos descritos en el presente documento. Cualquiera de los ARNd descritos en esta sección puede usarse en los métodos de edición y tratamiento de ARN descritos en este documento. Se pretende que cualquiera de las características y parámetros descritos en este documento para los ARNd se puedan combinar entre sí, como si todas y cada una de las combinaciones se describieran individualmente. Los ARNd descritos en el presente documento no comprenden un ARNtracr, ARNcr o ARNg usados en un sistema CRISPR/Cas.

En algunas modalidades, se proporciona un ARN de reclutamiento de desaminasa (ARNd) para la desaminación de una adenosina objetivo en un ARN objetivo mediante el reclutamiento de una ADAR, que comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo.

En un aspecto, el presente proporciona una construcción que comprende cualquiera de los ARN de reclutamiento de desaminasas descritos en el presente documento. En determinadas modalidades, la construcción es un vector viral (como un vector lentivirus) o un plásmido. En algunas modalidades, el vector viral es un AAV, como AAV8. En algunas modalidades, la construcción codifica un solo ARNd. En algunas modalidades, la construcción codifica una pluralidad (por ejemplo, aproximadamente cualquiera de 1, 2, 3, 4, 5, 10, 20 o más) ARNd.

En un aspecto, la presente solicitud proporciona una biblioteca que comprende una pluralidad de ARN de reclutamiento de desaminasas o una pluralidad de las construcciones descritas en el presente documento.

En un aspecto, la presente solicitud proporciona una composición o una célula hospedera que comprende el ARN de reclutamiento de desaminasas o la construcción descrita en el presente documento. En determinadas modalidades, la célula hospedera es una célula procariota o una célula eucariota. Preferiblemente, la célula hospedera es una célula de mamífero. Lo más preferiblemente, la célula hospedera es una célula humana.

En determinadas modalidades según cualquiera de los ARNd, construcciones, bibliotecas o composiciones descritas en el presente documento, la secuencia de ARN complementaria comprende una citidina, adenosina o uridina directamente opuesta a la adenosina objetivo que se va a editar en el ARN objetivo. En determinadas modalidades, la secuencia de ARN complementario comprende además una o más guanosinas que están cada una directamente opuesta a una adenosina no objetivo en el ARN objetivo. En ciertas modalidades, el vecino más cercano 5' del objetivo A es un nucleótido seleccionado entre U, C, A y G con preferencia $U > C \approx A > G$ y el vecino más cercano 3' del objetivo A es un nucleótido seleccionado entre G, C, A y U con preferencia $G > C > A \approx U$. En algunas modalidades, el vecino más cercano 5' del objetivo A es U. En algunas modalidades, el vecino más cercano 5' del objetivo A es C o A. En algunas modalidades, el vecino más cercano 3' del objetivo A es G. En algunas modalidades, el vecino más cercano 3' del objetivo A es C.

En ciertas modalidades de acuerdo con cualquiera de los ARNd, construcciones, bibliotecas o composiciones descritas en este documento, el objetivo A está en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN objetivo. En ciertas modalidades, el motivo de tres bases es UAG, y el ARNd comprende una A directamente opuesta a la U en el motivo de tres bases, una C directamente opuesta al objetivo A y una C, G o U directamente opuesta a la G en el motivo de tres bases. En ciertas modalidades, el motivo de tres bases es UAG en el ARN objetivo, y el ARNd comprende ACC, ACG o ACU que es opuesto a la UAG del ARN objetivo.

En algunas modalidades, el ARNd comprende un desajuste de citidina directamente opuesto al objetivo A en el ARN objetivo. En algunas modalidades, el desajuste de citidina está cerca

del centro de la secuencia de ARN complementario, tal como a 20, 15, 10, 5, 4, 3, 2 o 1 nucleótido del centro de la secuencia de ARN complementario. En algunas modalidades, el desajuste de citidina está al menos a 5 nucleótidos del extremo 5' de la secuencia de ARN complementaria. En algunas modalidades, el desajuste de citidina está al menos a 20 nucleótidos del extremo 3' de la secuencia de ARN complementaria.

En determinadas modalidades de acuerdo con cualquiera de los ARNd, construcciones, bibliotecas o composiciones descritas en el presente documento, el ARNd comprende más de aproximadamente 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240 o 250 nucleótidos. En ciertas modalidades, el ARNd es aproximadamente cualquiera de 40-260, 45-250, 50-240, 60-230, 65-220, 70-210, 70-200, 70-190, 70-180, 70-170, 70-160, 70-150, 70-140, 70-130, 70-120, 70-110, 70-100, 70-90, 70-80, 75-200, 80-190, 85-180, 90-170, 95-160, 100-150 o 105-140 nucleótidos de longitud.

En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de cualquiera de las SEQ ID NO: 25-44, 142-205, 341-342.

El ARNd de la presente solicitud comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo. La secuencia de ARN complementaria es perfectamente complementaria o sustancialmente complementaria al ARN objetivo para permitir la hibridación de la secuencia de ARN complementaria al ARN objetivo. En algunas modalidades, la secuencia de ARN complementario tiene una complementariedad de secuencia del 100% como ARN objetivo. En algunas modalidades, la secuencia de ARN complementaria es al menos aproximadamente cualquiera de 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% o más complementaria a más de un tramo continuo de al menos aproximadamente cualquiera de 20, 40, 60, 80, 100, 150, 200 o más nucleótidos en el ARN objetivo. En algunas modalidades, el ARNd formado por hibridación entre la secuencia de ARN complementaria y el ARN objetivo tiene uno o más (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más) pares de base que no son Watson-Crick (es decir, desajustes).

ADAR, por ejemplo, las enzimas ADAR humanas editan estructuras de ARN de cadena doble (ARNds) con especificidad variable, dependiendo de varios factores. Un factor importante es el grado de complementariedad de las dos cadenas que forman la secuencia de ARNds. La complementariedad perfecta entre el ARNd y el ARN objetivo generalmente hace que el dominio catalítico de ADAR desamine las adenosinas de una manera no discriminativa. La especificidad y la eficacia de ADAR se pueden modificar introduciendo desajustes en la región de ARNds. Por ejemplo, se recomienda preferiblemente el desajuste A-C para aumentar la especificidad y la eficacia de la desaminación de la adenosina que se va a editar. Por el contrario, en las otras posiciones A (adenosina) distintas del objetivo A (es decir, "A no objetivo"), el desajuste de G-A puede reducir la edición fuera del objetivo. No se requiere necesariamente una complementariedad perfecta para la formación de ARNds

entre el ARNd y su ARN objetivo, siempre que exista una complementariedad sustancial para la hibridación y formación del ARNd entre el ARNd y el ARN objetivo. En algunas modalidades, la secuencia de ARNd o región de ARN de cadena sencilla de la misma tiene al menos aproximadamente cualquiera de 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% de complementariedad de secuencia con el ARN objetivo, cuando se alinea de manera óptima. La alineación óptima se puede determinar con el uso de cualquier algoritmo adecuado para alinear secuencias, cuyos ejemplos no limitantes incluyen el algoritmo Smith-Waterman, el algoritmo Needleman-Wimsch, algoritmos basados en la transformada de Burrows-Wheeler (por ejemplo, el alineador de Burrows Wheeler).

Los nucleótidos vecinos de la adenosina objetivo también afectan la especificidad y eficacia de la desaminación. Por ejemplo, el vecino más cercano 5' de la adenosina objetivo que se va a editar en la secuencia de ARN objetivo tiene la preferencia $U > C \approx A > G$ y el vecino más cercano 3' de la adenosina objetivo que se va a editar en la secuencia de ARN objetivo tiene la preferencia $G > C > A \approx U$ en términos de especificidad y eficiencia de desaminación de adenosina. En algunas modalidades, cuando la adenosina objetivo puede estar en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN objetivo, la especificidad y eficiencia de la desaminación de la adenosina son más altas que las adenosinas en otros motivos de tres bases. En algunas modalidades, donde la adenosina objetivo que se va a editar está en el motivo de tres bases UAG, UAC, UAA, UAU, CAG, CAC, AAG, AAC o AAA, la eficacia de la desaminación de la adenosina es mucho mayor que la de las adenosinas en otros motivos. Con respecto al mismo motivo de tres bases, diferentes diseños de ARNd también pueden conducir a diferente eficiencia de desaminación. Tomando el motivo de tres bases UAG como ejemplo, en algunas modalidades, cuando el ARNd comprende citidina (C) directamente opuesta a la adenosina objetivo que se va a editar, adenosina (A) directamente opuesta a la uridina, y citidina (C), guanosina (G) o uridina (U) directamente opuesta a la guanosina, la eficiencia de desaminación de la adenosina objetivo es mayor que la que se utiliza con otras secuencias de ARNd. En algunas modalidades, cuando el ARNd comprende ACC, ACG o ACU opuesto a UAG del ARN objetivo, la eficiencia de edición de la A en la UAG del ARN objetivo puede alcanzar aproximadamente el 25% al 30%.

Además de las adenosinas objetivo, puede haber una o más adenosinas en el ARN objetivo, que no es deseable editar. Con respecto a estas adenosinas, es preferible reducir su eficiencia de edición tanto como sea posible. Se encuentra mediante esta invención que cuando la guanosina está directamente opuesta a una adenosina en el ARN objetivo, la eficacia de desaminación disminuye significativamente. Por lo tanto, para disminuir la desaminación fuera del objetivo, los ARNd pueden diseñarse para que comprendan una o más guanosinas directamente opuestas a una o más adenosinas distintas de la adenosina objetivo que se va a editar en el ARN objetivo.

El nivel deseado de especificidad y eficiencia de la edición de la secuencia de ARN objetivo puede depender de diferentes aplicaciones. Siguiendo las instrucciones de la presente solicitud de patente, los expertos en la técnica serán capaces de diseñar un ARNd que tenga una secuencia complementaria o sustancialmente complementaria a la secuencia de ARN objetivo de acuerdo con sus necesidades y, con algo de prueba y error, obtener los resultados deseados. Como se usa en este documento, el término "desajuste" se refiere a nucleótidos opuestos en un ARN de cadena doble (ARNds) que no forman pares de bases perfectos de acuerdo con las reglas de acoplamiento de bases de Watson-Crick. Los pares de bases que no coinciden incluyen, por ejemplo, pares de bases G-A, C-A, U-C, A-A, G-G, C-C, U-U. Tomando la coincidencia de A-C como ejemplo, donde se va a editar un objetivo A en el ARN objetivo, se diseña un ARNd para que comprenda una C opuesta a la A que se va a editar, generando un desajuste de A-C en el ARNds formado por hibridación entre el ARN objetivo y ARNd.

En algunas modalidades, el ARNds formado por hibridación entre el ARNd y el ARN objetivo no comprende un desajuste. En algunas modalidades, el ARNds formado por hibridación entre el ARNd y el ARN objetivo comprende uno o más, tal como cualquiera de 1, 2, 3, 4, 5, 6, 7 o más desajustes (por ejemplo, el mismo tipo de diferentes tipos de desajustes). En algunas modalidades, el ARNds formado por hibridación entre el ARNd y el ARN objetivo comprende uno o más tipos de desajustes, por ejemplo, 1, 2, 3, 4, 5, 6, 7 tipos de desajustes seleccionados del grupo que consiste en G-A, C-A, U-C, A-A, G-G, C-C y U-U.

Los nucleótidos con desajuste en el ARNds formados por hibridación entre el ARNd y el ARN objetivo pueden formar protuberancias, que pueden promover la eficiencia de edición del ARN objetivo. Puede haber una (que solo se forma en la adenosina objetivo) o más protuberancias formadas por los desajustes. Los desajustes adicionales que inducen protuberancias pueden ser corriente arriba y/o corriente abajo de la adenosina objetivo. Las protuberancias pueden ser protuberancias de un solo desajuste (causadas por un par de bases de desajuste) o abultamientos de múltiples desajustes (causadas por más de un par de bases de desajuste consecutivas, preferiblemente dos o tres pares de bases de desajuste consecutivos).

La secuencia de ARN complementaria en el ARNd es de cadena sencilla. El ARNd puede ser completamente de cadena sencilla o tener una o más (por ejemplo, 1, 2, 3 o más) regiones de cadena doble y/o una o más regiones de tallo bucle. En algunas modalidades, la secuencia de ARN complementaria es al menos aproximadamente cualquiera de 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 o más nucleótidos. En ciertas modalidades, la secuencia de ARN complementaria es aproximadamente cualquiera de 40-260, 45-250, 50-240, 60-230, 65-220, 70-220, 70-210, 70-200, 70-190, 70-180, 70-170, 70-160, 70-150, 70-140, 70-130, 70-120, 70-110, 70-100, 70-90, 70-80, 75-200, 80-190, 85-180, 90-170, 95-160, 100-200, 100-150,

100-175, 110-200, 110-175, 110-150 o 105-140 nucleótidos de longitud. En algunas modalidades, la secuencia de ARN complementaria tiene una longitud de aproximadamente 71 nucleótidos. En algunas modalidades, la secuencia de ARN complementaria tiene una longitud de aproximadamente 111 nucleótidos.

5 En algunas modalidades, el ARNd, además de la secuencia de ARN complementaria, comprende además regiones para estabilizar el ARNd, por ejemplo, una o más regiones de cadena doble y/o regiones de tallo bucle. En algunas modalidades, la región de cadena doble o la región del tallo bucle del ARNd comprende no más de aproximadamente cualquiera de 200, 150, 100, 50, 40, 30, 20, 10 o menos pares de bases. En algunas modalidades, el ARNd no comprende un tallo bucle o una
10 región de cadena doble. En algunas modalidades, el ARNd comprende un dominio de reclutamiento de ADAR. En algunas modalidades, el ARNd no comprende un dominio de reclutamiento de ADAR.

 El ARNd puede comprender una o más modificaciones. En algunas modalidades, el ARNd tiene uno o más nucleótidos modificados, incluida la modificación de la nucleobase y/o la modificación de la estructural fundamental. Las modificaciones ejemplares del ARN incluyen, pero no
15 se limitan a, modificación de la estructura fundamental de fosforotioato, sustituciones 2' en la ribosa (tales como sustituciones 2'-O-metilo y 2'-fluoro), LNA y L-ARN. En algunas modalidades, el ARNd no tiene modificaciones en la nucleobase o la estructura fundamental.

 La presente solicitud también contempla una construcción que comprende el ARNd descrito en el presente documento. El término "construcción", como se usa en este documento, se
20 refiere a moléculas de ADN o ARN que comprenden una secuencia de nucleótidos codificante que puede transcribirse en ARN o expresarse en proteínas. En algunas modalidades, la construcción contiene uno o más elementos reguladores operativamente unidos a la secuencia de nucleótidos que codifica el ARN o la proteína. Cuando la construcción se introduce en una célula hospedera, en condiciones adecuadas, la secuencia de nucleótidos codificante en la construcción se puede transcribir
25 o expresar.

 En algunas modalidades, la construcción comprende un promotor que está operativamente enlazado o conectado espacialmente a la secuencia de nucleótidos codificante, de modo que el promotor controla la transcripción o expresión de la secuencia de nucleótidos codificante. Un promotor puede colocarse 5' (corriente arriba) de una secuencia de nucleótidos codificante bajo su
30 control. La distancia entre el promotor y la secuencia codificante puede ser aproximadamente la misma que la distancia entre ese promotor y el gen que controla en el gen del que se deriva el promotor. Como se sabe en la técnica, la variación en esta distancia se puede acomodar sin la pérdida de la función del promotor. En algunas modalidades, la construcción comprende una UTR 5' y/o una UTR 3' que regula la transcripción o expresión de la secuencia de nucleótidos codificante.

35 En algunas modalidades, la construcción es un vector que codifica cualquiera de los

ARNd descritos en la presente solicitud. El término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido. Los vectores incluyen, pero no se limitan a, moléculas de ácido nucleico que son de cadena sencilla, de cadena doble o parcialmente de cadena doble; moléculas de ácido nucleico que comprenden uno o más extremos libres, sin extremos libres
5 (por ejemplo, circulares); moléculas de ácido nucleico que comprenden ADN, ARN o ambos; y otras variedades de polinucleótidos conocidas en la técnica. Un tipo de vector es un "plásmido", que se refiere a un bucle circular de ADN de doble cadena en el que se pueden insertar segmentos de ADN adicionales, como mediante técnicas estándar de clonación molecular. Ciertos vectores son capaces de replicación autónoma en una célula hospedera en la que se introducen (por ejemplo, vectores
10 bacterianos que tienen un origen bacteriano de replicación y vectores de mamíferos episomales). Otros vectores (por ejemplo, vectores de mamíferos no episomales) se pueden integrar en el genoma de una célula hospedera tras la introducción en la célula hospedera, y por lo tanto se replicaron junto con el genoma hospedero. Además, ciertos vectores son capaces de dirigir la transcripción o expresión de secuencias de nucleótidos codificantes a las que están operativamente enlazados. Dichos vectores se
15 denominan en el presente documento "vectores de expresión".

Los vectores de expresión recombinantes pueden comprender un ácido nucleico de la invención en una forma adecuada para la transcripción o expresión del ácido nucleico en una célula hospedera. En algunas modalidades, el vector de expresión recombinante incluye uno o más
20 elementos reguladores, que pueden seleccionarse sobre la base de las células hospederas que se utilizarán para la transcripción o expresión, que están operativamente enlazadas a la secuencia de ácido nucleico que se va a transcribir o expresar. Dentro de un vector de expresión recombinante, "enlazada operativamente" pretende dar a entender que la secuencia de nucleótidos de interés está enlazada al elemento o elementos reguladores en una manera que permite la expresión de la secuencia de nucleótidos (por ejemplo en un sistema de transcripción/traducción in vitro o en una
25 célula hospedera cuando el vector se introduce dentro de la célula hospedera).

En algunas modalidades, se proporciona una construcción (por ejemplo, un vector, tal como un vector viral) que comprende una secuencia de nucleótidos que codifica el ARNd. En algunas modalidades, se proporciona una construcción (por ejemplo, un vector, tal como un vector viral) que comprende una secuencia de nucleótidos que codifica ADAR. En algunas modalidades, se proporciona
30 una construcción que comprende una primera secuencia de nucleótidos que codifica el ARNd y una segunda secuencia de nucleótidos que codifica ADAR. En algunas modalidades, la primera secuencia de nucleótidos y la segunda secuencia de nucleótidos están unidas operativamente al mismo promotor. En algunas modalidades, la primera secuencia de nucleótidos y la segunda secuencia de nucleótidos están unidas operativamente a diferentes promotores. En algunas modalidades, el promotor es
35 inducible. En algunas modalidades, la construcción no codifica la ADAR. En algunas modalidades, el

vector comprende además secuencia o secuencias de ácido nucleico que codifican un inhibidor de ADAR3 (por ejemplo, ARNsh o ARNsi de ADAR3) y/o un estimulador de interferón (por ejemplo, IFN- α). En algunas modalidades, la construcción es un AAV, como AAV8.

Métodos de tratamiento

5 Los métodos y composiciones de edición de ARN descritos en el presente documento pueden usarse para tratar o prevenir una enfermedad o afección en un individuo, incluidas, entre otras, enfermedades genéticas hereditarias y resistencia a fármacos.

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula de un individuo (por ejemplo, un individuo humano) *ex vivo*, que comprende editar el ARN objetivo usando cualquiera de los métodos de edición de ARN descritos en este documento.

En algunas modalidades, se proporciona un método para editar un ARN objetivo en una célula de un individuo (por ejemplo, un individuo humano) *ex vivo*, que comprende introducir un ARNd o una construcción que codifica el ARNd en la célula del individuo, en donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y en donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo. En algunas modalidades, el ARN objetivo está asociado con una enfermedad o afección del individuo. En algunas modalidades, la enfermedad o afección es una enfermedad genética hereditaria o una enfermedad o afección asociada con una o más mutaciones genéticas adquiridas (por ejemplo, resistencia a fármacos). En algunas modalidades, el método comprende además obtener la célula del individuo.

En algunas modalidades, se proporciona un método para restaurar la expresión o actividad de un ARN o proteína objetivo codificada por el ARN objetivo en una célula de un individuo que tiene una enfermedad o afección, que comprende editar el ARN objetivo utilizando cualquiera de los métodos de edición de ARN descrita en este documento, en donde el ARN objetivo tiene una mutación de G a A asociada con la enfermedad o afección. En algunas modalidades, el método se lleva a cabo en una célula aislada de un individuo *ex vivo*. En algunas modalidades, el método se lleva a cabo *in vivo*.

En algunas modalidades, se proporciona un método para tratar o prevenir una enfermedad o afección en un individuo (por ejemplo, un individuo humano), que comprende editar un ARN objetivo asociado con la enfermedad o afección en una célula del individuo usando cualquiera de los métodos de la edición del ARN descrita en este documento.

En algunas modalidades, se proporciona un método para tratar o prevenir una enfermedad o afección en un individuo (por ejemplo, un individuo humano), que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo asociado con la enfermedad o afección, y en donde el ARNd es capaz de reclutar una ADAR para desaminar un

objetivo A en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el método comprende además cultivar la célula que tiene el ARN editado. En algunas modalidades, el método

5 comprende además administrar la célula que tiene el ARN editado al individuo. En algunas modalidades, la enfermedad o afección es una enfermedad genética hereditaria o una enfermedad o afección asociada con una o más mutaciones genéticas adquiridas (por ejemplo, resistencia a fármacos).

En algunas modalidades, se proporciona un método para tratar o prevenir una

10 enfermedad o afección en un individuo (por ejemplo, un individuo humano), que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo asociado con la enfermedad o afección, y en donde el ARNd es capaz de reclutar una ADAR expresada endógenamente de la célula hospedera para desaminar un objetivo A en el ARN objetivo. En algunas

15 modalidades, el método comprende además cultivar la célula que tiene el ARN editado. En algunas modalidades, el método comprende además administrar la célula que tiene el ARN editado al individuo. En algunas modalidades, la enfermedad o afección es una enfermedad genética hereditaria o una enfermedad o afección asociada con una o más mutaciones genéticas adquiridas (por ejemplo, resistencia a fármacos).

En algunas modalidades, se proporciona un método para tratar o prevenir una

20 enfermedad o afección en un individuo (por ejemplo, un individuo humano), que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd al individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con un ARN objetivo asociado con la enfermedad o afección, y en donde el ARNd es capaz de reclutar una ADAR para

25 desaminar un objetivo A en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, la enfermedad o afección es una enfermedad genética hereditaria o una enfermedad o afección asociada con una o más mutaciones genéticas adquiridas (por ejemplo, resistencia a fármacos).

Las enfermedades y afecciones adecuadas para el tratamiento utilizando los métodos

30 de la presente solicitud incluyen enfermedades asociadas con una mutación, tales como una mutación de G a A, por ejemplo, una mutación de G a A que da como resultado una mutación sin sentido, codón de terminación temprano, empalme aberrante o empalme alternativo en un transcripto de ARN. Ejemplos de mutaciones asociadas a enfermedades que pueden restaurarse mediante los métodos de

35 la presente solicitud incluyen, pero no se limitan a, TP53^{W53X} (por ejemplo, 158G> A) asociado con

cáncer, IDUA^{W402X} (por ejemplo, mutación TGG> TAG en el exón 9) asociado con mucopolisacaridosis tipo I (MPS I), COL3A1^{W1278X} (por ejemplo, mutación 3833G> A) asociado con el síndrome de Ehlers-Danlos, BMPR2^{W298X} (por ejemplo, 893G> A) asociado con hipertensión pulmonar primaria, AHI1^{W725X} (por ejemplo, 2174G> A) asociado con el síndrome Joubert, FANCC^{W506X} (por ejemplo, 1517G> A) asociado con anemia de Fanconi, MYBPC3^{W1098X} (por ejemplo, 3293G> A) asociado con miocardiopatía hipertrófica familiar primaria e IL2RG^{W237X} (por ejemplo, 710G> A) asociado con inmunodeficiencia combinada grave ligada a X. En algunas modalidades, la enfermedad o afección es un cáncer. En algunas modalidades, la enfermedad o afección es una enfermedad monogenética. En algunas modalidades, la enfermedad o afección es una enfermedad poligenética.

10 En algunas modalidades, se proporciona un método para tratar un cáncer asociado con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el
15 ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es TP53^{W53X} (por ejemplo, 158G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 195, 196 o 197.

20 En algunas modalidades, se proporciona un método para tratar o prevenir un cáncer con un ARN objetivo que tiene una mutación (por ejemplo, mutación G> A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd al individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para
25 desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es TP53^{W53X} (por ejemplo, 158G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 195, 196 o 197.

30 En algunas modalidades, se proporciona un método para tratar MPS I (por ejemplo, el síndrome de Hurler o síndrome de Scheie) asociado con un ARN objetivo que tiene una mutación (por ejemplo, mutación G> A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una
35 ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo.

En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es IDUA^{W402X} (por ejemplo, mutación TGG> TAG en el exón 9). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 204 o 205.

En algunas modalidades, se proporciona un método para tratar o prevenir MPS I (por ejemplo, el síndrome de Hurler o síndrome de Scheie) con un ARN objetivo que tiene una mutación (por ejemplo, mutación G> A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd para el individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es IDUA^{W402X} (por ejemplo, mutación TGG> TAG en el exón 9). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 204 o 205.

En algunas modalidades, se proporciona un método para tratar o prevenir el síndrome de Ehlers-Danlos asociado con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en un célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y en donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es COL3A1^{W1278X} (por ejemplo, mutación 3833G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 198.

En algunas modalidades, se proporciona un método para tratar o prevenir el síndrome de Ehlers-Danlos con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd para el individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una

construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es COL3A1^{W1278X} (por ejemplo, mutación 3833G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 198.

En algunas modalidades, se proporciona un método para tratar la hipertensión pulmonar primaria asociada con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, en donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y en el que el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es BMPR2^{W298X} (por ejemplo, 893G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 199.

En algunas modalidades, se proporciona un método para tratar o prevenir la hipertensión pulmonar primaria con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd para el individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es BMPR2^{W298X} (por ejemplo, 893G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 199.

En algunas modalidades, se proporciona un método para tratar el síndrome de Joubert asociado con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es AHI1^{W725X} (por ejemplo, 2174G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 200.

En algunas modalidades, se proporciona un método para tratar o prevenir el síndrome de Joubert con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd al individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es AHI1^{W725X} (por ejemplo, 2174G> A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 200.

En algunas modalidades, se proporciona un método para tratar la anemia de Fanconi asociada con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es FANCC^{W506X} (por ejemplo, 1517G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 201.

En algunas modalidades, se proporciona un método para tratar o prevenir la anemia de Fanconi con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd a la individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es FANCC^{W506X} (por ejemplo, 1517G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 201.

En algunas modalidades, se proporciona un método para tratar la miocardiopatía hipertrófica familiar primaria asociada con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende introducir un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una ADAR para

desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es MYBPC3^{W1098X} (por ejemplo, 3293G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 202.

En algunas modalidades, se proporciona un método para tratar o prevenir la miocardiopatía hipertrófica familiar primaria con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd para el individuo, en donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o afección, y en donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es MYBPC3^{W1098X} (por ejemplo, 3293G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 202.

En algunas modalidades, se proporciona un método para tratar la inmunodeficiencia combinada grave ligada al cromosoma X asociada con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende la introducción de un ARNd o una construcción que codifica el ARNd en una célula aislada del individuo *ex vivo*, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en la célula aislada. En algunas modalidades, el método comprende introducir la ADAR o una construcción que codifica la ADAR en la célula aislada. En algunas modalidades, el ARN objetivo es IL2RG^{W237X} (por ejemplo, 710G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 203.

En algunas modalidades, se proporciona un método para tratar o prevenir la inmunodeficiencia combinada grave ligada al cromosoma X con un ARN objetivo que tiene una mutación (por ejemplo, mutación G>A) en un individuo, que comprende administrar una cantidad eficaz de un ARNd o una construcción que codifica el ARNd al individuo, donde el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo asociado con la enfermedad o condición, y donde el ARNd es capaz de reclutar una ADAR para desaminar un objetivo A en el ARN objetivo, rescatando así la mutación en el ARN objetivo. En algunas modalidades, la ADAR es una ADAR expresada endógenamente en las células del individuo. En algunas modalidades, el método

comprende administrar la ADAR o una construcción que codifica la ADAR al individuo. En algunas modalidades, el ARN objetivo es IL2RG^{W237X} (por ejemplo, 710G>A). En algunas modalidades, el ARNd comprende la secuencia de ácido nucleico de SEQ ID NO: 203.

5 Como se usa en el presente documento, "tratamiento" o "que trata" es un enfoque para obtener resultados beneficiosos o deseados, incluidos los resultados clínicos. Para los propósitos de esta invención, los resultados clínicos beneficiosos o deseados incluyen, pero no se limitan a, uno o más de los siguientes: disminuir uno más de los síntomas resultantes de la enfermedad, disminuir la extensión de la enfermedad, estabilizar la enfermedad (por ejemplo, prevenir o retrasar el empeoramiento de la enfermedad), prevenir o retrasar la propagación (por ejemplo, metástasis) de la enfermedad, prevenir o retrasar la aparición o recurrencia de la enfermedad, retrasar o ralentizar la progresión de la enfermedad, mejorar el estado de la enfermedad, proporcionar una remisión (ya sea parcial o total) de la enfermedad, disminuyendo la dosis de uno o más medicamentos necesarios para tratar la enfermedad, retrasando la progresión de la enfermedad, aumentando la calidad de vida y/o prolongando la supervivencia. También se incluye por "tratamiento" una reducción de las
10 consecuencias patológicas de la enfermedad o afección. Los métodos de la invención contemplan uno o más de estos aspectos del tratamiento.

Los términos "individuo", "sujeto" y "paciente" se usan indistintamente en el presente documento para describir un mamífero, incluidos los seres humanos. Un individuo incluye, pero no se limita a, humano, bovino, caballo, felino, canino, roedor o primate. En algunas modalidades, el
20 individuo es humano. En algunas modalidades, un individuo padece una enfermedad o afección, como resistencia a los fármacos. En algunas modalidades, el individuo necesita tratamiento.

Como se entiende en la técnica, una "cantidad eficaz" se refiere a una cantidad de una composición (por ejemplo, ARNd o construcciones que codifican el ARNd) suficiente para producir un resultado terapéutico deseado (por ejemplo, reducir la gravedad o duración de, estabilizar la gravedad de o eliminar uno o más síntomas de una enfermedad o afección). Para uso terapéutico, los resultados
25 beneficiosos o deseados incluyen, por ejemplo, disminuir uno o más síntomas resultantes de la enfermedad (bioquímicos, histológicos y/o conductuales), incluidas sus complicaciones y fenotipos patológicos intermedios presentados durante el desarrollo de la enfermedad, aumentando la calidad de vida de los que padecen la enfermedad o afección, disminuir la dosis de otros medicamentos necesarios para tratar la enfermedad, potenciar el efecto de otro medicamento, retrasar la progresión de la enfermedad y/o prolongar la supervivencia de los pacientes.

Generalmente, las dosis, programas y vías de administración de las composiciones (por ejemplo, ARNd o construcción que codifica el ARNd) se pueden determinar según el tamaño y la condición del individuo y según la práctica farmacéutica estándar. Las vías de administración
35 ejemplares incluyen intravenosa, intraarterial, intraperitoneal, intrapulmonar, intravesicular,

intramuscular, intratraqueal, subcutánea, intraocular, intratecal o transdérmica.

Los métodos de edición de ARN de la presente solicitud no solo pueden usarse en células animales, por ejemplo células de mamíferos, sino que también pueden usarse en la modificación de ARN de plantas u hongos, por ejemplo, en plantas u hongos que tienen ADAR expresada endógenamente. Los métodos descritos en este documento pueden usarse para generar plantas y hongos modificados genéticamente con propiedades mejoradas.

Composiciones, kits y artículos manufacturados

También se proporcionan en el presente documento composiciones (tales como composiciones farmacéuticas) que comprenden cualquiera de los ARNd, construcciones, bibliotecas o células hospederas que tienen ARN editado como se describe en el presente documento.

En algunas modalidades, se proporciona una composición farmacéutica que comprende cualquiera de los ARNd o construcciones que codifican el ARNd descrito en este documento, y un vehículo, excipiente o estabilizador farmacéuticamente aceptable. Se han descrito vehículos, excipientes y estabilizadores farmacéuticamente aceptables ejemplares, por ejemplo, en Remington's Pharmaceutical Sciences 16ª edición, Osol, A. Ed. (1980). Los vehículos, excipientes o estabilizadores aceptables no son tóxicos para los receptores en las dosis y concentraciones empleadas, e incluyen reguladores de pH tales como fosfato, citrato y otros ácidos orgánicos; antioxidantes que incluyen ácido ascórbico y metionina; conservantes (tales como cloruro de octadecil dimetilbencil amonio; cloruro de hexametonio; cloruro de benzalconio, cloruro de bencetonio; fenol, alcohol butílico o bencílico; alquil parabenos tales como metil o propilparabeno; catecol; resorcinol; ciclohexanol; 3-pentanol; y m-cresol); polipéptidos de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas, como albúmina sérica, gelatina o inmunoglobulinas; polímeros hidrófilos como olivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparagina, histidina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos que incluyen glucosa, manosa o dextrinas; agentes quelantes como EDTA; azúcares como sacarosa, manitol, trehalosa o sorbitol; contraiones formadores de sal tales como sodio; complejos metálicos (por ejemplo, complejos Zn-proteína); y/o agentes tensoactivos no iónicos como TWEEN™, PLURONICS™ o polietilenglicol (PEG). En algunas modalidades, se proporcionan formulaciones liofilizadas. Las composiciones farmacéuticas que se utilizarán para la administración *in vivo* deben ser estériles. Esto se logra fácilmente, por ejemplo, mediante filtración a través de membranas de filtración estériles.

Se proporcionan además kits o artículos de fabricación útiles para cualquiera de los métodos de edición de ARN o métodos de tratamiento descritos en este documento, que comprenden cualquiera de los ARNd, construcciones, composiciones, bibliotecas o células hospederas editadas como se describe en este documento.

En algunas modalidades, se proporciona un kit para editar un ARN objetivo en una

célula hospedera, que comprende un ARNd, en el que el ARNd comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, en el que el ARNd es capaz de reclutar una ADAR para desaminar una A en el ARN objetivo. En algunas modalidades, el kit comprende además una ADAR o una construcción que codifica una ADAR. En algunas modalidades, el kit comprende además un inhibidor de ADAR3 o una construcción del mismo. En algunas modalidades, el kit comprende además un estimulador de interferón o una construcción del mismo. En algunas modalidades, el kit comprende además una instrucción para llevar a cabo cualquiera de los métodos de edición de ARN descritos en el presente documento.

Los kits de la presente solicitud están en un envasado adecuado. El envasado adecuado incluye, pero no se limita a, viales, botellas, frascos, envases flexibles (por ejemplo, Mylar o bolsas de plástico selladas) y similares. Los kits pueden proporcionar opcionalmente componentes adicionales tales como transfección o reactivos de transducción, medio de cultivo celular, reguladores de pH e información interpretativa.

Por tanto, la presente solicitud también proporciona artículos manufacturados. El artículo de fabricación puede comprender un recipiente y una etiqueta o prospecto sobre el recipiente o asociado con él. Los recipientes adecuados incluyen viales (tales como viales sellados), botellas, tarros, envases flexibles y similares. En algunas modalidades, el recipiente contiene una composición farmacéutica y puede tener un puerto de acceso estéril (por ejemplo, el recipiente puede ser una bolsa de solución intravenosa o un vial que tiene un tapón perforable por una aguja de inyección hipodérmica). El recipiente que contiene la composición farmacéutica puede ser un vial de usos múltiples, que permite administraciones repetidas (por ejemplo, de 2 a 6 administraciones) de la formulación reconstituida. El prospecto se refiere a las instrucciones que habitualmente se incluyen en los paquetes comerciales de productos terapéuticos que contienen información sobre las indicaciones, el uso, la dosis, la administración, las contraindicaciones y/o advertencias relativas al uso de dichos productos. Además, el artículo de fabricación puede comprender además un segundo recipiente que comprende un regulador de pH farmacéuticamente aceptable, tal como agua bacteriostática para inyección (BWFI), solución salina amortiguada con fosfato, solución de Ringer y solución de dextrosa. Puede incluir además otros materiales deseables desde un punto de vista comercial y de usuario, incluidos otros reguladores de pH, diluyentes, filtros, agujas y jeringas.

Los kits o el artículo de fabricación pueden incluir múltiples dosis unitarias de las composiciones farmacéuticas e instrucciones de uso, empaquetadas en cantidades suficientes para su almacenamiento y uso en farmacias, por ejemplo, farmacias hospitalarias y farmacias de compuestos.

Modalidades ejemplares

Entre las modalidades proporcionadas en este documento se encuentran:

1. Un método para editar un ARN objetivo en una célula hospedera, que comprende introducir un

- ARN de reclutamiento de desaminasas (ARNd) o una construcción que codifica el ARNd en la célula hospedera, donde el ARNd comprende una secuencia de ARN complementaria que hibrida con el ARN objetivo, y el ARN de reclutamiento de desaminasa es capaz de reclutar una adenosina desaminasa que actúa sobre el ARN (ADAR) para desaminar una adenosina objetivo en el ARN objetivo.
2. El método de la modalidad 1, en el que la secuencia de ARN comprende una citidina, adenosina o uridina directamente opuesta a la adenosina objetivo en el ARN objetivo.
 3. El método de la modalidad 2, en el que la secuencia de ARN comprende un desajuste de citidina directamente opuesto a la adenosina objetivo en el ARN objetivo.
 4. El método de la modalidad 3, en el que el desajuste de citidina se localiza al menos a 20 nucleótidos del extremo 3' de la secuencia complementaria y al menos a 5 nucleótidos del extremo 5' de la secuencia complementaria en el ARNd.
 5. El método de la modalidad 4, en el que el desajuste de citidina se localiza dentro de los 10 nucleótidos del centro (por ejemplo, en el centro) de la secuencia complementaria en el ARNd.
 6. El método de cualquiera de las modalidades 1-5, en el que la secuencia de ARN comprende además una o más guanosinas, cada una opuesta a una adenosina no objetivo en el ARN objetivo.
 7. El método de cualquiera de las modalidades 1-6, en el que la secuencia complementaria comprende dos o más nucleótidos erróneos consecutivos opuestos a una adenosina no objetivo en el ARN objetivo.
 8. El método de cualquiera de las modalidades 1-7, en el que el vecino más cercano 5' de la adenosina objetivo en el ARN objetivo es un nucleótido seleccionado entre U, C, A y G con preferencia $U > C > A > G$ y el vecino más cercano 3' de la adenosina objetivo en el ARN objetivo es un nucleótido seleccionado entre G, C, A y U con preferencia $G > C > A \approx U$.
 9. El método de cualquiera de las modalidades 1-8, en el que la adenosina objetivo está en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN objetivo.
 10. El método de la modalidad 9, en el que el motivo de tres bases es UAG, y en el que el ARN de reclutamiento de desaminasas comprende una A directamente opuesta a la uridina en el motivo de tres bases, una citidina directamente opuesta a la adenosina objetivo y una citidina, guanosina o uridina directamente opuesta a la guanosina en el motivo de tres bases.
 11. El método de cualquiera de las modalidades 1 a 10, en el que el ARN de reclutamiento de desaminasas tiene una longitud de aproximadamente 40 a 260 nucleótidos.
 12. El método de la modalidad 11, en el que el ARN de reclutamiento de desaminasas tiene una longitud de aproximadamente 60-230 nucleótidos.
 13. El método de la modalidad 11 o 12, en el que el ARNd tiene más de aproximadamente 70

- nucleótidos de longitud.
14. El método de cualquiera de las modalidades 11-13, en el que el ARNd tiene una longitud de aproximadamente 100 a aproximadamente 150 (por ejemplo, aproximadamente 110-150) nucleótidos.
 - 5 15. El método de cualquiera de las modalidades 1-14, en el que el ARN objetivo es un ARN seleccionado del grupo que consiste en un ARN pre-mensajero, un ARN mensajero, un ARN ribosómico, un ARN de transferencia, un ARN largo no codificante y un ARN pequeño.
 16. El método de la modalidad 15, en el que el ARN objetivo es un ARN pre-mensajero.
 17. El método de cualquiera de las modalidades 1-16, en donde la ADAR es expresada
 - 10 endógenamente por la célula hospedera.
 18. El método de cualquiera de las modalidades 1-16, en el que la ADAR es exógena a la célula hospedera.
 19. El método de la modalidad 18, que comprende además introducir la ADAR en la célula hospedera.
 20. El método de la modalidad 18 o 19, en el que ADAR comprende una mutación E1008.
 - 15 21. El método de cualquiera de las modalidades 1-20, en el que el ARN de reclutamiento de desaminasas es un ARN de cadena sencilla.
 22. El método de cualquiera de las modalidades 1-20, en el que la secuencia de ARN complementaria es de cadena sencilla y en el que el ARN de reclutamiento de desaminasas comprende además una o más regiones de cadena doble.
 - 20 23. El método de cualquiera de las modalidades 1-22, en el que el ARNd no comprende un dominio de reclutamiento de ADAR (por ejemplo, un dominio de unión a DSB, un dominio GluR2 o un dominio MS2).
 24. El método de una cualquiera de las modalidades 1-23, en el que el ARNd no comprende un nucleótido modificado químicamente (por ejemplo, modificación con 2'-O-metilo o fosforotioato).
 - 25 25. El método de la modalidad 24, que comprende introducir una construcción que codifica el ARNd en la célula hospedera, en el que el construcción es un vector viral (por ejemplo, vector lentiviral) o un plásmido.
 26. El método de cualquiera de las modalidades 1-25, en el que la desaminación de la adenosina objetivo en el ARN objetivo da como resultado una mutación sin sentido, un codón de terminación
 - 30 temprano, empalme aberrante o empalme alternativo en el ARN objetivo, o la reversión de una mutación sin sentido, un codón de terminación temprano, empalme aberrante o empalme alternativo en el ARN objetivo.
 27. El método de la modalidad 26, en donde la desaminación de la adenosina objetivo en el ARN objetivo da como resultado una mutación puntual, truncamiento, alargamiento y/o plegamiento
 - 35 incorrecto de la proteína codificada por el ARN objetivo, o una proteína funcional, de longitud

completa, plegada correctamente y/o de tipo salvaje mediante la inversión de una mutación sin sentido, un codón de terminación temprano, empalme aberrante o empalme alternativo en el ARN objetivo.

28. El método de cualquiera de las modalidades 1-27, en donde la célula hospedera es una célula eucariota.
29. El método de la modalidad 28, en el que la célula hospedera es una célula de mamífero.
30. El método de la modalidad 29, en el que la célula hospedera es una célula humana o de ratón.
31. El método de la modalidad 29 o 30, en donde ADAR es ADAR1 y/o ADAR2.
32. El método de cualquiera de las modalidades 1 a 31, en el que la célula hospedera es una célula primaria.
33. El método de la modalidad 32, en el que la célula hospedera es una célula T.
34. El método de la modalidad 32, en el que la célula hospedera es una célula posmitótica.
35. El método de cualquiera de las modalidades 1-34, que comprende además la introducción de un inhibidor de ADAR3 en la célula hospedera.
36. El método de cualquiera de las modalidades 1 a 35, que comprende además la introducción de un estimulador de interferón en la célula hospedera.
37. El método de cualquiera de las modalidades 1-36, que comprende introducir una pluralidad de ARNd, cada uno de los cuales se dirige a un ARN objetivo diferente.
38. El método de cualquiera de las modalidades 1-37, en el que la eficiencia de editar el ARN objetivo es al menos aproximadamente el 5% (por ejemplo, al menos aproximadamente el 10%, 20% o 30%).
39. El método de cualquiera de las modalidades 1 a 38, en el que el ARNd no induce una respuesta inmune.
40. Un ARN editado o una célula hospedera que tiene un ARN editado producido por el método de cualquiera de las modalidades 1-39.
41. Un método para tratar o prevenir una enfermedad o afección en un individuo, que comprende editar un ARN objetivo asociado con la enfermedad o afección en una célula del individuo según cualquiera de las modalidades 1-39.
42. El método de la modalidad 41, en el que la enfermedad o afección es una enfermedad genética hereditaria o una enfermedad o afección asociada con una o más mutaciones genéticas adquiridas.
43. El método de la modalidad 41 o 42, en el que el ARN objetivo tiene una mutación de G a A.
44. El método de cualquiera de las modalidades 41-43, en el que la enfermedad o afección es una enfermedad o afección monogenética.
45. El método de cualquiera de las modalidades 41-44, en el que la enfermedad o afección es una

enfermedad o afección poligenética.

46. El método de cualquiera de las modalidades 41 a 45, en donde:
 - (i) el ARN objetivo es TP53 y la enfermedad o afección es cáncer;
 - (ii) el ARN objetivo es IDUA y la enfermedad o afección es mucopolisacaridosis tipo I (MPS I);
 - 5 (iii) el ARN objetivo es COL3A1 y la enfermedad o afección es el síndrome de Ehlers-Danlos;
 - (iv) el ARN objetivo es BMPR2 y la enfermedad o afección es el síndrome de Joubert;
 - (v) el ARN objetivo es FANCC y la enfermedad o afección es anemia de Fanconi;
 - (vi) el ARN objetivo es MYBPC3 y la enfermedad o afección es una miocardiopatía hipertrófica familiar primaria; o
 - 10 (vii) el ARN objetivo es IL2RG y la enfermedad o afección es inmunodeficiencia combinada grave ligada al cromosoma X.
47. Un ARN de reclutamiento de desaminasa (ARNd) para la desaminación de una adenosina objetivo en un ARN objetivo mediante el reclutamiento de una adenosina desaminasa que actúa sobre el ARN (ADAR), que comprende una secuencia de ARN complementaria que se hibrida con el ARN
 - 15 objetivo.
48. El ARN de reclutamiento de desaminasa de la modalidad 47, en el que la secuencia de ARN comprende una citosina, adenosina o U directamente opuesta a la adenosina objetivo en el ARN objetivo.
49. El ARNd de la modalidad 48, en el que la secuencia de ARN comprende un desajuste de citidina
 - 20 directamente opuesto a la adenosina objetivo en el ARN objetivo.
50. El ARNd de la modalidad 49, en el que el desajuste de citidina está ubicado al menos a 20 nucleótidos del extremo 3' de la secuencia complementaria, y al menos a 5 nucleótidos del extremo 5' de la secuencia complementaria en el ARNd.
51. El ARNd de la modalidad 50, en el que el desajuste de citidina se localiza dentro de los 10
 - 25 nucleótidos del centro (por ejemplo, en el centro) de la secuencia complementaria en el ARNd.
52. El ARN de reclutamiento de desaminasa de cualquiera de las modalidades 47-51, en el que la secuencia de ARN comprende además una o más guanosinas, cada una directamente opuesta a una adenosina no objetivo en el ARN objetivo.
53. El ARNd de cualquiera de las modalidades 47-51, en el que la secuencia complementaria
 - 30 comprende dos o más nucleótidos erróneos consecutivos opuestos a una adenosina no objetivo en el ARN objetivo.
54. El ARN de reclutamiento de desaminasa de cualquiera de las modalidades 47-53, en el que la adenosina objetivo está en un motivo de tres bases seleccionado del grupo que consiste en UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA y GAU en el ARN
 - 35 objetivo.

55. El ARN de reclutamiento de desaminasa de la modalidad 54, en el que el motivo de tres bases es UAG, y en el que el ARNd comprende una adenosina directamente opuesta a la uridina en el motivo de tres bases, una citosina directamente opuesta a la adenosina objetivo y una citidina, guanosina o uridina directamente opuesta a la guanosina en el motivo de tres bases.
- 5 56. El ARN de reclutamiento de desaminasas de la modalidad 55, en el que el motivo de tres bases es UAG en el ARN objetivo, y en el que el ARN de reclutamiento de desaminasas comprende ACC, ACG o ACU opuesto a UAG del ARN objetivo.
57. El ARN de reclutamiento de desaminasas de cualquiera de las modalidades 47-56, en el que el ARN de reclutamiento de desaminasas tiene una longitud de aproximadamente 40-260 nucleótidos.
- 10 58. El ARNd de la modalidad 57, en el que el ARNd tiene más de aproximadamente 70 nucleótidos de longitud.
59. El ARNd de la modalidad 57 o 58, en el que el ARNd tiene aproximadamente 100 a aproximadamente 150 nucleótidos (por ejemplo, aproximadamente 110-150) de longitud.
60. El ARNd de cualquiera de las modalidades 47-59, en el que el ARNd no comprende un dominio de reclutamiento de ADAR (por ejemplo, un dominio de unión a DSB, un dominio GluR2 o un dominio MS2).
- 15 61. El ARNd de cualquiera de las modalidades 47-60, en el que el ARNd no comprende un nucleótido modificado químicamente (por ejemplo, modificación con 2'-O-metil o fosforotioato).
62. Una construcción que codifica el ARN de reclutamiento de desaminasas de cualquiera de las modalidades 47-61.
- 20 63. La construcción de la modalidad 62, en la que la construcción es un vector viral (por ejemplo, un vector lentiviral) o un plásmido.
64. Una biblioteca que comprende una pluralidad de ARN de reclutamiento de desaminasas de cualquiera de las modalidades 47-61 o la construcción de la modalidad 62 o 63.
- 25 65. Una composición que comprende el ARN de reclutamiento de desaminasas de cualquiera de las modalidades 47-61, la construcción de la modalidad 62 o 63, o la biblioteca de la modalidad 64.
66. Una célula hospedera que comprende el ARN de reclutamiento de desaminasas de cualquiera de las modalidades 47-61 o la construcción de la modalidad 62 o 63.
67. La célula hospedera de la modalidad 66, en la que la célula hospedera es una célula eucariota.
- 30 68. La célula hospedera de la modalidad 66 o 67, en la que la célula hospedera es una célula primaria.
69. Un kit para editar un ARN objetivo en una célula hospedera, que comprende un ARN de reclutamiento de desaminasas, en el que el ARN de reclutamiento de desaminasas comprende una secuencia de ARN complementaria que se hibrida con el ARN objetivo, en el que el ARN de reclutamiento de desaminasas es capaz de reclutar una ADAR para desaminar una adenosina
- 35 objetivo en el ARN objetivo.

EJEMPLO

Se pretende que los ejemplos siguientes sean puramente ejemplares de la presente solicitud y, por lo tanto, no se debe considerar que limitan la invención de ninguna manera. Los siguientes ejemplos y descripción detallada se ofrecen a modo de ilustración y no a modo de limitación.

Materiales y métodos

Construcción de plásmidos

El reportero de fluorescencia dual se clonó mediante PCR amplificando mCherry y EGFP (se eliminó el primer codón ATG de EGFP) que codificaba el ADN, el enlazador 3×GS y la secuencia de ADN dirigida se agregaron mediante cebadores durante la PCR. Luego, los productos de la PCR se escindieron y se unieron mediante la enzima de restricción BsmB1 (Thermo) de tipo II y la ADN ligasa T4 (NEB), que luego se insertaron en la estructura fundamental pLenti (pLenti-CMV-MCS-SV-Bsd, Stanley Cohen Lab, Stanford University).

El ADN de dLbuCas13 se amplificó por PCR a partir de los plásmidos Lbu (Addgene #83485). La ADAR1DD y ADAR2DD se amplificaron a partir de ADNc de Adar1(p150) y ADNc de Adar2, ambos obsequios del laboratorio de Han en la Universidad de Xiamen. La ADAR1DD o ADAR2DD se fusionaron con el ADN de dLbuCas13 mediante superposición-PCR, y los productos de PCR fusionados se insertaron en la estructura fundamental pLenti.

Para la expresión de ARNd en células de mamíferos, las secuencias de ARNd se sintetizaron directamente (para ARNd cortos) y se reasociaron o amplificaron mediante PCR sintetizando ADNss superpuesto, y los productos se clonaron en los vectores correspondientes bajo expresión de U6 mediante clonación Golden-gate.

La Adar1 de longitud completa (p110) y Adar1 (p150) se amplificaron por PCR a partir de ADNc de Adar1 (p150), y el Adar2 de longitud completa se amplificó por PCR a partir de ADNc de Adar2, que luego se clonaron en la estructura fundamental pLenti, respectivamente.

Para las tres versiones de los reporteros de fluorescencia dual (reportero-1, -2 y -3), las secuencias codificantes de *mCherry* y EGFP (se eliminó el codón de inicio ATG de EGFP) se amplificaron por PCR, se digirieron usando BsmBI (Thermo Fisher Scientific, ER0452), seguido por ligadura mediada por ligasa de ADN T4 (NEB, M0202L) con enlazadores GGGGS. El producto de ligación se insertó posteriormente en la estructura fundamental pLenti-CMV-MCS-PURO.

Para la construcción que expresa dLbuCas13-ADAR_{DD} (E1008Q), el gen ADAR1_{DD} se amplificó a partir de la construcción de ADAR1^{p150} (un regalo del laboratorio de Jiahuai Han, Universidad de Xiamen). El gen dLbuCas13 se amplificó mediante PCR a partir del plásmido

Lbu_C2c2_R472A_H477A_R1048A_H1053A (Addgene #83485). La ADAR1DD (variante hiperactiva de E1008Q) se generó por superposición-PCR y luego se fusionó con dLbuCas13. Los productos de ligación se insertaron en la estructura fundamental pLenti-CMV-MCS-BSD.

Para la construcción que expresa ARNar, las secuencias de ARNar se sintetizaron y se clonaron con Golden-gate en la estructura fundamental pLenti-sARNg-lib 2.0 (Addgene #89638), y la transcripción del ARNar fue impulsada por el promotor hU6. Para las construcciones que expresan ADAR, la ADAR1^{p110} y ADAR1^{p150} de longitud completa se amplificaron por PCR a partir de la construcción de ADAR1^{p150}, y la ADAR2 de longitud completa se amplificó por PCR a partir de la construcción de ADAR2 (un obsequio del laboratorio de Jiahui Han, Universidad de Xiamen). A continuación, los productos amplificados se clonaron en la estructura fundamental pLenti-CMV-MCS-BSD.

Para las construcciones que expresan genes con mutaciones patógenas, secuencias codificantes completas de TP53 (ordenadas a Vigenebio) y otros 6 genes relevantes para la enfermedad (*COL3A1*, *BMPR2*, *AHI1*, *FANCC*, *MYBPC3* e *IL2RG*, obsequios del laboratorio de Jianwei Wang, Instituto de biología de patógenos, Academia China de Ciencias Médicas) se amplificaron a partir de las construcciones que codifican los genes correspondientes con la introducción de mutaciones G>A mediante mutagénesis PCR. Los productos amplificados se clonaron en la estructura fundamental pLenti-CMV-MCS-mCherry mediante el método de clonación de Gibson⁵⁹.

Líneas celulares de mamíferos y cultivo celular

Se cultivaron líneas celulares de mamíferos en medio Eagle modificado de Dulbecco (10-013-CV, Corning, Tewksbury, MA, E.U.A.), agregando suero bovino fetal al 10% (Lanzhou Bailing Biotechnology Co., Ltd., Lanzhou, China), suplementado con penicilina-estreptomicina al 1% en CO₂ al 5% a 37°C. La línea celular Adar1-KO se adquirió de EdiGene China, y EdiGene China también proporcionó los resultados de la genotipificación.

Las líneas celulares HeLa y B16 eran del laboratorio de Z. Jiang (Universidad de Pekín). Y la línea celular HEK293T era del laboratorio de C. Zhang (Universidad de Pekín). La línea celular RD era del laboratorio de J Wang (Instituto de Biología de Patógenos, Colegio Médico de la Unión de Pekín y Academia China de Ciencias Médicas). Las líneas celulares SF268 eran del Centro Celular, Instituto de Ciencias Médicas Básicas, Academia China de Ciencias Médicas. Las líneas celulares A549 y SW13 eran de EdiGene Inc. Las líneas celulares HepG2, HT29, NIH3T3 y MEF se mantuvieron en nuestro laboratorio en la Universidad de Pekín. Estas líneas de células de mamíferos se cultivaron en medio Eagle modificado de Dulbecco (Corning, 10-013-CV) con suero bovino fetal al 10% (CellMax, SA201.02), adicionalmente suplementado con penicilina-estreptomicina al 1% bajo CO₂ al 5% a 37°C. A menos que se describa lo contrario, las células se transfectaron con el reactivo de transfección de ADN de X-tremeGENE HP (Roche, 06366546001) de acuerdo con las instrucciones del fabricante.

Los fibroblastos pulmonares primarios humanos (#3300) y las células epiteliales bronquiales primarias humanas (#3210) se adquirieron en ScienCell Research Laboratories, Inc. y se cultivaron en medio de fibroblastos (ScienCell, #2301) y medio de células epiteliales bronquiales (ScienCell, #3211), respectivamente. Ambos medios se complementaron con suero bovino fetal (BI) al 15% y penicilina-estreptomicina al 1%. Las células primarias GM06214 y GM01323 se solicitaron al Instituto Coriell de Investigación Médica y se cultivaron en medio Eagle modificado de Dulbecco (Corning, 10-013-CV) con suero bovino fetal (BI) al 15% y penicilina-estreptomicina al 1%. Todas las células se cultivaron bajo CO₂ al 5% a 37°C.

Transfección del sistema reportero, análisis FACS y secuenciación de

Sanger

Para los experimentos de edición de reportero de fluorescencia dual, se sembraron células 293T-WT o células 293T-Adar1-KO en placas de 6 pozos (6×10^5 células/pozo), 24 horas después, se cotransfectaron 1.5 µg de plásmidos reporteros y 1.5 µg de plásmidos de ARNd utilizando el reactivo de transfección de ADN X-tremeGENE HP (06366546001; Roche, Mannheim, Alemania), de acuerdo con los protocolos del proveedor. De 48 a 72 horas después, se recogieron las células y se realizó el análisis FACS. Para confirmar aún más la edición del ARNm reportero, clasificamos las células positivas para EGFP de las células 293T-WT transfectadas con plásmidos reporteros y ARNd utilizando un citómetro de flujo FACS Aria (BD Biosciences), seguido de aislamiento total de ARN (TIANGEN, DP430). Luego, el ARN se transcribió de forma inversa en ADNc mediante RT-PCR (TIANGEN, KR103-04), y el locus objetivo se amplificó por PCR con los pares de cebadores correspondientes (23 ciclos de PCR) y los productos de PCR se purificaron para la secuenciación de Sanger.

Para los experimentos de sobreexpresión y rescate de Adar1(p110), Adar1(p150) o Adar2, se sembraron células 293T-WT o células 293T-Adar1-KO en placas de 12 pozos (2.5×10^5 células/pozo), 24 horas después, 0.5 µg de plásmidos reporteros, 0.5 µg de plásmidos de ARNd y 0.5 µg de plásmidos Adar1/2 (estructura fundamental pLenti como control) se contranfectaron utilizando el reactivo de transfección de DNA X-tremeGENE HP (06366546001, Roche, Mannheim, Alemania). De 48 a 72 horas después, se recogieron las células y se realizó el análisis FACS.

Para los experimentos de ARNm endógeno, se sembraron células 293T-WT en placas de 6 pozos (6×10^5 células/pozo). Cuando aproximadamente un 70% de confluencia, se transfectaron 3 µg de plásmidos de ARNd utilizando el reactivo de transfección de ADN X-tremeGENE HP (06366546001, Roche, Mannheim, Alemania). 72 horas más tarde, se recogieron las células y se clasificaron las células positivas para GFP o positivas para BFP (según el fabricante de fluorescencia correspondiente) mediante FACS para el siguiente aislamiento de ARN.

Aislamiento y cultivo de células T primarias humanas

Se aislaron células T humanas primarias a partir de productos de leucaféresis de un

donante humano sano. Brevemente, las células mononucleares de sangre periférica (PBMC) se aislaron mediante centrifugación Ficoll (Dakewei, AS1114546), y las células T se aislaron mediante selección magnética negativa usando un kit de aislamiento de células T humanas EasySep (STEMCELL, 17951) de PBMC. Después del aislamiento, las células T se cultivaron en medio X-vivo15, FBS al 10% e IL2 (1000 U/ml) y se estimularon con DynaBeads CD3/CD28 (ThermoFisher, 11131D) durante 2 días. Los productos de leucaféresis de donantes sanos se adquirieron de AllCells LLC China. Todos los donantes sanos dieron su consentimiento informado.

Construcción de la línea de células reporteras y paquete de lentivirus

El plásmido de expresión se cotransfectó en células HEK293T-WT, junto con dos plásmidos de empaquetado viral, pR8.74 y pVSVG (Addgene) mediante el reactivo de transfección de ADN X-tremeGENE HP. 72 horas después, se recogió el virus sobrenadante y se almacenó a -80°C. Las células HEK293T-WT se infectaron con lenti-virus, 72 horas más tarde, las células positivas para mCherry se clasificaron mediante FACS y se cultivaron para seleccionar líneas celulares de un solo clon que expresan de manera estable el sistema reportero de fluorescencia dual con un fondo de EGFP muy bajo mediante el método de dilución limitante.

Para las líneas celulares reporteras estables, las construcciones reporteras (estructura fundamental pLenti-CMV-MCS-PURO) se cotransfectaron en células HEK293T, junto con dos plásmidos de empaquetamiento viral, pR8.74 y pVSVG. 72 horas después, se recogió el virus sobrenadante y se almacenó a -80°C. Las células HEK293T se infectaron con lentivirus, luego las células positivas para mCherry se clasificaron mediante FACS y se cultivaron para seleccionar líneas celulares de un solo clon que expresan de manera estable el sistema indicador de fluorescencia dual sin un fondo de EGFP detectable. Las líneas celulares HEK293T ADAR1^{-/-} y TP53^{-/-} se generaron de acuerdo con un método previamente informado⁶⁰. El ARNsg que dirige ADAR1 y el ADN del donante amplificado por PCR que contiene el gen resistente a puromicina accionado por CMV se cotransfectaron en células HEK293T. Luego, las células se trataron con puromicina 7 días después de la transfección. Se aislaron clones individuales de células resistentes a puromicina seguido de verificación mediante secuenciación y Western blot.

Edición del ARN de transcritos de expresión endógena o exógena

Para evaluar la edición de ARN en el reportero de fluorescencia dual, se sembraron células HEK293T o células HEK293T ADAR1^{-/-} en placas de 6 pozos (6 x 10⁵ células/pozo). 24 horas más tarde, las células se cotransfectaron con 1.5 µg de plásmidos reporteros y 1.5 µg de plásmidos de ARNar. Para examinar el efecto de la expresión de la proteína ADAR1^{p110}, ADAR1^{p150} o ADAR2, se evaluó la eficiencia de edición mediante la proporción positiva de EGFP y la secuenciación profunda.

Se sembraron células HEK293T ADAR1^{-/-} en placas de 12 pozos (2.5 x 10⁵ células/pozo). 24 horas después, las células se cotransfectaron con 0.5 µg de plásmidos reporteros,

0.5 µg de plásmidos de ARNar y 0.5 µg de plásmidos ADAR1/2 (estructura fundamental pLenti como control). La eficiencia de edición se evaluó mediante la proporción positiva de EGFP y la secuenciación profunda.

5 Para evaluar la edición de ARN en los transcritos de ARNm endógeno, se sembraron células HEK293T en placas de 6 pozos (6×10^5 células/pozo). Veinticuatro horas más tarde, las células se transfectaron con 3 µg de plásmidos de ARNar. La eficacia de la edición se evaluó mediante secuenciación profunda.

10 Para evaluar la eficacia de la edición de ARN en múltiples líneas celulares, se sembraron células de $8-9 \times 10^4$ (RD, SF268, HeLa) o 1.5×10^5 (HEK293T) en placas de 12 pozos. Para las células difíciles de transfectar, como HT29, A549, HepG2, SW13, NIH3T3, MEF y B16, se sembraron $2-2.5 \times 10^5$ células en una placa de 6 pozos. Veinticuatro horas más tarde, los reporteros y el plásmido de ARNar se cotransfectaron en estas células. La eficiencia de edición se evaluó mediante la proporción positiva de EGFP.

15 Para evaluar la proporción positiva de EGFP, de 48 a 72 horas después de la transfección, las células se clasificaron y recolectaron mediante análisis de clasificación de células activadas por fluorescencia (FACS). La señal de mCherry se utilizó como marcador de selección fluorescente para las células que expresan el reportero/ARNar, y los porcentajes de células EGFP⁺/mCherry⁺ se calcularon como la lectura para la eficiencia de edición.

20 Para la cuantificación de NGS de la velocidad de edición de A a I, de 48 a 72 horas después de la transfección, las células se clasificaron y recogieron mediante el ensayo FACS y luego se sometieron a aislamiento de ARN (TIANGEN, DP420). Luego, los ARN totales se transcribieron de manera inversa en ADNc mediante RT-PCR (TIANGEN, KR103-04) y el locus dirigido se amplificó por PCR con los cebadores correspondientes enumerados en el siguiente cuadro.

Nombre del cebador	Secuencia (5'--->3')
mCherry-SpeI-F	TATAACTAGTATGGTGAGCAAGGGCGAGGAG (SEQ ID NO: 206)
mCherry-BsmBI-R1	TATACGTCTCATCTACAGATTCTTCCGGCGTGATACCTTC (SEQ ID NO: 207)
EGFP-BsmBI-F1 (Reportero-1)	TATACGTCTCATAGAGATCCCCGGTCGCCACCGTGAGCAAGGGCGAGGAG CTG (SEQ ID NO: 208)
EGFP-AscI-R	TATAGGCGCGCCTTACTTGTACAGCTCGTCCATGCC (SEQ ID NO: 209)
mCherry-BsmBI-R2	TATACGTCTCAAGGCGCTGCCTCCTCCGCCGCTGCCTCCTCCGCCGCTGCC TCCTCCGCCCTGCAGCTTGTACAGCTCGTCCATGCCGCCGGTG (SEQ ID NO: 210)
EGFP-BsmBI-F2 (Reportero-2)	TATACGTCTCAGCCTGCTCGCGATGCTAGAGGGCTCTGCCAGTGAGCAAGG GCGAGGAGCTG (SEQ ID NO: 211)

LbuCas13-SpeI-F	TATAACTAGTATGGTGGATTACAAGGATGACGACGATAAGATGAAAGTGAC GAAGGTAGGAGGCATTTTCG (SEQ ID NO: 212)
LbuCas13-AscI-R	ATATGGCGCGCCGTTTTTCAGACTTTTTCTCTTCCATTTGTATTCAAACATAA TCTTCAC (SEQ ID NO: 213)
hADAR1 _{DD} -AscI-F	TATAGGCGCGCCAGGCGGAGGAGGCAGCGGCGGAGGAGGCAGCCTCCTC CTCTCAAGGTCCCCAGAAGC (SEQ ID NO: 214)
hADAR1 _{DD} -SbfI-R	TATACCTGCAGGCTACACCTTGCGTTTTTTCTTGGGTACTGGGCAGAGATAA AAGTTCTTTTCC (SEQ ID NO: 215)
sec-Profunda-F (Reportero-1)	CACTCCACCGCGGCATGGACGAG (SEQ ID NO: 216)
sec-Profunda-R (Reportero-1)	CACGCTGAACTTGTGGCCGTTTACGTCG (SEQ ID NO: 217)
ADAR1-p150-SpeI-F	TATAACTAGTATGAATCCGCGGCAGGGGTATTCCCTCAGC (SEQ ID NO: 218)
ADAR1-p150-AscI-R	TATAGGCGCGCCCTACTTATCGTCGTCATCCTTGAATCTACTGGGCAGAGA TAAAAGTTCTTTTCCTCTGG (SEQ ID NO: 219)
ADAR2-SpeI-F	TATAACTAGTATGGATATAGAAGATGAAGAAAACATGAGTTC (SEQ ID NO: 220)
ADAR2-AscI-R	TATAGGCGCGCCCTACTTATCGTCGTCATCCTTGAATCGGGCGTGAGTGA GAACTGGTCCTGCTCG (SEQ ID NO: 221)
ADAR1-p110-SpeI-F	TATAACTAGTATGGCCGAGATCAAGGAGAAAATCTGC (SEQ ID NO: 222)
ADAR1-p110-AscI-R	TATAGGCGCGCCCTACTTATCGTCGTCATCCTTGAATCTACTGGGCAGAGA TAAAAGTTCTTTTCCTCTGG (SEQ ID NO: 223)
KRAS-sec-profunda-F	CGCCATTTTCGGACTGGGAG (SEQ ID NO: 224)
KRAS-sec-profunda-R	AGAGACAGGTTTCTCCATCAATTAC (SEQ ID NO: 225)
PPIB-sec-profunda-F	GAGCCCGCGAGCAACC (SEQ ID NO: 226)
PPIB-sec-profunda-R	GCAGCAGGAAGAAGACGGAC (SEQ ID NO: 227)
FANCC-sec-profunda-F1 (sitio TAC)	AGAAGCAGTTGAAGACCAGACTC (SEQ ID NO: 228)
FANCC-sec-profunda-R (sitio TAC)	GGCCTTCACCTGGACCATAG (SEQ ID NO: 229)
FANCC-sec-profunda-F2 (sitio TAC)	AGAGAAGCAGTTGAAGACCAGA (SEQ ID NO: 230)
FANCC-sec-profunda-R2 (sitio TAC)	CGGCCTTCACCTGGACCATA (SEQ ID NO: 231)
FANCC-sec-profunda-F3 (sitio TAC)	CAGAGAAGCAGTTGAAGACCAGA (SEQ ID NO: 232)
FANCC-sec-profunda-R3 (sitio TAC)	CGGCCTTCACCTGGACCATA (SEQ ID NO: 233)

SMAD4-sec-profunda-F1	TTTGTGAAAGGCTGGGGACC (SEQ ID NO: 234)
SMAD4-sec-profunda-R1	ACAGGATTGTATTTGTAGTCCACC (SEQ ID NO: 235)
SMAD4-sec-profunda-F2	AGGATGAGTTTTGTGAAAGGCTG (SEQ ID NO: 236)
SMAD4-sec-profunda-R2	ATTTTGTAGTCCACCATCCTGATA (SEQ ID NO: 237)
SMAD4-sec-profunda-F3	GATGAGTTTTGTGAAAGGCTGG (SEQ ID NO: 238)
SMAD4-sec-profunda-R3	ATTTTGTAGTCCACCATCCTGATAA (SEQ ID NO: 239)
TRAPPC12-sec-profunda-F	CGAAGAGAACGAGACCGCAT (SEQ ID NO: 240)
TRAPPC12-sec-profunda-R	GAAGATGGTGCACACCGGG (SEQ ID NO: 241)
TARDBP-sec-profunda-F	GACAGATGCTTCATCAGCAGTG (SEQ ID NO: 242)
TARDBP-sec-profunda-R	CGAACAAAGCCAAACCCCTTT (SEQ ID NO: 243)
COL3A1-sec-profunda-F	TCTGTTAATGGACAAATAGAAAGCC (SEQ ID NO: 244)
COL3A1-sec-profunda-R	GGAACATTCAAAGGATTGGCACT (SEQ ID NO: 245)
BMPR2-sec-profunda-F	AGTCACTGCAGATGGACGCA (SEQ ID NO: 246)
BMPR2-sec-profunda-R	ATCTCGATGGGAAATTGCAGGT (SEQ ID NO: 247)
AHI1-sec-profunda-F	TCAGAGTTTTACCTCATCCTTCTTT (SEQ ID NO: 248)
AHI1-sec-profunda-R	CCTGAATACATATGATGACCTTCAG (SEQ ID NO: 249)
FANCC-sec-profunda-F (Sitio2)	AGGGCACAGACACAGACCTC (SEQ ID NO: 250)
FANCC-sec-profunda-R (Sitio2)	AGGGCTTTCAATGCCAAGACG (SEQ ID NO: 251)
MYBPC3-sec-profunda-F	TGACAAGCCAAGTCCTCCC (SEQ ID NO: 252)
MYBPC3-sec-profunda-R	ATTGCCAATGATGAGCTCTGG (SEQ ID NO: 253)
IL2RG-sec-profunda-F	TTATAGACATAAGTTCTCCTTGCCT (SEQ ID NO: 254)
IL2RG-sec-profunda-R	TCAATCCCATGGAGCCAACA (SEQ ID NO: 255)
1-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTTAAGTAGAGGCCGCGCACTCCACCGGCGGC (SEQ ID NO: 256)
2-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTATCATGCTTAGCCGCGCACTCCACCGGCGGC (SEQ ID NO: 257)
3-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTGATGCACATCTGCCGCGCACTCCACCGGCGGC (SEQ ID NO: 258)

4-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTCGATTGCTCGACGCCGCGCCACTCCACCGGCG GC (SEQ ID NO: 259)
5-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTTCGATAGCAATTCGCCGCGCCACTCCACCGGC GGC (SEQ ID NO: 260)
6-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTATCGATAGTTGCTTGCCGCGCCACTCCACCGGC GGC (SEQ ID NO: 261)
7-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTGATCGATCCAGTTAGGCCGCGCCACTCCACCG GCGGC (SEQ ID NO: 262)
8-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTCGATCGATTGAGCCTGCCGCGCCACTCCACC GGCGGC (SEQ ID NO: 263)
9-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTACGATCGATACACGATCGCCGCGCCACTCCACC GGCGGC (SEQ ID NO: 264)
10-sec-profunda-F (Reportero-3)	TACACGACGCTCTTCCGATCTTACGATCGATGGTCCAGAGCCGCGCCACTCCA CCGGCGGC (SEQ ID NO: 265)
1-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTTAAGTAGAGTCGCCGTCCAGCTCGACCAG (SEQ ID NO: 266)
2-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTATCATGCTTATCGCCGTCCAGCTCGACCAG (SEQ ID NO: 267)
3-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTGATGCACATCTTCGCCGTCCAGCTCGACCA G (SEQ ID NO: 268)
4-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTCGATTGCTCGACTCGCCGTCCAGCTCGACC AG (SEQ ID NO: 269)
5-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTTCGATAGCAATTCTCGCCGTCCAGCTCGAC CAG (SEQ ID NO: 270)
6-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTATCGATAGTTGCTTTCGCCGTCCAGCTCGAC CAG (SEQ ID NO: 271)
7-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTGATCGATCCAGTTAGTCGCCGTCCAGCTCG ACCAG (SEQ ID NO: 272)
8-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTCGATCGATTTGAGCCTTCGCCGTCCAGCTC GACCAG (SEQ ID NO: 273)
9-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTACGATCGATACACGATCTCGCCGTCCAGCT CGACCAG (SEQ ID NO: 274)
10-sec-profunda-R (Reportero-3)	AGACGTGTGCTCTTCCGATCTTACGATCGATGGTCCAGATCGCCGTCCAGC TCGACCAG (SEQ ID NO: 275)
ST3GAL1-sec-profunda-F	GGGGAACTCGGGCAACCT (SEQ ID NO: 276)

ST3GAL1-sec-profunda-R	GAATCGGATCTGCCCCGTG (SEQ ID NO: 277)
EHD2-sec-profunda-F	CATCGAGGCCAAGCTGGAA (SEQ ID NO: 278)
EHD2-sec-profunda-R	GTAGTGAGGAGGGAGACCCC (SEQ ID NO: 279)
OSTM1-AS1-sec-profunda-F	AAGCCTCCTTCCTTCCCCAA (SEQ ID NO: 280)
OSTM1-AS1-sec-profunda-R	ATCGATACACTCCCTAGCCCA (SEQ ID NO: 281)
IL6-qPCR-F1	ACAAATTCGGTACATCCTCGAC (SEQ ID NO: 282)
IL6-qPCR-R1	TTCAGCCATCTTTGGAAGTT (SEQ ID NO: 283)
INF- β -qPCR-F1	ACGCCGATTGACCATCTAT (SEQ ID NO: 284)
INF- β -qPCR-R1	TAGCCAGGAGGTTCTCAACA (SEQ ID NO: 285)
GAPDH-F1	GGCATGGACTGTGGTCATGAG (SEQ ID NO: 286)
GAPDH-R1	TGCACCACCAACTGCTTAGC (SEQ ID NO: 287)
Reportero-1-qPCR-F	CCCCGTAATGCAGAAGAAGACC (SEQ ID NO: 288)
Reportero-1-qPCR-R	GTCCTTCAGCTTCAGCCTCTG (SEQ ID NO: 289)
PPIB-qPCR-F	AACGCAACATGAAGGTGCTC (SEQ ID NO: 290)
PPIB-qPCR-R	ACCTTGACGGTGACTTTGGG (SEQ ID NO: 291)
KRAS-qPCR-F	CAGTGCAATGAGGGACCACT (SEQ ID NO: 292)
KRAS-qPCR-R	AGGACCATAGGTACATCTTCAGAG (SEQ ID NO: 293)
SMAD4-qPCR-F	CGAACGAGTTGTATCACCTGGA (SEQ ID NO: 294)
SMAD4-qPCR-R	CGATGGCTGTCCCTCAAAGT (SEQ ID NO: 295)
FANCC-qPCR-F	AGTTGCTCTTTTCACTCAAGGTC (SEQ ID NO: 296)
FANCC-qPCR-R	TTCTCTCTGAGTTCAGACGCT (SEQ ID NO: 297)
PPIB-sec-profunda-F (sitio AAG)	TACACGACGCTCTTCCGATCTTAAGTAGAGTGGCACAGGAGGAAAGAGCAT C (SEQ ID NO: 298)
PPIB-sec-profunda-R (sitio AAG)	AGACGTGTGCTCTTCCGATCTTAAGTAGAGGCACCACCTCCATGCCCTC (SEQ ID NO: 299)
PPIB-sec-profunda-F (sitio CAG)	TACACGACGCTCTTCCGATCTTAAGTAGAGCATCGCAGACTGCGGCAAG (SEQ ID NO: 300)
PPIB-sec-profunda-R (sitio CAG)	AGACGTGTGCTCTTCCGATCTTAAGTAGAGAGTCCATGGGCCTGTGGAATG T (SEQ ID NO: 301)

FANCC-sec-profunda-F2 (sitio AAG/CAG)	GAAAACTGGCCCGAGAGC (SEQ ID NO: 302)
FANCC-sec-profunda-R2 (sitio AAG/CAG)	CTGAGTCTGGGCTGAGGGAC (SEQ ID NO: 303)
IDUA-sec-profunda-F	CGCTTCCAGGTCAACAACAC (SEQ ID NO: 304)
IDUA-sec-profunda-R	CTCGCGTAGATCAGCACCG (SEQ ID NO: 305)
p53-sec-profunda-F	CCCCTCTGAGTCAGGAAACAT (SEQ ID NO: 306)
p53-sec-profunda-R	GAAGATGACAGGGGCCAGG (SEQ ID NO: 307)
IFN- β -qPCR-F	TAGCACTGGCTGGAATGAG (SEQ ID NO: 308)
IFN- β -qPCR-R	GTTTCGGAGGTAACCTGTAAG (SEQ ID NO: 309)
ISG56-qPCR-F	TACAGCAACCATGAGTACAA (SEQ ID NO: 310)
ISG56-qPCR-R	TCAGGTGTTTCACATAGGC (SEQ ID NO: 311)
ISG54-qPCR-F	CTGCAACCATGAGTGAGAA (SEQ ID NO: 312)
ISG54-qPCR-R	CCTTTGAGGTGCTTTAGATAG (SEQ ID NO: 313)
IL-6-qPCR-F	GCCCTGAGAAAGGAGACAT (SEQ ID NO: 314)
IL-6-qPCR-R	CTGTTCTGGAGGTACTCTAGGTAT (SEQ ID NO: 315)
IL-8-qPCR-F	TTTGAAGAGGGCTGAGAA (SEQ ID NO: 316)
IL-8-qPCR-R	TGTTCTGGATATTTTCATGG (SEQ ID NO: 317)
RANTES-qPCR-F	CATCTGCCTCCCCATATTCC (SEQ ID NO: 318)
RANTES-qPCR-R	TCCATCCTAGCTCATCTCCAAA (SEQ ID NO: 319)
IL-12-qPCR-F	TGCTCCAGAAGGCCAGAC (SEQ ID NO: 320)
IL-12-qPCR-R	TTCATAAATACTACTAAGGCACAGG (SEQ ID NO: 321)
IL-1 β -qPCR-F	ACAGATGAAGTGCTCCTTCCA (SEQ ID NO: 322)
IL-1 β -qPCR-R	GTCGGAGATTCTGCTAGCTGGAT (SEQ ID NO: 323)
MCP1-qPCR-F	CATTGTGGCCAAGGAGATCTG (SEQ ID NO: 324)
MCP1-qPCR-R	CTTCGGAGTTTGGGTTTGCTT (SEQ ID NO: 325)
MIP1A-qPCR-F	CATCACTTGCTGCTGACACG (SEQ ID NO: 326)
MIP1A-qPCR-R	TGTGGAATCTGCCGGGAG (SEQ ID NO: 327)

IP10-qPCR-F	CTGACTCTAAGTGGCATT (SEQ ID NO: 328)
IP10-qPCR-R	TGATGGCCTTCGATTCTG (SEQ ID NO: 329)
GAPDH-qPCR-F2	CGGAGTCAACGGATTTGGTCGTA (SEQ ID NO: 330)
GAPDH-qPCR-R2	AGCCTTCTCCATGGTGGTGAAGAC (SEQ ID NO: 331)

Los productos de PCR se purificaron para secuenciación Sanger o NGS (Illumina HiSeq X Ten).

Secuenciación profunda

5 Para los experimentos de edición de ARNm endógeno, se sembraron células 293T-WT en placas de 6 pozos (6×10^5 células/pozo). Cuando hay aproximadamente un 70% de confluencia, las células HEK293 se transfectaron con 3 μ g de ARNd utilizando el reactivo de transfección de ADN X-tremeGENE HP (Roche). 72 horas más tarde, se clasificaron las células positivas para GFP o positivas para BFP (según el marcador de fluorescencia correspondiente) mediante FACS, seguido de
10 aislamiento de ARN. Luego, el ARN aislado se transcribió de forma inversa en ADNc mediante RT-PCR, y se amplificó el locus del gen objetivo específico con los pares de cebadores correspondientes (23 ciclos de PCR) y se secuenciaron en un Illumina NextSeq.

Prueba en múltiples líneas celulares

Además de las células HEK293T (control positivo) y HEK293T ADAR1^{-/-} (control
15 negativo), una línea celular de ratón (NIH3T3) y siete líneas celulares humanas (RD, HeLa, SF268, A549, HepG2, HT-29, SW13) se originan de diferentes tejidos y órganos fueron seleccionados para realizar el experimento. Para las líneas celulares con mayor eficiencia de transfección, se sembraron en placa aproximadamente $8-9 \times 10^4$ células (RD, HeLa, SF268) o 1.5×10^5 (HEK293T) en cada pozo de la placa de 12 pozos, como para las (A549, HepG2, HT-29, SW13, NIH3T3) que son difíciles de transfectar,
20 se sembraron $2-2.5 \times 10^5$ células en placas de 6 pozos. Y todas estas células se mantuvieron en medio Eagle modificado de Dulbecco (DMEM, Corning) suplementado con suero bovino fetal al 10% (FBS, CellMax) con CO₂ al 5% a 37°C. 24 horas más tarde, el reportero CG2 y el plásmido de ARNd 71nt (35-C-35) se cotransfectaron en diferentes tipos de células con el reactivo de transfección de ADN X-tremeGENE HP (Roche). 48 horas después de la transfección, las células se tripsinizaron y analizaron
25 mediante FACS (BD). Debido a que las células con baja eficiencia de transfección tenían bastante menos células positivas para mCherry y BFP, aumentamos el número total de células para el análisis FACS a 1×10^5 para las células colocadas en placas de 6 pozos.

Análisis de edición de ARN para sitios específicos

Para el análisis de secuenciación profunda, se generó un índice utilizando la secuencia
30 del sitio objetivo (corriente arriba y corriente abajo 20-nt) de las secuencias que cubren el ARNar. Las

lecturas se alinearon y cuantificaron utilizando BWA versión 0.7.10-r789. Samtools clasificó las BAM de alineación y los sitios de edición de ARN se analizaron utilizando REDitools versión 1.0.4. Los parámetros son los siguientes: -U [AG o TC] -t 8 -n 0.0 -T 6-6 -e -d -u. Toda la conversión A>G significativa dentro de la región de selección de ARNar calculada mediante la prueba exacta de Fisher (valor de $p < 0.05$) se consideró como ediciones por ARNar. Las conversiones, a excepción de la adenosina dirigida, fueron ediciones fuera del objetivo. Las mutaciones que aparecieron en los grupos control y experimental simultáneamente fueron consideradas como SNP.

Análisis de secuenciación de ARN en todo el transcriptoma

Los plásmidos que expresan ARN₁₅₁ Ctrl o ARNar₁₅₁-PPIB con casete de expresión de BFP se transfectaron en células HEK293T. Las células BFP⁺ se enriquecieron con FACS 48 horas después de la transfección y los ARN se purificaron con el kit ARNprep Pure Micro (TIANGEN, DP420). A continuación, se purificó el ARNm utilizando el módulo de aislamiento magnético de ARNm de NEBNext Poli(A) (New England Biolabs, E7490), procesado con el kit de preparación de bibliotecas de ARN NEBNext Ultra II para Illumina (New England Biolabs, E7770), seguido de un análisis de secuenciación profunda con la plataforma Illumina HiSeq X Ten (extremo emparejado 2x150 pb; 30G para cada muestra). Para excluir el efecto inespecífico causado por la transfección, incluimos el grupo simulado en el que solo tratamos células con reactivo de transfección. Cada grupo contenía cuatro réplicas.

La tubería de análisis bioinformático fue referido al trabajo de Vogel et al²². El control de calidad del análisis se realizó mediante FastQC, y el recorte de calidad se basó en Cutadapt (los primeros 6 pb de cada lectura se recortaron y hasta 20 pb se recortaron en calidad). Se utilizaron textos AWK para filtrar los ARNar introducidos. Después de recortar, se filtraron las lecturas con longitudes inferiores a 90-nt. Posteriormente, las lecturas filtradas se mapearon en el genoma de referencia (GRCh38-hg38) mediante el software STAR⁶¹. Usamos el GATK Haplotypecaller⁶² para llamar a las variantes. Los archivos VCF sin procesar generados por GATK fueron filtrados y anotados por GATK VariantFiltration, bcftools y ANNOVAR⁶³. Se filtraron las variantes en dbSNP, 1000 Genoma⁶⁴, EVS. Las variantes compartidas en cuatro réplicas de cada grupo se seleccionaron luego como los sitios de edición de ARN. El nivel de edición de ARN del grupo simulado se consideró como el fondo, y los objetivos globales de ARN₁₅₁ Ctrl y ARNar₁₅₁-PPIB se obtuvieron restando las variantes en el grupo simulado.

Para evaluar si LEAPER perturba la homeostasis de edición natural, analizamos los sitios de edición global compartidos por el grupo simulado y el grupo ARNar₁₅₁-PPIB (o grupo ARN₁₅₁ Ctrl). Las tasas de edición de ARN diferencial en los sitios de edición nativos de A a I se evaluaron con el análisis del coeficiente de correlación de Pearson. Las correlaciones de Pearson de la tasa de edición entre el grupo simulado y el grupo ARNar₁₅₁-PPIB (o grupo ARN₁₅₁ Ctrl) se calcularon y anotaron en la

figura 6.

$$\rho(X,Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

X significa la tasa de edición de cada sitio en el grupo simulado; Y significa la tasa de edición de cada sitio en el grupo ARN₁₅₁ Ctrl (figura 6A) o el grupo ARN₁₅₁-PPIB (figura 6B); σ_X es la desviación estándar de X; σ_Y es la desviación estándar de Y; μ_X es la media de X; μ_Y es la media de Y; E es la expectativa.

Los datos de ARN-Sec se analizaron para el interrogatorio de posibles cambios transcripcionales inducidos por eventos de edición de ARN. El análisis de la expresión génica de todo el transcriptoma se realizó utilizando el software HISAT2 y STRINGTIE⁶⁵. Usamos Cutadapt y FastQC para el control de calidad de los datos de secuenciación. Las lecturas de secuenciación se mapearon luego al genoma de referencia (GRCh38-hg38) utilizando HISAT2, seguido del análisis del coeficiente de correlación de Pearson como se mencionó anteriormente.

Western blot

Usamos los anticuerpos primarios monoclonales de ratón respectivamente contra ADAR1 (Santa Cruz, sc-271854), ADAR2 (Santa Cruz, sc-390995), ADAR3 (Santa Cruz, sc-73410), p53 (Santa Cruz, sc-99), KRAS (Sigma, SAB1404011); GAPDH (Santa Cruz, sc-47724) y β -tubulina (CWBioTech, CW0098). El anticuerpo secundario IgG de cabra anti-ratón conjugado con HRP (H+L, 115-035-003) se adquirió de Jackson ImmunoResearch. Se clasificaron 2×10^6 células para su lisis y se cargó una cantidad igual de cada lisado para SDS-PAGE. Luego, las proteínas de muestra se transfirieron a la membrana de PVDF (Bio-Rad Laboratories) y se inmunotransfirieron con anticuerpos primarios contra una de las enzimas ADAR (anti-ADAR1, 1:500; anti-ADAR2, 1:100; anti-ADAR3, 1:800), seguido de incubación de anticuerpos secundarios (1:10,000) y exposición. La β -tubulina se volvió a sondear en la misma membrana de PVDF después de eliminar las proteínas ADAR con el regulador de pH de extracción (CWBioTech, CW0056). Los experimentos se repitieron tres veces. El análisis semicuantitativo se realizó con el software Image Lab.

Ensayo de expresión de citocinas

Se sembraron células HEK293T en placas de 12 pozos (2×10^5 células/pozo). Cuando aproximadamente un 70% de confluencia, las células se transfectaron con 1.5 μ g de ARN₁₅₁. Como un control positivo, se transfectó 1 μ g de poli(I: C) (Invitrogen, tlr-picw). Cuarenta y ocho horas después, las células se recolectaron y se sometieron a aislamiento de ARN (TIANGEN, DP430). Luego, los ARN totales se transcribieron de forma inversa en ADNc mediante RT-PCR (TIANGEN, KR103-04) y se midió la expresión de IFN- β e IL-6 mediante PCR cuantitativa (TAKARA, RR820A). Las secuencias de los cebadores se enumeraron en el cuadro anterior.

Ensayo de actividad reguladora transcripcional de p53

Los plásmidos que expresan ADNc de TP53^{W53X} y los plásmidos que expresan ARNar se cotransfectaron en células HEK293T TP53^{-/-}, junto con plásmidos reporteros en cis p53-Luciérnaga-luciferasa (YRGene, VXS0446) y plásmidos Renilla-luciferasa (un regalo del Laboratorio de Z. Jiang, Universidad de Pekín) para detectar la actividad reguladora transcripcional de p53. 48 horas después, las células se recolectaron y analizaron con el sistema de ensayo de luciferasa Promega Dual-Glo (Promega, E4030) de acuerdo con el protocolo del fabricante. Brevemente, se agregaron 150 µl de reactivo de luciferasa Dual-Glo al sedimento celular cosechado y, 30 minutos después, se midió la luminiscencia de luciérnaga agregando 100 µl de reactivo de luciferasa Dual-Glo (lisis celular) a una placa blanca de 96 pozos con el lector Infinite M200 (TECAN). 30 minutos más tarde, se añadieron secuencialmente 100 µl de terminación Dual-Glo y reactivo Glo a cada pozo para medir la luminiscencia de Renilla y calcular la relación entre la luminiscencia de luciérnaga y la luminiscencia de Renilla.

Electroporación en celdas primarias

Para la electroporación de plásmidos que expresan ARNar en fibroblastos pulmonares primarios humanos o células epiteliales bronquiales primarias humanas, se electroporaron 20 µg de plásmidos con el dispositivo Nucleofector™ 2b (Lonza) y el kit Basic Nucleofector™ (Lonza, VPI-1002), y el programa de electroporación fue U-023. Para la electroporación de plásmidos que expresan ARNar en células T primarias humanas, se electroporaron 20 µg de plásmidos en T primaria humana con el dispositivo Nucleofector™ 2b (Lonza) y el kit Nucleofector™ de células T humanas (Lonza, VPA-1002), y el programa de electroporación fue T-024. Cuarenta y ocho horas después de la electroporación, las células se clasificaron y recolectaron mediante el ensayo FACS y luego se sometieron a la siguiente secuenciación profunda para el ensayo de edición de ARN dirigido. La eficacia de la electroporación se normalizó según el marcador de fluorescencia.

Para la electroporación de ARNar quimiosintético o ARN de control en células T primarias humanas o células GM06214 primarias, se disolvió oligo de ARN en 100 µl de medio opti-MEM (Gibco, 31985070) con la concentración final 2 µM. A continuación, se resuspendieron 1 x 10E6 células GM06214 o 3 x 10E6 células T con la mezcla de electroporación anterior y se electroporaron con dispositivo Agile Pulse In Vivo (BTX) a 450 V durante 1 ms. Luego, las células se transfirieron a un medio de cultivo caliente para los siguientes ensayos.

Ensayo de actividad catalítica de α-L-iduronidasa (IDUA)

El sedimento de células recolectadas se resuspendió y se lisó con 28 µl de Triton X-100 al 0.5% en regulador de pH 1xPBS en hielo durante 30 minutos. Y luego se agregaron 25 µl de la lisis celular a 25 µl de sustrato de 4-metilumbeliferil-α-L-iduronidasa 190 µM (Cayman, 2A-19543-500), que se disolvió en regulador de pH de formiato de sodio 0.4 M que contenía Triton X-100 al 0.2%, pH

Prueba del método de edición de ARN de la invención en un reportero

El reportero-1 de fluorescencia dual comprende la secuencia de mCherry (SEQ ID NO:1), la secuencia que comprende el enlazador 3xGS y la A dirigida (SEQ ID NO:2), y la secuencia de eGFP (SEQ ID NO:3).

30 CTGCAG ***GGCGGAGGAGGCAGCGGCGGAGGAGGCAGCGGCGGAGGAGGCAGC***AGAAGGTATACACG
CCGGAAGAATCTGT**A**GAGATCCCCGGTCGCCACC (secuencia que comprende el enlazador 3×GS (se
muestra como caracteres en cursiva y negrita) y la A dirigida (se muestra como A más grande y en
negrita)) (SEQ ID NO:2)

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC
35 CACAAGTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCA

CCACCGGAAGCTGCCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTCAGCCG
 CTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACC
 ATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACC
 GCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAA
 5 CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACA
 TCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC
 TGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGT
 CCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA (secuencia de
 eGFP) (SEQ ID NO:3)

- 10 El reportero-2 de fluorescencia dual comprende la secuencia de mCherry (SEQ ID NO: 1), la secuencia que comprende el enlazador 3×GS (mostrada en cursiva y negrita) y la A dirigida (mostrada como A más grande y en negrita) (SEQ ID NO:4) , y secuencia de eGFP (SEQ ID NO: 3).

CTGCAG ***GGCGGAGGAGGCAGCGCGGAGGAGGCAGCGCGGAGGAGGCAGCGC***
 CTGCTCGCGATGCT***AGAGGGCTCTGCCA*** (secuencia que comprende el enlazador 3xGS (se muestra
 15 como caracteres en cursiva y negrita) y la A dirigida (se muestra como caracteres más grandes y en negrita)) (SEQ ID NO:4))

El reportero-3 de fluorescencia dual comprende la secuencia de mCherry (SEQ ID NO: 1), la secuencia que comprende el enlazador 1xGS (que se muestra como caracteres en cursiva y negrita) y la A dirigida (SEQ ID NO:5), y la secuencia de eGFP (SEQ ID NO:3).

- 20 CTGCAG ***GGCGGAGGAGGCAGCGC***CTGCTCGCGATGCT***AGAGGGCTCTGCCA***
 (secuencia que comprende el enlazador 1xGS (que se muestra en cursiva y negrita) y la A dirigida (que se muestra como A más grande y en negrita)) (SEQ ID NO:5)

Clonamos mCherry-enlazador 3×GS-TAG-EGFP en pLenti-estructura fundamental, y el plásmido reportero se empaquetó en lentivirus, que infectó células 293T que construían una línea celular estable que expresaba el reportero de fluorescencia dual. Luego, seleccionamos un solo clon con un fondo de fluorescencia de EGFP bajo como sistema reportero. Colocamos en mosaico guía de ARNcr de Lbuc2c2 con espaciadores de 28 a 78 nucleótidos de largo a través de la adenosina dirigida para probar el diseño de ARNcr óptimo. Encontramos que las guías de ARNcr más largas conferían una mayor eficiencia positiva de EGFP. Sorprendentemente, cuando transfectamos plásmidos de ARNcr dirigidos sin cotransfección de ningún plásmido que exprese dC2c2-ADARDD, la proteína EGFP se expresa sustancialmente. Por ejemplo, la guía de ARNcr con la secuencia:
 25 GGACCACCCCAAAAUGAAUAUAACCAAAACUGAACAGCUCCUCGCCCUGCUCACUGGCAGAGCCCUCCAG
 CAUCGCGAGCAGGCGCUGCCUCCUCCGCC (SEQ ID NO: 6) confería una eficiencia positiva de EGFP superior al 25%. Esto indica que la adenina en el codón de terminación UAG se edita en gran medida.
 30 Por el contrario, el ARNcr aleatorio no pudo convertir las células EGFP negativas en positivas (figuras

6A, 6B y 6C). Basándonos en estos resultados, inferimos que la sobreexpresión de un transcripto de ARN por sí solo podría aprovechar la enzima ADAR endógena para editar el ARN.

Además, eliminamos la secuencia de ARN de andamio en las guías de ARN, creando un ARN guía lineal. Encontramos ARN de 70 nucleótidos de longitud
 5 (AAACCGAGGGAUCAUAGGGGACUGAAUCCACCAUUCUUCUCCCAAUCCUGCAACUCCUUCUCCCCUGC(SEQ ID NO: 7)) complementario al ARN dirigido con un desajuste A-C podría convertir de manera eficiente las células negativas EGFP en células positivas EGFP, mientras que el ARN aleatorio 70-nt (UGAACAGCUCCUCGCCCUGCUCACUGGCAGAGCCUCCAGCAUCGCGAGCAGGCGUGCCUCCUCCGCC(SEQ ID NO: 8)) no pudo (figuras 1A, 1B, 1C y 1D). Por lo tanto, designamos este ARN como ARNd
 10 (ARN de reclutamiento de desaminasas). Para verificar que la ADAR endógena celular podría ser reclutada para llevar a cabo la desaminación de la adenina por el ARNd, realizamos experimentos en las líneas celulares 293T de doble desactivadas Adar1 p110 y Adar1 p150 (figuras 6E y 6F). Porque Adar1 se expresa de forma ubicua, mientras que Adar2 se expresa principalmente en el cerebro a alto nivel. Así que propusimos que la desaminación de adenina dirigida por ARNd estaba mediada
 15 principalmente por Adar1 pero no por Adar2. Como se esperaba, el ARNd dirigido no podría desencadenar la expresión de EGFP en células 293T-Adar1-/-, pero la sobreexpresión de Adar1 p110, p150 o Adar2 exógena podría rescatar la expresión de EGFP en células 293T-Adar1-/- (figuras 1E y 1F), lo que sugiere que en las células 293T, el ARNd podría reclutar Adar1 o Adar2 para mediar en la desaminación de adenina en un ARN objetivo. Además, encontramos que Adar1-p110 y Adar2 tienen
 20 una mayor actividad de edición que Adar1-p150 (figura 1G y la figura 6G), posible debido a la diferente localización celular de Adar1-p110 y Adar1-p150.

Para determinar que la restauración de la fluorescencia de EGFP se debió a los eventos de edición de ARN de direccionamiento, medimos directamente la edición mediada por ARNd de los transcriptos del reportero-2 a través de RT-PCR seguida de secuenciación de Sanger dirigida y
 25 secuenciación de próxima generación. Los resultados de la secuenciación mostraron la conversión de bases A a G en la adenina objetivo (sitio de desajuste A-C) y la tasa de edición podría alcanzar el 13% (figura 6H y la figura 1H). Además, también observamos una edición leve de A a G durante las ventanas de secuencia cerca de la adenina dirigida, muy posiblemente debido al aumento de las regiones de ARN dúplex, más tarde, intentaríamos deshacernos de la edición inesperada con varias
 30 estrategias.

EJEMPLO 2

Optimización de los factores para diseñar ARNd

35 A continuación, nos propusimos optimizar el ARNd para lograr una mayor eficiencia de

edición. Primero, nuestro objetivo fue determinar qué base en el sitio opuesto de la adenina objetivo favorece la edición. Estudios anteriores mostraron que la base opuesta de la adenosina dirigida afectaría la edición de manera eficiente. Por lo tanto, diseñamos ARNd 71nt con un desajuste N (A, U, C y G) en la posición media opuesta al objetivo A. Según los resultados de FACS, encontramos que los

5 cuatro ARNd diferentes se editan de manera eficiente de la siguiente manera: C> A> U> G (figuras 2A y 2B). Recientemente, se ha informado de que una pequeña burbuja en el sitio UAG objetivo puede ser beneficiosa para la eficiencia de la edición. Por lo tanto, diseñamos ARNd que contienen dos o tres bases de desajuste con el sitio UAG objetivo para probar nuestra hipótesis. Se diseñaron y construyeron 16 ARNd 71-nt diferentes en el vector ARNd con marcador BFP utilizando el método de

10 clonación Golden Gate. Encontramos que los ARNd con secuencia CCA y GCA son de la mayor eficiencia, lo que significa que la pequeña burbuja contribuye poco a la edición A-I, al menos en el caso del sitio objetivo UAG. Además, cuatro ARNd de la secuencia de NCA tienen un mayor porcentaje de células positivas para GFP, lo que lleva a la conclusión de que el par de bases U-A complementarias puede ser importante para la edición de ADAR (figuras 2C y 2D). Posteriormente, probamos la eficiencia de

15 diferentes longitudes de ARNd basados en el reportero. Los ARNd se diseñaron con un desajuste C en la posición media con diferentes longitudes que van desde 31-nt a 221-nt. Descubrimos que la eficiencia de edición aumenta con un ARNd más largo. El pico de edición del sistema reportero se encuentra en ARNd 171-nt. El ARNd de 51-nt podría iluminar el sistema indicador con una buena eficiencia (18%) (figuras 2E y 2F). Finalmente, examinamos si la posición del desajuste C del ARNd

20 afecta la eficiencia de edición. Los ARNd se mantuvieron con la misma longitud de 71-nt, se diseñó un desajuste C en una posición diferente desde el comienzo de la transcripción. Basándonos en los resultados de FACS, encontramos que la ubicación del desajuste opuesto C podría afectar la eficiencia de edición, y el desajuste C ubicado en 5' o 3' del ARNd tiene una eficiencia menor (figuras 2G y 2H).

Se construyeron 16 reporteros diferentes que comprenden secuencias objetivo que

25 contienen todos los posibles motivos de 3 bases mediante clonación de Gibson, y luego se clonaron en la estructura fundamental pLenti (pLenti-CMV-MCS-SV-Bsd, Stanley Cohen Lab, Universidad de Stanford). Las secuencias objetivo se muestran a continuación.

Secuencias objetivo que contienen todos los motivos de 3 bases posibles:

TAT:

30 ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCTATAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 9)

TAA:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCTAAAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 10)

35 TAC:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCTACAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 11)

TAG:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCTAGAGGGCTCTGCCAGTG
5 AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 12)

AAT:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCAATAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 13)

AAA:

10 ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCAAAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 14)

AAC:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCAACAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 15)

15 AAG:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCAAGAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 16)

CAT:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCATAGGGCTCTGCCAGTG
20 AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 17)

CAA:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCAAAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 18)

CAC:

25 ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCACAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 19)

CAG:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCAGAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 20)

30 GAT:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCATAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 21)

GAA:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCCAAAGGGCTCTGCCAGTG
35 AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 22)

GAC:

ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCGACAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 23)

GAG:

- 5 ATGGACGAGCTGTACAAGCTGCAGGGCGGAGGAGGCAGCGCCTGCTCGCGATGCGAGAGGGCTCTGCCAGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC (SEQ ID NO: 24)

Los ARNd se mantuvieron con la misma longitud de 111 pb y diseñaron un desajuste C en el centro hacia el objetivo A.

- En el grupo de cultivo celular de 12 pozos, se sembraron en placa 2×10^5 células
10 HEK293T en cada pozo y cada experimento se realizó para tres repeticiones. 24 horas después, se
cotransfectaron 0.5 μ g de plásmido de ARNd y 0.5 μ g de plásmido objetivo reportero a las células
utilizando el reactivo de transfección de DNA X-tremeGENE HP (Roche). 48 horas más tarde, las células
se tripsinizaron y se seleccionaron para células positivas para mCherry a través de FACS (BD). Se
recogió un total de 4×10^5 células y se extrajo el ARN total usando el kit de células/bacterias ARNprep
15 puro (TIANGEN DP430). Los ADNc se sintetizaron a partir de 2 μ g de ARN total usando el kit
Quantscript RT (TIANGEN KR103-04). Y las 111 regiones objetivo se amplificaron mediante PCR y se
enviaron para una secuenciación profunda.

- Encontramos que los 16 motivos de 3 bases diferentes pueden editarse a través de un
método de edición de ARN ejemplar de la presente solicitud, aunque con una eficiencia variable. En
20 resumen, los resultados indican que el vecino más cercano 5' de A que se va a editar tiene la
preferencia $U > C \approx A > G$ y el vecino más cercano 3' de A que se va a editar tiene la preferencia
 $G > C > A \approx U$. Los datos se presentaron como gráfico de barras en la figura 3A o mapa de calor de la
figura 3B.

25

EJEMPLO 3

Edición de ARN transcrito a partir de genes endógenos

- A continuación, probamos si el ARNd pudiera mediar el ARNm transcrito de genes
endógenos. Diseñamos ARNd dirigido a cuatro genes KRAS, PPIB, β -Actina y GAPDH. Para el ARNm de
30 KRAS, diseñamos ARNd de 91, 111, 131, 151, 171 y 191 nucleótidos de longitud (figura 4A) con
secuencias como se muestra a continuación.

91-nt KRAS-ARNd

UAGCUGUAUCGUCAAGGCACUCUUGCCUACGCCACCAGCUCCAACCACCACAAGUUUAUAUUCAGUCAUUU
UCAGCAGGCCUCUCUCCCGC (SEQ ID NO: 25)

- 35 111-nt KRAS-ARNd

GAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCUUGCCUACGCCACCAGCUCCAACUACCACAAGUUUAUUAU
UCAGUCAUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGC (SEQ ID NO: 26)

131-nt KRAS-ARNd

UCCACAAAAUGAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCUUGCCUACGCCACCAGCUCCAACUACCAC
5 AAGUUUAUUAUUCAGUCAUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGCCGCUGAGCCU (SEQ ID NO:
27)

151-nt KRAS-ARNd

AUCAUAUUCGUCCACAAAAUGAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCUUGCCUACGCCACCAGCUC
CAACCACCACAAGUUUAUUAUUCAGUCAUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGCCGCUGAGCCUC
10 UGGCCCCGC (SEQ ID NO: 28)

171-nt KRAS-ARNd

CUAUUGUUGGAUCAUAUUCGUCCACAAAAUGAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCUUGCCUACG
CCACCAGCUCCAACCACCACAAGUUUAUUAUUCAGUCAUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGCC
GCUGAGCCUCUGGCCCCGCCGCCGCCUUC (SEQ ID NO: 29)

15 191-nt KRAS-ARNd

UAGGAAUCCUCUAUUGUUGGAUCAUAUUCGUCCACAAAAUGAUUCUGAAUUAGCUGUAUCGUCAAGGCACU
CUUGCCUACGCCACCAGCUCCAACCACCACAAGUUUAUUAUUCAGUCAUUUUCAGCAGGCCUCUCUCCCGCA
CCUGGGAGCCGCUGAGCCUCUGGCCCCGCCGCCGCCUUCAGUGCCUGCG (SEQ ID NO: 30)

Los resultados de la secuenciación de próxima generación mostraron que los ARNd podrían editar el
20 ARNm de KRAS objetivo con hasta un 11.7% de eficiencia de edición (figura 4B). Para el ARNm de PPIB
endógeno, los tres sitios se enfocaron: sitio1, sitio2 y sitio3. Diseñamos ARNd de 151 nucleótidos de
largo para cada sitio (figura 4C) con secuencias como se muestra a continuación.

PPIB-ARNd de 151-nt (sitio 1)

GAGGCGCAGCAUCCACAGGCGGAGGCGAAAGCAGCCCGGACAGCUGAGGCCGGAAGAGGGUGGGGCCGCG
25 GUGGCCAGGGAGCCGGCGCCGCCACGCGCGGGUGGGGGGACUGGGGUUGCUCGCGGGCUCGGGGCGGG
CGGCGGGCGCCG (SEQ ID NO: 31)

PPIB-ARNd de 151-nt (sitio 2)

UCCUGUAGCUAAGGCCACAAAAUUAUCCACUGUUUUUGGAACAGUCUUUCCGAAGAGACCAAAGAUCACCC
GGCCCACAUCUUAUCUCCAUAUUCGUAGGUCAAAUACACCUUGACGGUGACUUUGGGCCCCUUCUUCUUC
30 UCAUCGGCC (SEQ ID NO: 32)

PPIB-ARNd de 151-nt (sitio 3)

GCCUGGAUCAUGAAGUCCUUGAUUACACGAUGGAAUUUGCUGUUUUUGUAGCCAAAUCCUUCUCUCCU
GUAGCCAAGGCCACAAAAUUAUCCACUGUUUUUGGAACAGUCUUUCCGAAGAGACCAAAGAUCACCCGGCC
UACAUCUUA (SEQ ID NO: 33)

35 Los resultados de la secuenciación de próxima generación mostraron que el ARNd podría editar el sitio

1 del ARNm de PPIB de manera eficiente con una tasa de edición de hasta el 14% (figura 4D). Para el sitio 2 y el sitio 3 del ARNm de PPIB, la eficiencia de edición fue del 1.5% y del 0.6% (figura 4E y 4F). Para el ARNm de β -actina endógeno, seleccionamos dos sitios objetivo y diseñamos ARNd para cada sitio (figura 4G) con secuencias como se muestra a continuación.

5 72-nt β -Actina-ARNd (sitio 1)

GCGCAAGUUAGGUUUUGUCAAGAAAGGGUGUAACGCAACCAAGUCAUAGUCCGCCUAGAAGCAUUUGCGG
UG (SEQ ID NO: 34)

131-nt β -Actina-ARNd (sitio 1)

10 GCCAUGCCAAUCUCAUCUUGUUUUCUGCGCAAGUUAGGUUUUGUCAAGAAAGGGUGUAACGCAACCAAGU
CAUAGUCCGCCUAGAAGCAUUUGCGGUGGACGAUGGAGGGGCCGACUCGUCAUACUCCUG (SEQ ID NO:
35)

70-nt β -Actina-ARNd (sitio 2)

GGACUCCUGUAACAACGCAUCUCAUUAUUUGGAAUGACCAUUAACAAACAAUGUGCAAUCAAAGUC
(SEQ ID NO: 36)

15 Encontramos que el ARNd podría editar el ARNm de β -actina tanto en el sitio 1 como
en el sitio 2, con una eficiencia de edición de hasta el 1.4% para cada sitio (figura 4H y la figura 8A).
También observamos que el ARNd más largo confería una mayor eficiencia de edición, con un 0.6%
para el ARNd-71nt y un 1.4% para el ARNd-131nt (figura 3H). Para otro gen de limpieza GAPDH,
usamos ARNd 71nt
20 (CAAGGUGCGGCUCGCGCCCUCCCUUUAAGGGGUCCACAUGGCAACUGUGAGGAGGGGAGAUUCAGU
G (SEQ ID NO: 37)), y la eficiencia de edición es del 0.3%, tal vez debido a la corta longitud del ARNd
(figura 8B).

EJEMPLO 4

25 **Análisis fuera del objetivo en un método LEAPER ejemplar**

Para la aplicación terapéutica, la precisión de la edición es fundamental. A
continuación, intentamos caracterizar la especificidad de un sistema de edición de ARN ejemplar de la
presente solicitud. Seleccionamos el sitio PPIB endógeno 1 y el sitio KRAS para el análisis. Para el sitio
30 1 de PPIB, pudimos ver que durante las regiones cubiertas de ARNd, había varias bases A flanqueando
el A76 objetivo, como A22, A30, A33, A34, A39, A49, A80, A91, A107 y A140. Reveló que las bases A
flanqueantes apenas se editaron, mientras que la base A76 objetivo (desajuste A-C) mostró hasta un
14% de eficiencia de edición (figuras 5A y 5B).

En cuanto al sitio KRAS, pudimos ver en la región cubierta de ARNd, hay muchas
35 adeninas flanqueando la base A56 dirigida, hasta 29 bases A flanqueantes. A partir de los resultados

de la edición de ARNm de KRAS, encontramos que mientras que la base A56 dirigida (desajuste A-C) mostraba hasta un 11.7% de eficiencia de edición, la adenina flanqueante podía editarse (figuras 5C y 5D). Se editaron una variedad de adeninas no dirigidas, mientras que las adeninas como A41, A43, A45, A46, A74, A79 mostraron más edición. Encontramos que el vecino 5' más cercano de esas bases A no editadas era G o C, mientras que el vecino 5' más cercano de esas adeninas editadas eficientemente era T o A. Con base en esta observación, nos propusimos diseñar ARNd para minimizar el fuera de destino edición de aquellas adeninas que son propensas a ser editadas. En nuestro estudio, hemos encontrado que ADAR prefería el desajuste A-C a A-A, A-U, y el desajuste A-G era el menos preferido. Por lo tanto, propusimos que para las bases A fuera de la focalización en las que el vecino más cercano de 5' era U o A, la falta de desajuste A-G podría reducir o disminuir los efectos fuera de la focalización. Un estudio anterior informó que el desajuste A-G podría bloquear la edición de desaminación por ADAR.

Así que a continuación diseñamos tres tipos de variantes de ARNd 91-nt y cuatro tipos de variantes de ARNd 111-nt (con secuencias como se muestra a continuación) que contienen diferentes combinaciones de desajustes A-G basados en los resultados estadísticos de la figura 5D y conocimiento existente: ARNd-AG1 (A41, A46, A74); ARNd-AG2 (A41, A43, A45, A46, A74, A79); ARNd-AG3 (A31, A32, A33, A41, A43, A45, A46, A47, A74, A79); ARNd-AG4 (A7, A31, A32, A33, A40, A41, A43, A45, A46, A47, A74, A79, A95) (figura 4E).

KRAS-ARNd-91-AG2

UAGCUGUAUCGUCAAGGCACUCgUGCCgACGCCACCAGCUCCAACcACCACAAGgggAgAgUCAGUCAgggUC
AGCAGGCCUCUCUCCCGC (SEQ ID NO: 38)

KRAS-ARNd-91-AG3

UAGCUGUAUCGUCAAGGCACUCUUGCCgACGCCACCAGCUCCAACcACCACAAGUgUAUAUgUCAGUCAUUUU
CAGCAGGCCUCUCUCCCGC (SEQ ID NO: 39)

25 KRAS-ARNd-91-AG4

UAGCUGGAUCGUCAAGGCACUCGUGCCGACGCCACCAGCUCCAACCACCACAAGGGGAGAGGCAGUCAGGG
UCAGCAGGCCUCUCUCCCGC (SEQ ID NO: 40)

KRAS-ARNd-111-AG1

GAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCUUGCCgACGCCACCAGCUCCAACcACCACAAGUgUAUAUgU
30 CAGUCAUUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGC (SEQ ID NO:41)

KRAS-ARNd-111-AG2

GAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCgUGCCgACGCCACCAGCUCCAACcACCACAAGUggAgAgU
CAGUCAUUUUUCAGCAGGCCUCUCUCCCGCACCUGGGAGC (SEQ ID NO:42)

KRAS-ARNd-111-AG3

35 GAUUCUGAAUUAGCUGUAUCGUCAAGGCACUCgUGCCgACGCCACCAGCUCCAACcACCACAAGgggAgAgU

CAGUCAgggUCAGCAGGCCUCUCUCCCGCACCUGGGAGC (SEQ ID NO:43)

KRAS-ARNd-111-AG4

GCUCCCCGGUGCGGGAGAGAGGCCUGCUGACCCUGACUGCCUCUCCCCUUGUGGUGGUUGGAGCUGGUGG

CGUCGGCACGAGUGCCUUGACGAUCCAGCUAAUUCAGAAUC (SEQ ID NO: 44)

- 5 Luego, estos ARNds se transfectaron en células HEK293T y un vector vacío y un control de ARNd no dirigido a 71-nt:

(TCTCAGTCCAATGTATGGTCCGAGCACAAGCTCTAATCAAAGTCCGCGGGTGTAGACCGGTTGCCATAGGA

- (SEQ ID NO: 45)) se utilizaron como controles negativos. Para los ARNd 91-nt, los resultados de la secuenciación profunda mostraron que la edición en el objetivo (A56) se redujo al 2.8% para el ARNd-91-AG2, al 2.3% para el ARNd-91-AG3 y al 0.7% para el ARNd-91-AG4, en comparación con la eficiencia de edición en el objetivo (A56) del 7.9% para el ARNd-91 sin desajuste de A-G (figura 4F). Para los ARNd 91-nt, la edición en el objetivo (A56) se redujo al 5.1% para el ARNd-111-AG2 y al 4.9% para el ARNd-111-AG3 en comparación con la eficiencia de la edición en el objetivo (A56) del 15.1% para el ARNd- 111 sin desajuste de A-G (figura 4F), lo que indica un ARNd más largo que podría tener más desajustes A-G. Entonces, a continuación, seleccionamos ARNd 111-nt para un análisis detallado fuera del objetivo. La edición de las bases A flanqueantes se borró drásticamente excepto para A7 y A79 (figura 4G). Para la base A7, el efecto fuera del objetivo podría prevenirse mediante un diseño adicional de desajuste A-G en este sitio, que está ausente en el diseño actual de ARNd. Para la base A79, la introducción de dos desajustes A-G adyacentes A78/A79 podría ayudar a eliminar los efectos fuera del objetivo. Basándose en tales resultados, la aplicación de los sistemas de edición de ARN de la presente solicitud para curar enfermedades genéticas es muy prometedora y alentadora.

EJEMPLO 5

Probando un método LEAPER ejemplar en múltiples líneas celulares

25

- A través de los resultados en células HEK293T, supusimos que el ARN de doble cadena formado por el ARNd lineal y su ARN objetivo podría reclutar una proteína ADAR endógena para la edición A-I. Para confirmar la hipótesis, elegimos más líneas celulares para probar nuestro método de edición de ARN. Los resultados se muestran en la figura 9. Esos resultados en múltiples líneas celulares demostraron la universalidad de nuestro método de edición de ARN. En primer lugar, a pesar de las diversas eficiencias de edición, el uso de ARNd para reclutar ADAR endógena fue adecuado para múltiples líneas celulares humanas, que se originó a partir de 7 tejidos y órganos diferentes. Además, este método no solo podría funcionar en células humanas, sino también en células de ratón, lo que brinda la posibilidad de realizar experimentos en un ratón.

35

EJEMPLO 6

Aprovechamiento de ADAR endógena para la edición de ARN

En un intento por explorar una plataforma de edición de ARN eficiente, fusionamos el dominio desaminasa del mutante E1008Q hiperactivo ADAR1 (ADAR1^{DD})⁴⁰ con el LbuCas13 inactivo catalítico (dCas13a), un efector CRISPR dirigido a ARN guiado por ARN⁴¹ (figura 10A). Para evaluar la eficiencia de edición de ARN, construimos un reportero sustituto que alberga genes *mCherry* y EGFP ligados por una secuencia que comprende una región codificante de 3x GGGGS y un codón de terminación UAG en marco (reportero-1, figura 10B). Las células transfectadas con reportero solo expresaron la proteína mCherry, mientras que la edición dirigida en el UAG del transcripto reportero podría convertir el codón de terminación en UIG y, en consecuencia, permitir la expresión de EGFP corriente abajo. Dicho reportero nos permite medir la eficiencia de edición de A a I a través del monitoreo del nivel de EGFP. A continuación, diseñamos ARNcr impulsados por el promotor hU6 (ARN de CRISPR) que contienen andamios 5' sometidos al reconocimiento de Cas13a y longitudes variables de secuencias espaciadoras para el direccionamiento (ARNcr^{Cas13a}, siguiendo las secuencias de ARNcr de LbuCas13).

Secuencias de ARNcr de LbuCas13

Nombre	Secuencia	Fuente
Andamio de ARNcr de LbuCas13/Cas13a	Ggaccaccccaaaaugaaggggacuaaaac (SEQ ID NO: 46)	Figura 10
ARNcr ₇₀ Ctrl	Aaaccgagggaucauaggggacugaauccacauucucucccaauc ccugcaacuccuucuccccugc (SEQ ID NO: 47)	Figura 10
Espaciador de ARNcr ₁₅	gcagagccucCagc (SEQ ID NO: 48)	Figura 10
Espaciador de ARNcr ₂₂	cucacuggcagagccucCagc (SEQ ID NO: 49)	Figura 10
Espaciador de ARNcr ₂₈	cccuugcucacuggcagagccucCagc (SEQ ID NO: 50)	Figura 10
Espaciador de ARNcr ₃₅	cucucgcccucgucacuggcagagccucCagc (SEQ ID NO: 51)	Figura 10
Espaciador de ARNcr ₄₀	cucucgcccucgucacuggcagagccucCagcaucgc (SEQ ID NO: 52)	Figura 10
Espaciador de ARNcr ₄₇	ugaacagcucucgcccucgucacuggcagagccucCagcaucgc (SEQ ID NO: 53)	Figura 10

Espaciador de ARNcr ₇₀	ugaacagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccucccgcc (SEQ ID NO: 54)	Figura 10
-----------------------------------	---	-----------

- Todas las secuencias complementarias a los transcritos objetivo contienen CCA opuesto al codón UAG para introducir un emparejamiento incorrecto de citidina (C) con la adenosina (A) (figura 10B) porque la desaminación de adenosina se produce preferentemente en el sitio de desajuste AC^{13,14}. Para probar la longitud óptima del ARNcr, se coexpresaron ARNcr no dirigidos o dirigidos de diferentes longitudes con proteínas dCas13a-ADAR1_{DD} en células HEK293T que expresan de manera estable el reportero-1. Se observaron efectos de edición de ARN evidentes indicados por la aparición de expresión de EGFP con ARNcr que contenían secuencias coincidentes de al menos 40-nt de longitud, y cuanto más largos eran los ARNcr, mayor era el porcentaje positivo de EGFP (figura 10C). Sorprendentemente, la expresión de ARNcr^{Cas13a} larga sola pareció suficiente para activar la expresión de EGFP, y la coexpresión de dCas13a-ADAR1_{DD} disminuyó bastante la actividad de ARNcr (figuras 10C, 10d). La expresión de EGFP era claramente dependiente de la secuencia porque el ARN de control de 70-nt (exclusivo del andamio 5' para el cálculo de la longitud) no podía activar la expresión de EGFP (figuras 10C, 10D).
- Con el sorprendente hallazgo de que ciertos ARNcrCas13a de ingeniería larga permitían la edición de ARN independiente de dCas13a-ADAR1_{DD}, decidimos eliminar la secuencia de andamio de reclutamiento de Cas13a del ARNcr. Debido a que el ARNcr₇₀ tenía la actividad más alta para desencadenar la expresión de EGFP (figuras 10C, 10D), elegimos el mismo ARN guía de 70-nt de largo sin el andamio de reclutamiento de Cas13a para una prueba adicional (figura 11A y las siguientes secuencias de ARNars y control ARN utilizados en los ejemplos).

Nombre	Secuencia (5' ---> 3')	Fuente
ARNcr ₇₀ Ctrl	Aaaccgagggaucauaggggacugaauccaccauucucccauuccugcaacuccuucuccc cugc (SEQ ID NO: 55)	Figura 11
ARNar ₇₀	ugaacagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccuccc cgcc (SEQ ID NO: 56)	
ARNcr ₇₁ Ctrl	Ucucaguccaauguaugguccgagcacaagcucuaaucaaaguccgaggguagaccgguugcc auagga (SEQ ID NO: 57)	figura 14 y figura 16
ARNar ₇₁	acagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccuccuccg gcug (SEQ ID NO: 58)	
ARNar ₇₁ -CAA	acagcuccucgcccugcucacuggcagagcccucAagcaucgcgagcaggcgugccuccuccg gcug (SEQ ID NO: 59)	figura 16A
ARNar ₇₁ -CUA	acagcuccucgcccugcucacuggcagagcccucUagcaucgcgagcaggcgugccuccuccg gcug (SEQ ID NO: 60)	
ARNar ₇₁ -CGA	acagcuccucgcccugcucacuggcagagcccucGagcaucgcgagcaggcgugccuccuccg gcug	

	(SEQ ID NO: 61)	
ARNar ₇₁ -GCA	acagcuccucgcccugcucacuggcagagcccuGCAgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 62)	figura 16B, 16C
ARNar ₇₁ -UCA	acagcuccucgcccugcucacuggcagagcccuUCAgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 63)	
ARNar ₇₁ -ACA	acagcuccucgcccugcucacuggcagagcccuACAagcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 64)	
ARNar ₇₁ -CCU	acagcuccucgcccugcucacuggcagagcccuCCUgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 65)	
ARNar ₇₁ -GCU	acagcuccucgcccugcucacuggcagagcccuGCUgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 66)	
ARNar ₇₁ -UCU	acagcuccucgcccugcucacuggcagagcccuUCUgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 67)	
ARNar ₇₁ -ACU	acagcuccucgcccugcucacuggcagagcccuACUgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 68)	
ARNar ₇₁ -CCC	acagcuccucgcccugcucacuggcagagcccuCCCgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 69)	
ARNar ₇₁ -GCC	acagcuccucgcccugcucacuggcagagcccuGCCgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 70)	
ARNar ₇₁ -UCC	acagcuccucgcccugcucacuggcagagcccuUCCgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 71)	
ARNar ₇₁ -ACC	acagcuccucgcccugcucacuggcagagcccuACCgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 72)	
ARNar ₇₁ -CCG	acagcuccucgcccugcucacuggcagagcccuCCGgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 73)	
ARNar ₇₁ -GCG	acagcuccucgcccugcucacuggcagagcccuGCUgcaucgcgagcaggcgugccuccuccgc cgug (SEQ ID NO: 74)	
ARNar ₇₁ -UCG	acagcuccucgcccugcucacuggcagagcccuUCGgcaucgcgagcaggcgugccuccuccgc cgug(SEQ ID NO: 75)	
ARNar ₇₁ -ACG	acagcuccucgcccugcucacuggcagagcccuACGgcaucgcgagcaggcgugccuccuccgc cgug(SEQ ID NO: 76)	
ARNar ₃₁ -Reportero-1	acuggcagagccucCagcaucgcgagcagg (SEQ ID NO: 77)	figura 16D y figura 27
ARNar ₅₁ -Reportero-1	gcccugcucacuggcagagccucCagcaucgcgagcaggcgugccucc (SEQ ID NO: 78)	
ARNar ₉₁ -Reportero-1	acagcuccucgcccugcucacuggcagagccucCagcaucgcgagcaggcgugccuccuccgcc gug(SEQ ID NO: 79)	
ARNar ₁₁₁ -Reportero-1	accccggugaacagcuccucgcccugcucacuggcagagccucCagcaucgcgagcaggcgug ccuccuccgcccugccuccuccgc (SEQ ID NO: 80)	

ARNar ₁₃₁ -Reporter-1	gcucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagccucC agcaucgcgagcaggcgugccuccuccgcccugcuccuccgcccugcuccuccgcccugc (SEQ ID NO: 81)
ARNar ₁₅₁ -Reporter-1	ucgccguccagcucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucacuggca gagccucCagcaucgcgagcaggcgugccuccuccgcccugcuccuccgcccugcuccucc gccugcagcuuguaca (SEQ ID NO: 82)
ARNar ₁₇₁ -Reporter-1	gccguuuacgucgcccagcucgaccaggaugggcaccaccccgugaacagcuccucgccc gucacuggcagagccucCagcaucgcgagcaggcgugccuccuccgcccugcuccuccgccc cugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 83)
ARNar ₁₉₁ -Reporter-1	ugaacuuguggccguuuacgucgcccagcucgaccaggaugggcaccaccccgugaacagcu ccucgcccugcucacuggcagagccucCagcaucgcgagcaggcgugccuccuccgcccugcc uccuccgcccugcuccuccgcccugcagcuuguacagcucguccaugccgcccug (SEQ ID NO: 84)
ARNar ₂₁₁ -Reporter-1	ccggacacgcugaacuuguggccguuuacgucgcccagcucgaccaggaugggcaccaccccg gugaacagcuccucgcccugcucacuggcagagccucCagcaucgcgagcaggcgugccuccu ccgcccugcuccuccgcccugcuccuccgcccugcagcuuguacagcucguccaugccgccc uggaguggcggc (SEQ ID NO: 85)
ARNar ₃₁ -Reporter-2	gcgaccgggaucucCacagauucuccggc (SEQ ID NO: 86)
ARNar ₅₁ -Reporter-2	gcucacgguggcgaccgggaucucCacagauucuccggcguguauaccu (SEQ ID NO: 87)
ARNar ₇₁ -Reporter-2	ccucgcccugcucacgguggcgaccgggaucucCacagauucuccggcguguauaccuucg cugccu(SEQ ID NO: 88)
ARNar ₉₁ -Reporter-2	gugaacagcuccucgcccugcucacgguggcgaccgggaucucCacagauucuccggcgugu auaccuucugcugccuccuccgccc (SEQ ID NO: 89)
ARNar ₁₁₁ -Reporter-2	caccaccccgugaacagcuccucgcccugcucacgguggcgaccgggaucucCacagauuc ccggcguguauaccuucugcugccuccuccgcccugcuccucc (SEQ ID NO: 90)
ARNar ₁₃₁ -Reporter-2	ccaggaugggcaccaccccgugaacagcuccucgcccugcucacgguggcgaccgggaucucC acagauucuccggcguguauaccuucugcugccuccuccgcccugcuccuccgcccugccu (SEQ ID NO: 91)
ARNar ₁₅₁ -Reporter-2	uccagcucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucacgguggcgacc gggaucucCacagauucuccggcguguauaccuucugcugccuccuccgcccugcuccuccg ccgugccuccuccgccc (SEQ ID NO: 92)
ARNar ₁₇₁ -Reporter-2	cgcgacguauccagcucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucac gguggcgaccgggaucucCacagauucuccggcguguauaccuucugcugccuccuccgccc ugccuccuccgcccugcuccuccgcccugcagcuugu (SEQ ID NO: 93)
ARNar ₁₉₁ -Reporter-2	uguggccguuuacgucgcccagcucgaccaggaugggcaccaccccgugaacagcuccucg ccuugcucacgguggcgaccgggaucucCacagauucuccggcguguauaccuucugcugccu ccuccgcccugcuccuccgcccugcuccuccgcccugcagcuuguacagcucgucc (SEQ ID NO: 94)

ARNar ₂₁₁ -Reporter-2	acgcugaacuuguggccguuuacgucgccguccagcucgaccaggaugggcaccaccccgugaac agcuccucgcccugcucacgguggcgacgggggaucucCacagauucuuccggcguguauaccu ucugcugccuccuccgccgugccuccuccgccgugccuccuccgccgucagcuuguacagcucg uccaugccgccg (SEQ ID NO: 95)	
ARNar ₇₁ (C+70)-Reporter-1	Cagcaucgcgagcaggcgugccuccuccgccgugccuccuccgccgugccuccuccgccgca gcuu (SEQ ID NO: 96)	figura 16E
ARNar ₇₁ (5+C+65)-Reporter-1	cccucCagcaucgcgagcaggcgugccuccuccgccgugccuccuccgccgugccuccuccgcc ugc (SEQ ID NO: 97)	
ARNar ₇₁ (10+C+60)-Reporter-1	cagagcccucCagcaucgcgagcaggcgugccuccuccgccgugccuccuccgccgugccuccu ccgc (SEQ ID NO: 98)	
ARNar ₇₁ (15+C+55)-Reporter-1	acuggcagagcccuccCagcaucgcgagcaggcgugccuccuccgccgugccuccuccgccgug ccucc (SEQ ID NO: 99)	
ARNar ₇₁ (20+C+50)-Reporter-1	ugcucacuggcagagcccucCagcaucgcgagcaggcgugccuccuccgccgugccuccuccgcc gcug (SEQ ID NO: 100)	
ARNar ₇₁ (25+C+45)-Reporter-1	gcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccuccuccgccgugccuccu ccgc (SEQ ID NO: 101)	
ARNar ₇₁ (30+C+40)-Reporter-1	uccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccuccuccgccgugc cucc (SEQ ID NO: 102)	
ARNar ₇₁ (40+C+30)-Reporter-1	ggugaacagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgugccucc uccgc (SEQ ID NO: 103)	
ARNar ₇₁ (45+C+25)-Reporter-1	accccgugaacagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggcgug ccucc (SEQ ID NO: 104)	
ARNar ₇₁ (50+C+20)-Reporter-1	gcaccaccccgugaacagcuccucgcccugcucacuggcagagcccucCagcaucgcgagcaggc gcug (SEQ ID NO: 105)	
ARNar ₇₁ (55+C+15)-Reporter-1	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagcccucCagcaucgca gcagg (SEQ ID NO: 106)	
ARNar ₇₁ (60+C+10)-Reporter-1	accaggaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagcccucCagcau cgca (SEQ ID NO: 107)	
ARNar ₇₁ (65+C+5)-Reporter-1	gcucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagcccucC agcau (SEQ ID NO: 108)	
ARNar ₇₁ (70+C)-Reporter-1	guccagcugaccaggaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagc ccucC (SEQ ID NO: 109)	
ARNar ₇₁ (C+70)-Reporter-2	Cacagauucuuccggcguguauaccuucugcugccuccuccgccgugccuccuccgccgugccu ccucc (SEQ ID NO: 110)	

ARNar ₇₁ (5+C+65)-Reportero-2	aucucCacagauucuuccggcguguauaccuucugcugccuccuccgcccgcugccuccuccgcccgc ugccu (SEQ ID NO: 111)	figura 16F, 16G
ARNar ₇₁ (10+C+60)-Reportero-2	cggggaucucCacagauucuuccggcguguauaccuucugcugccuccuccgcccgcugccuccucc gccgc (SEQ ID NO: 112)	
ARNar ₇₁ (15+C+55)-Reportero-2	gcgaccgggggaucucCacagauucuuccggcguguauaccuucugcugccuccuccgcccgcugccu ccucc (SEQ ID NO: 113)	
ARNar ₇₁ (20+C+50)-Reportero-2	cgguaggcgaccgggggaucucCacagauucuuccggcguguauaccuucugcugccuccuccgcccgc ugccu (SEQ ID NO: 114)	
ARNar ₇₁ (25+C+45)-Reportero-2	gcucacggugcgaccgggggaucucCacagauucuuccggcguguauaccuucugcugccuccucc cgcccgc (SEQ ID NO: 115)	
ARNar ₇₁ (30+C+40)-Reportero-2	cccuugcucacggugcgaccgggggaucucCacagauucuuccggcguguauaccuucugcugcc uccucc (SEQ ID NO: 116)	
ARNar ₇₁ (40+C+30)-Reportero-2	cagcuccucgcccugcucacggugcgaccgggggaucucCacagauucuuccggcguguauaccu ucugc (SEQ ID NO: 117)	
ARNar ₇₁ (45+C+25)-Reportero-2	gugaacagcuccucgcccugcucacggugcgaccgggggaucucCacagauucuuccggcgugu auaccu (SEQ ID NO: 118)	
ARNar ₇₁ (50+C+20)-Reportero-2	ccccggugaacagcuccucgcccugcucacggugcgaccgggggaucucCacagauucuuccggc gugua (SEQ ID NO: 119)	
ARNar ₇₁ (55+C+15)-Reportero-2	caccaccccgugaacagcuccucgcccugcucacggugcgaccgggggaucucCacagauucu ccggc (SEQ ID NO: 120)	
ARNar ₇₁ (60+C+10)-Reportero-2	augggcaccaccccgugaacagcuccucgcccugcucacggugcgaccgggggaucucCacaga uucuu (SEQ ID NO: 121)	
ARNar ₇₁ (65+C+5)-Reportero-2	ccaggaugggcaccaccccgugaacagcuccucgcccugcucacggugcgaccgggggaucucC acaga (SEQ ID NO: 122)	
ARNar ₇₁ (70+C)-Reportero-2	cucgaccaggaugggcaccaccccgugaacagcuccucgcccugcucacggugcgaccggggga ucucC (SEQ ID NO: 123)	
ARNar ₁₁₁ -CCA-Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagccucCagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 124)	figura 16F, 16G
ARNar ₁₁₁ -GCA-Reportero-3 (UAC)	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagccucCagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 125)	
ARNar ₁₁₁ -UCA-Reportero-3 (UAA)	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagccucCagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 126)	
ARNar ₁₁₁ -ACA-Reportero-3 (UAU)	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagccucCagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 127)	

ARNar ₁₁₁ -CCG-Reportero-3 (CAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 128)	figura 17
ARNar ₁₁₁ -GCG-Reportero-3 (CAC)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 129)	
ARNar ₁₁₁ -UCG-Reportero-3 (CAA)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuuCggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 130)	
ARNar ₁₁₁ -ACG-Reportero-3 (CAU)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuaCggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 131)	
ARNar ₁₁₁ -CCU-Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 132)	
ARNar ₁₁₁ -GCU-Reportero-3 (AAC)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 133)	
ARNar ₁₁₁ -ACU-Reportero-3 (AAA)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuaCugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 134)	
ARNar ₁₁₁ -UCU-Reportero-3 (AAU)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuuCugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 135)	
ARNar ₁₁₁ -CCC-Reportero-3 (GAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 136)	
ARNar ₁₁₁ -GCC-Reportero-3 (GAC)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccucCcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 137)	
ARNar ₁₁₁ -UCC-Reportero-3 (GAA)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuuCcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 138)	figura 17B
ARNar ₁₁₁ -ACC-Reportero-3 (GAU)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagccuaCcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 139)	
ARNcr111 Ctrl	Uaccgcuacagccacgcugauuucagcuauaccugcccguauaaagggacguucacaccgcgau guucucugcugggaauugcgcauauucaggauuaaagaagugc (SEQ ID NO: 140)	
ARNcr151 Ctrl	Acuacaguugcuccgauuuuaggcuacgucauaggcacuaacuuuugcguggugaacgg acuuccucgaguuaccagaagagacuacaaaacuccuuuccauugcgaguuacggagucuggc ucaguugggcaggaggcacu (SEQ ID NO: 141)	
ARNar ₅₁ -PPIB	cggaagaggguggggcccgguggcCagggagccggcgccgacgcgcgg (SEQ ID NO: 142)	figura 17B
ARNar ₇₁ -PPIB	cagcugagggccgaagaggguggggcccgguggcCagggagccggcgccgacgcgcgggugg ggggga (SEQ ID NO: 143)	
ARNar ₁₁₁ -PPIB	ggaggcgaaagcagcccggacagcugaggccggaagaggguggggcccgguggcCagggagcc ggcgccgacgcgcggguggggggacugggguugcucgcccgcuc (SEQ ID NO: 144)	

ARNar ₁₅₁ -PPIB	gaggcgcagcauccacaggcggaggcgaaagcagcccgacagcugaggccggaagagggugggg ccgcgguggcCagggagccggcgccgacgcggguggggggacuggggguucgcggggcu ccggcgggcgggcgccg (SEQ ID NO: 145)
ARNar ₅₁ -KRAS	ucuugccuacgccaccagcuccaacCaccacaaguuuauuucagucuuuu (SEQ ID NO: 146)
ARNar ₇₁ -KRAS	gucaaggcacucuugccuacgccaccagcuccaacCaccacaaguuuauuucagucuuuuucagc aggcc
ARNar ₁₁₁ -KRAS	GauucugaaauagcuguaucgucaaggcacucuugccuacgccaccagcuccaacCaccacaagu uuauuucagucuuuuucagcaggccucucuccgcaccugggagc (SEQ ID NO: 147)
ARNar ₁₅₁ -KRAS	aucauuucguccacaaaugauucugaaauagcuguaucgucaaggcacucuugccuacgccac cagcuccaacCaccacaaguuuauuucagucuuuuucagcaggccucucuccgcaccugggagc cgugagccucuggcccg (SEQ ID NO: 148)
ARNar ₅₁ -SMAD 4	ucggcagguagaaguacuucgucCaggagcuggagggcccguguaagu (SEQ ID NO: 149)
ARNar ₇₁ -SMAD 4	gggucugcaaucggcagguagaaguacuucgucCaggagcuggagggcccguguaagugaa uuucaau (SEQ ID NO: 150)
ARNar ₁₁₁ -SMA D4	gaccucagucuaaagguuguggucgcaaucggcagguagaaguacuucgucCaggagcug gagggcccguguaagugaaauucauccagcaagguguuucuuuga (SEQ ID NO: 151)
ARNar ₁₅₁ -SMA D4	uaagggcccaacgguaaaagaccucagucuaaagguuguggucgcaaucggcagguaga aguacuucgucCaggagcuggagggcccguguaagugaaauucauccagcaagguguuucuu ugaugcucugucuuagguaaucc (SEQ ID NO: 152)
ARNar ₅₁ -FANC C (sitio TAC)	ugggggguucggcugccgacauagCaauugcucugccaccaucucagccc (SEQ ID NO: 153)
ARNar ₇₁ -FANC C (sitio TAC)	agcagggccgugggggguucggcugccgacauagCaauugcucugccaccaucucagcccaucc uccgaa (SEQ ID NO: 154)
ARNar ₁₁₁ -FANC C (sitio TAC)	aguagaaggccaagagccacagcaggccgugggggguucggcugccgacauagCaauugcuc gccaccaucucagcccaucccgaagugaaugaacaggaaccagc (SEQ ID NO: 155)
ARNar ₁₅₁ -FANC C (sitio TAC)	ccuccaucacggggccguagaaggccaagagccacagcaggccgugggggguucggcug ccgacauagCaauugcucugccaccaucucagcccaucccgaagugaaugaacaggaaccagc ucuaaagggaccuccgag (SEQ ID NO: 156)
ARNar ₁₅₁ -PPIB (sitio AAG)	gcaaacaccacatgcttgccatctagccaggctgtcttgactgtctgatgaagaactgggagccgttgg gtcCttgcctcgcttgccatgctcaccagccaggccgtagtgcctcagttgaagttctcatcggggaag cgctca (SEQ ID NO: 157)
ARNar ₁₅₁ -PPIB (sitio CAG)	gggagtgggtccgctccaccagatgccagcaccggggccagtgagctcagagccctgtggcgactac agggccCgcacagacggtcactcaagaaagatgtcctgtgcttactccttgccgatggcaaagggtt ctccactcga (SEQ ID NO: 158)
ARNar ₁₅₁ -FANC C (sitio AAG)	tgcattttgtaaatagatactagcagattgtcccaagatgtgtacagctattctcacagcccagcagggc acCtactccacaatgcgtggccacaggtcatcacctgtcctgtggccctggcgagcctgatccctcacgc gggac (SEQ ID NO: 159)
ARNar ₁₅₁ -FANC C (sitio CAG)	gctcattctcacagcccagcagggcacttactccacaatgcgtggccacaggtcatcacctgtcctgtg cccCggcgagcctgatccctacgcccgggacccacagggcctgctgctttagacttgagttcgagct ctttaag

figura
17C

	(SEQ ID NO: 160)	
ARNar ₁₅₁ -IDUA (sitio CAG)	tcggccgggcccctggggcggtgggcgtggccaggacgcccaccgtgtggtgtgttccaggacggtcc cgcccCgcgacacttcggcccagagctgctcctatccagcagcgccagcagccccatggccgtgagcac cggttgcgca (SEQ ID NO: 161)	
ARNar ₁₁₁ -TARD BP	ugaccagucuuagaucuuucugaccugcaccuaaagaacuucuccaaagguaCaaaauacuc uuucagguccuguuucggguuuuuuccaugggagaccaacacauu (SEQ ID NO: 162)	figura 17D
ARNar ₁₁₁ -CGA- Reportero-3 (UAG)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccucGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 163)	
ARNar ₁₁₁ -GGA- Reportero-3 (UAC)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccugGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 164)	
ARNar ₁₁₁ -UGA- Reportero-3 (UAA)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 165)	
ARNar ₁₁₁ -AGA- Reportero-3 (UAU)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 166)	
ARNar ₁₁₁ -CGG- Reportero-3 (CAG)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccucGggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 167)	
ARNar ₁₁₁ -GGG- Reportero-3 (CAC)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccugGggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 168)	
ARNar ₁₁₁ -UGG- Reportero-3 (CAA)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 169)	figura 17G
ARNar ₁₁₁ -AGG- Reportero-3 (CAU)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGggcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 170)	
ARNar ₁₁₁ -CGU- Reportero-3 (AAG)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccucGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 171)	
ARNar ₁₁₁ -GGU- Reportero-3 (AAC)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccugGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 172)	
ARNar ₁₁₁ -AGU- Reportero-3 (AAA)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 173)	
ARNar ₁₁₁ -UGU- Reportero-3 (AAU)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccuGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 174)	
ARNar ₁₁₁ -CGC- Reportero-3 (GAG)	gaugggcaccaccccgguaacagcuccugcccuugcucacuggcagagcccucGcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 175)	

ARNar ₁₁₁ -GGC- Reportero-3 (GAC)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 176)	
ARNar ₁₁₁ -UGC- Reportero-3 (GAA)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 177)	
ARNar ₁₁₁ -AGC- Reportero-3 (GAU)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGcgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 178)	
ARNar ₁₁₁ -CGA- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 179)	figura 17H
ARNar ₁₁₁ -GGA- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 180)	
ARNar ₁₁₁ -UGA- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuUGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 181)	
ARNar ₁₁₁ -AGA- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuAGagcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 182)	
ARNar ₁₁₁ -CGU- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuUGucaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 183)	
ARNar ₁₁₁ -CGG- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGGgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 184)	
ARNar ₁₁₁ -CGC- Reportero-3 (UAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGCgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 185)	
ARNar ₁₁₁ -CGU- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 186)	
ARNar ₁₁₁ -GGU- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 187)	
ARNar ₁₁₁ -UGU- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuUGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 188)	
ARNar ₁₁₁ -AGU- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuAGugcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 189)	
ARNar ₁₁₁ -CGA- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGAgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 190)	
ARNar ₁₁₁ -CGC- Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccugcccuugcucacuggcagagcccuGCgcaucgcga gcaggcgugccuccuccgcccugcagcuuguacagcucguccau (SEQ ID NO: 191)	

ARNar ₁₁₁ -CGG-Reportero-3 (AAG)	gaugggcaccaccccgugaacagcuccucgcccugcucacuggcagagcccucGGgcaucgcga gcaggcgugccucccgcccugcagcuuguacagcucguccau (SEQ ID NO: 192)	
ARNar ₁₁₁ -KRAS-AG6	gauucugaauuagcuguaucgucaaggcacucgugccgacgccaccagcuccaacCaccacaagug gagagucagucuuuucagcaggccucucucccgaccugggagc (SEQ ID NO: 193)	figura 17I
ARNar ₁₁₁ -KRAS-AG9	gauucugaauuagcuggauucgucaaggcacucgggcccagccaccagcuccaacCaccacaagug gagagucagucuuuucagcaggccucucucccgaccggggagc (SEQ ID NO: 194)	
ARNar ₁₁₁ -TP53	gggagcagccucuggcauucugggagcucaucuggaccugggucucagugaacCauuguuca auaucguccgggacagcaucaaaucauccauugcuugggacggcaa (SEQ ID NO: 195)	figura 23
ARNar ₁₁₁ -TP53-AG1	gggagcagccucuggcauucugggagcucaucuggaccugggucucagugaacCauuguuca agaucguccgggacagcaucaaaucauccauugcuugggacggcaa (SEQ ID NO: 196)	
ARNar ₁₁₁ -TP53-AG4	gggagcagccucuggcagucgggagcucaucuggaccugggucucagugaacCauuguuca agaucguccgggacagcaucaaaucauccagucugugggacggcaa (SEQ ID NO: 197)	
ARNar ₁₁₁ -COL3 A1	cauauuacagaauaccuugauagcauccaauuugcauccuugguuagggucaaccCaguauucuc cacucugaguucaggauaggcagaauuucaggucucugcaguuucu (SEQ ID NO: 198)	figura 26
ARNar ₁₁₁ -BMP R2	gugaagauaagccaguccucuaagaaacagaauagcaagacggcaagagcuuaccCagucacuug uguggagacuuaaaauacuugcauaaagaucgauugggagauagucuc (SEQ ID NO: 199)	
ARNar ₁₁₁ -AHI1	gugaacgucaaacugcggaccaauauggcagaauucucucaucucaacuucCauauccgua ucauggaaucauagcauccuguaacuacuagcucuuacagcugg (SEQ ID NO: 200)	
ARNar ₁₁₁ -FANCC (Sitio 2)	gccauagaucucgugaguuaucucagcagugugagccaucagggugaugacauccCaggcgau guguggccuccaggagccagagcaggaaguaggagaaggugccu (SEQ ID NO: 201)	
ARNar ₁₁₁ -MYB PC3	caagacggugaaccacuccauggucuuucugcggcuucugcacuguguaccccCagagcuccg uguugccgacaucuccugggguggcuccacuccagagccacauaaag (SEQ ID NO: 202)	
ARNar ₁₁₁ -IL2R G	aggauucucuuuugaaguauugcuccccaguggauuggguggcuccauucacucCauugcuga gcacuuccacagaguggguuaaagcggcuccgaacacgaaacgugua (SEQ ID NO: 203)	
ARNar ₁₁₁ -IDUA -V1	gacgccaccgugugguugcuguccaggacggucccgccgacacucggccCagagcugcuc cucauccagcagcgccagcagccccauaggccgugagcaccggcuu (SEQ ID NO: 204)	figura 29
ARNar ₁₁₁ -IDUA -V2	gacgccaccgugugguugcuguccaggacggucccgccgacacucggccCagagcugcuc cucaucugcggggcgggggggggcgccgucggcguggggucguug (SEQ ID NO: 205)	

Resultó que este ARN guía lineal inducía una fuerte expresión de EGFP en cerca del 40% de las células totales que albergaban el reportero-1 (figura 11B, superior). Debido a que las proteínas ADAR endógenas podrían editar sustratos¹² de ARN de cadena doble (ARNds), razonamos que los ARN de guía largos podrían aparearse con los transcritos objetivo para formar sustratos de ARNds que a su vez reclutan proteínas ADAR endógenas para la edición dirigida. Por lo tanto, designamos dicho ARN

guía como ARNar (ARN de reclutamiento de ADAR).

Para verificar si las proteínas ADAR endógenas son realmente responsables de la observación anterior, nos propusimos examinar la edición de ARN mediada por ARNar en células deficientes en ADAR. Dado que el ARNm de ADAR2 era apenas detectable en las células HEK293T (figura 12A), generamos células HEK293T ADAR1^{-/-}, lo que hace que esta línea celular sea deficiente tanto en ADAR1 como en ADAR2 (figuras 11C, 11D). De hecho, el agotamiento de ADAR1 anuló las señales de EGFP inducidas por ARNar₇₀ (figura 11B, inferior). Además, la expresión exógena de ADAR1^{p110}, ADAR1^{p150} o ADAR2 en células HEK293T ADAR1^{-/-} (figura 11C, 11D) rescató con éxito la pérdida de inducción de EGFP por ARNar₇₀ (figura 11E, figura 12B), demostrando que la expresión del reportero de EGFP inducido por ARNar dependía únicamente de ADAR1 nativa, cuya actividad podría ser reconstituida por sus isoformas (p110 y p150) o ADAR2. El análisis de secuenciación de Sanger en la región de direccionamiento de ARNar₇₀ mostró un pico de solapamiento A/G en el sitio de adenosina predicho dentro de UAG, lo que indica una conversión significativa de A a I (G) (figura 11F). La secuenciación de próxima generación (NGS) confirmó además que la tasa de conversión de A a I era aproximadamente el 13% de los transcritos de reportero totales (figura 11G). El análisis de PCR cuantitativo mostró que ARNar₇₀ no redujo la expresión de transcritos dirigidos (figura 13), descartando el posible efecto de ARNi del ARNar. En conjunto, nuestros datos demostraron que el ARNar es capaz de generar un nivel significativo de edición en los transcritos específicos a través del desajuste A-C diseñado.

20

EJEMPLO 7

LEAPER permite la edición de ARN en múltiples líneas celulares

Debido a que la expresión de proteínas ADAR endógenas es un requisito previo para la edición de ARN mediada por LEAPER, probamos el rendimiento de LEAPER en un panel de líneas celulares originadas a partir de distintos tejidos, incluidos HT29, A549, HepG2, RD, SF268, SW13 y HeLa. Primero examinamos la expresión endógena de los tres tipos de proteínas ADAR utilizando análisis Western blotting. ADAR1 se expresó altamente en todas las líneas celulares ensayadas, y su identidad en el análisis Western blot se confirmó mediante el control negativo, línea HEK293T ADAR1^{-/-} (figuras 14A, 14B). Se detectó ADAR3 sólo en células HepG2 y HeLa (figuras 14A, 14B). ADAR2 no fue detectable en ninguna célula, un resultado que no se debió al fracaso del Western blot porque la proteína ADAR2 pudo detectarse a partir de células HEK293T que sobreexpresan ADAR2 (figuras 14A, 14B). Estos hallazgos concuerdan con informes anteriores de que ADAR1 se expresa de manera ubicua, mientras que las expresiones de ADAR2 y ADAR3 están restringidas a ciertos tejidos¹¹.

35

Luego nos propusimos probar las eficiencias de edición de un ARNar 71-nt rediseñado

(ARNar₇₁) dirigido al reportero-1 (figura 15A y secuencias de ARNar y ARN de control utilizados en este estudio enumerados anteriormente) en estas líneas celulares.

LEAPER trabajó en todas las células probadas para este ARNar₇₁, aunque con eficiencias variables (figura 14C). Estos resultados estaban de acuerdo con el informe anterior de que los niveles de proteína ADAR1/2 se correlacionan con el rendimiento de edición de ARN⁴², con la excepción de las células HepG2 y HeLa. Las correlaciones subóptimas de las eficiencias de edición con los niveles de ADAR1 probablemente se debieron a las abundantes expresiones de ADAR3 en estas dos líneas (figuras 14A, 14B) porque se ha informado que ADAR3 desempeña un papel inhibitor en la edición de ARN. Es importante destacar que LEAPER también trabajó en tres líneas celulares diferentes de origen de ratón (NIH3T3, fibroblasto embrionario de ratón (MEF) y B16) (figura 14D), allanando el camino para probar su potencial terapéutico a través de modelos animales y de enfermedades. Colectivamente, llegamos a la conclusión de que LEAPER es una herramienta versátil para un amplio espectro de tipos de células y para diferentes organismos.

15

EJEMPLO 8

Caracterización y optimización de LEAPER

Para caracterizar y optimizar mejor LEAPER, investigamos las opciones de nucleótidos opuestos a la adenosina dentro del triplete UAG del transcripto dirigido. En las células HEK293T, el ARNar₇₁ dirigido a reportero-1 mostró eficiencias de edición variables con un triplete cambiado (5'-CNA, N denota uno de A/U/C/G) opuesto al UAG objetivo (secuencias de ARNar y ARN de control utilizados en este estudio enumerados anteriormente). El desajuste A-C resultó en la mayor eficiencia de edición, y el desajuste A-G produjo las ediciones menos evidentes (figura 16A). Luego investigamos la preferencia de los nucleótidos que flanquean el desajuste A-C en el ARNar. Probamos las 16 combinaciones de sitios vecinos 5' y 3' que rodean la citidina (5'-N¹CN²) (secuencias de ARNar y ARN de control usados en este estudio enumerados anteriormente), y encontramos que la adenosina vecina 3' era necesaria para la edición eficiente, mientras que la adenosina es el nucleótido menos favorable en el sitio 5' (figuras 16B, 16C). Por lo tanto, llegamos a la conclusión de que el motivo CCA en el ARNar confiere la mayor eficiencia de edición dirigida al sitio UAG. Vale la pena señalar que la guanosina vecina 3'(5'-N¹CG) en el ARNar mostró un efecto inhibitor espectacular (figuras 16B, 16C).

La longitud del ARN pareció relevante para la eficiencia del ARNar al dirigir la edición en los transcriptos objetivo (figura 10C), en consonancia con un informe anterior⁴². Para comprender completamente este efecto, probamos ARNar con longitudes variables dirigidas a dos transcriptos reporteros diferentes: reportero-1 y reportero-2 (figuras 15A, 15B). Para cualquiera de los dos indicadores, se diseñaron y probaron ARNar de 10 tamaños diferentes, que van de 31-nt a 211-nt, con

35

triplete CCA (para direccionamiento de UAG) justo en el medio (secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente). Según las actividades de EGFP del reportero, la longitud del ARNar se correlacionó positivamente con la eficiencia de edición, para ambos reporteros, alcanzando un máximo de 111 a 191-nt (figura 16D). Aunque un ARNar₅₁ parecía funcionar, 71-nt era la longitud mínima para que el ARNar funcionara para ambos reporteros (figura 16D).

A continuación, investigamos el efecto de la posición de desajuste A-C dentro de un ARNar sobre la eficiencia de edición. Fijamos las longitudes de todos los ARNar para las pruebas en 71-nt y deslizamos el triplete ACC dirigido a UAG de 5' a 3' dentro de los ARNar (secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente). Resultó que colocar el desajuste A-C en la región media resultó en un alto rendimiento de edición, y los ARNar con los sitios de desajuste cerca del extremo 3' superaron a los cercanos al extremo 5' en ambos reporteros (figura 16E). Por conveniencia, colocamos el desajuste A-C en el centro de los ARNar para todos nuestros estudios posteriores.

También probamos la flexibilidad de direccionamiento de LEAPER e intentamos determinar si UAG en el objetivo es el único motivo sujeto a edición de ARN. Para las 16 combinaciones de tripletes (5'-N¹AN²) en reportero-3 (figura 15C), usamos los ARNar correspondientes con las longitudes fijas (111-nt) y aseguramos la coincidencia de la secuenciación perfecta para el ARNar y el reportero excepto para el sitio de edición (desajuste A-C) (figura 16F y secuencias de ARNar y ARN de control usadas en este estudio enumeradas anteriormente). Los resultados de NGS mostraron que se podían editar todos los motivos N¹AN². El UAN² y GAN² son los motivos más y menos preferibles, respectivamente (figuras 16F, 16G). En conjunto, la preferencia del vecino más cercano de la adenosina objetivo es 5' U>C≈A>G y 3' G>C>A≈U (figura 16G).

EJEMPLO 9

Edición de transcritos endógenos con LEAPER

A continuación, examinamos si LEAPER pudiera permitir la edición eficaz de transcritos endógenos. Usando ARNar de diferentes longitudes, nos enfocamos en los motivos UAG en los transcritos de los genes *PPIB*, *KRAS* y *SMAD4*, y un motivo UAC en el transcripto del gen *FANCC* (figura 17A, secuencias de ARNar y ARN de control usados en este estudio enumerado anteriormente). De manera alentadora, los sitios de adenosina dirigidos en los cuatro transcritos fueron editados por sus ARNar correspondientes con los cuatro tamaños, aunque con eficiencias variables según los resultados de NGS (figura 17B). De acuerdo con nuestra observación anterior, los ARNar más largos tendían a producir tasas de edición más altas. Es de destacar que el ARNar_{PPIB} de 151-nt editó ~50% de los transcritos totales del gen *PPIB* (figura 17B). Ningún ARNar mostró efectos de ARNi sobre sus

transcriptos dirigidos (figura 18A) o el nivel de proteína final (por ejemplo, KRAS, figura 18B). Además, LEAPER puede lograr una tasa de edición deseable en sitios que no son UAN (figura 17C y secuencias de ARNar y ARN de control usados en este estudio enumerados anteriormente), lo que muestra la flexibilidad de LEAPER en la edición de transcriptos endógenos. Para explorar más a fondo el poder de

5 LEAPER, probamos si podía apuntar simultáneamente a varios sitios. Observamos la edición multiplex de los transcriptos de *TARDBP* y *FANCC* mediante la coexpresión de dos ARNar (secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente), con una eficiencia incluso mayor que con ARNar individuales (figura 17D), lo que indica que LEAPER es ideal para editar varios objetivos en paralelo.

10 Es de destacar que ADAR1/2 tiende a desaminar de forma promiscua múltiples adenosinas en un dúplex de ARN⁴⁴ y el desajuste A-C no es el único motivo para guiar el cambio A-a-I (figura 16A). Por lo tanto, es razonable suponer que todas las adenosinas en los transcriptos objetivo dentro de las coberturas de ARNar están sujetas a niveles variables de edición, fuentes importantes de modificaciones no deseadas. Cuanto más largo sea el ARNar, mayor será la posibilidad de que se

15 produzcan tales desvíos. Por lo tanto, examinamos todos los sitios de adenosina dentro del ARNar que cubren las regiones en estos transcriptos dirigidos. Para los transcriptos de PPIB, se observó muy poca edición fuera del objetivo a lo largo de la ventana de secuenciación para tamaños variables de ARNar (figuras 17E, 17F). Sin embargo, en los casos de dirigirse a los genes KRAS, SMAD4 y FANCC, se detectaron múltiples ediciones fuera del objetivo (figuras 19A a 19F). Para KRAS en particular, 11 de las

20 30 adenosinas experimentaron conversiones sustanciales de A a I en la ventana de secuenciación del ARNar₁₁₁ (figuras 19A, 19B).

A continuación, intentamos desarrollar estrategias para minimizar esos efectos no deseados fuera del objetivo. Debido a que un desajuste A-G suprimió la edición para la direccionamiento de UAG (figura 16A), postulamos que emparejar una guanosina con una adenosina

25 no dirigida podría reducir la edición no deseada. Luego probamos el efecto del desajuste A-G sobre la adenosina en todas las combinaciones posibles de tripletes (5'-N¹AN²) como en el reportero-3 (figura 15C y secuencias de ARNar y ARN de control usadas en este estudio enumeradas anteriormente). El desajuste A-G de hecho disminuyó la edición en adenosina en todos los objetivos probados, excepto para el direccionamiento de UAG o AAG (~2%) (figura 17G), en comparación con el desajuste A-C

30 (figura 16F). Para reducir aún más las tasas de edición en sitios no deseados, continuamos probando el efecto de dos desajustes consecutivos. Resultó que el desajuste adicional en el nucleótido del extremo 3' del triplete opuesto a UAG o AAG, abolió su correspondiente edición de adenosina (figura 17H y secuencias de ARNar y ARN de control usadas en este estudio enumeradas anteriormente). A la luz de estos hallazgos, intentamos aplicar esta regla para reducir los objetivos fuera de objetivos en los

35 transcriptos de KRAS (figura 19A). Primero diseñamos un ARNar (ARNar₁₁₁-AG6) que creaba

desajustes de A-G en todos los motivos "propensos a la edición" cubiertos por ARNar₁₁₁ (figura 17I, figura 19A y secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente), incluida la AAU (el 61), UAU (el 63), UAA (el 65), AAA (el 66), UAG (el 94) y AAG (el 99). Este ARNar₁₁₁-AG6 eliminó la mayor parte de la edición fuera del objetivo, mientras que mantuvo una tasa de edición en el objetivo de ~ 5%. De acuerdo con los hallazgos en la figura 17G, el único desajuste A-G no pudo minimizar completamente la edición en el motivo AAG (el 99) (figura 17I y figura 19A). Luego agregamos más desajustes en ARNar₁₁₁-AG6, incluido un desajuste dual (5'-CGG opuesto al motivo objetivo 5'-AAG), más tres desajustes de A-G adicionales para mitigar la edición en las adenosinas 27, 98 y 115 (ARNar₁₁₁-AG9) (secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente). En consecuencia, logramos una especificidad mucho mejor para la edición, sin una pérdida adicional de la tasa de edición en el sitio objetivo (A76) (figura 17I). En resumen, LEAPER diseñado que incorpora reglas adicionales permite una edición de ARN más eficiente y precisa en los transcriptos endógenos.

15

EJEMPLO 10

Especificidad de edición de ARN de LEAPER

Además de los posibles efectos fuera del objetivo dentro de la región de ARNds cubierta de ARNar, también estábamos preocupados por los posibles efectos fuera del objetivo en otros transcriptos a través del acoplamiento parcial de bases de ARNar. Luego, realizamos un análisis de secuenciación de ARN en todo el transcriptoma para evaluar los efectos globales fuera del objetivo de LEAPER. Las células se transfectaron con un ARNcr₁₅₁ Ctrl o un ARNar específico de PPIB (ARNar₁₅₁-PPIB) que expresaban plásmidos antes de someterlas a análisis de ARN-sec. Identificamos seis potenciales desvíos en el grupo ARN₁₅₁ Ctrl (figura 20A) y cinco en el grupo ARNar₁₅₁-PPIB (figura 20B), y la tasa de PPIB dentro del objetivo basada en el análisis NGS fue ~37% (figura 20B). Un análisis más detallado reveló que todos los sitios, a excepción de los dos sitios de los transcriptos de *EIF2AK2*, estaban ubicados en las regiones SINE (Alu) o LINE (figuras 20A, 20B), ambos son propensos a la edición mediada por ADAR⁴⁵, lo que sugiere que estos desvíos no pueden derivarse del acoplamiento entre los transcriptos objetivo y el ARNar o el ARN de control. Es de destacar que en ambos grupos aparecieron dos transcriptos no dirigidos, *WDR73* y *SMYD4*, lo que sugiere que es poco probable que la edición de ARN dependa de la secuencia. De hecho, el análisis de energía libre mínima sugirió que todos estss posibles transcriptos fuera del objetivo no pudieron formar un dúplex estable con ARN₁₅₁ Ctrl o ARNar₁₅₁-PPIB (figura 20C). Para probar más a fondo si ARNar genera fuera de objetivos dependientes de la secuencia, seleccionamos sitios potenciales fuera de objetivo comparando la similitud de secuencia utilizando NCBI BLAST para ARNar₁₅₁-PPIB y ARNar₁₁₁-FANCC.

Los transcritos de *TRAPPC12* para ARNar₁₅₁-PPIB y tres sitios en los transcritos *ST3GAL1*, *OSTM1-AS1* y *EHD2* para ARNar₁₁₁-FANCC fueron los mejores candidatos (figura 20D y figura 21A). El análisis NGS reveló que no se pudo detectar edición en ninguno de estos sitios fuera del objetivo predichos (figura 20D y figura 21B). Estos resultados indican que LEAPER permite la edición eficiente en el sitio objetivo, mientras mantiene la especificidad de todo el transcriptoma sin ediciones fuera del objetivo dependientes de la secuencia detectables.

EJEMPLO 11

Evaluación de la seguridad de LEAPER en células de mamíferos

10

Debido a que los ARNar dependen de proteínas ADAR endógenas para editar los transcritos objetivo, nos preguntamos si la adición de ARNar exógenos afecta los eventos de edición de ARN nativo al ocupar demasiadas proteínas ADAR1 o ADAR2. Por lo tanto, analizamos los sitios de edición de ARN A-to-I compartidos por el grupo simulado y el grupo ARNar₁₅₁-PPIB de los resultados de secuenciación de ARN de todo el transcriptoma, y también se analizó la comparación entre el grupo simulado y el grupo ARN₁₅₁ Ctrl. Ni el grupo ARN₁₅₁ Ctrl ni el grupo ARNar₁₅₁-PPIB mostraron una diferencia significativa en comparación con el grupo simulado (figuras 22A, 22B), lo que indica que LEAPER tuvo poco impacto en la función normal de ADAR1 endógena para catalizar los eventos de edición nativos A-a-I.

20

Mientras tanto, realizamos un análisis diferencial de expresión génica utilizando datos de ARN-seq para verificar si el ARNar afecta la expresión génica global. Encontramos que ni ARN₁₅₁ Ctrl ni ARNar₁₅₁-PPIB afectaron la expresión génica global en comparación con el grupo simulado (figuras 22C, 22D). De acuerdo con nuestra observación anterior (figura 18A), los ARNar no mostraron ningún efecto de ARNi sobre la expresión de PPIB (figuras 22C, 22D).

25

Teniendo en cuenta que el ARNar forma dúplex de ARN con el transcripto objetivo y que el dúplex de ARN podría provocar una respuesta inmune innata, investigamos si la introducción de ARNar tiene tal efecto. Para probar esto, seleccionamos ARNar dirigidos a cuatro transcritos de genes que habían demostrado ser eficaces. No observamos ninguna inducción de ARNm de interferón- β (IFN- β) (figura 22E) o interleucina-6 (IL-6) (figura 22F), que son dos características de la activación inmune innata. Como un control positivo, un análogo sintético de ARN de cadena doble - poli(I:C) indujo una fuerte expresión de IFN- β e IL-6 (figuras 22E, 22F). LEAPER no parece inducir inmunogenicidad en las células objetivo, una característica importante para terapias seguras.

30

EJEMPLO 12

Recuperación de la actividad reguladora transcripcional de p53 por LEAPER

Ahora que hemos establecido un método novedoso para la edición de ARN sin la
 5 necesidad de introducir proteínas extrañas, intentamos demostrar su utilidad terapéutica. Primero nos
 enfocamos en el gen supresor de tumores *TP53*, que se sabe que juega un papel vital en el
 mantenimiento de la homeostasis celular, pero sufre mutaciones frecuentes en >50% de los cánceres
 humanos⁴⁶. La mutación c.158G>A en *TP53* es una mutación sin sentido clínicamente relevante
 (Trp53Ter), que da como resultado una proteína truncada no funcional⁴⁶. Diseñamos un ARNar₁₁₁ y
 10 dos ARNar alternativos (ARNar₁₁₁-AG1 y ARNar₁₁₁-AG4) (secuencias de ARNar y ARN de control usadas
 en este estudio enumeradas anteriormente), todos dirigidos a transcritos de *TP53*^{W53X} (figura 23A), y
 los dos últimos están diseñados para minimizar potenciales fuera de objetivos. Generamos la línea
 celular HEK293T *TP53*^{-/-} para eliminar los efectos de la proteína p53 nativa. Las tres formas de ARNar
 dirigidos a *TP53*^{W53X} convirtieron ~25-35% de los transcritos de *TP53*^{W53X} en el sitio de adenosina
 15 mutado (figura 23B), con reducciones variables de ediciones no deseadas para ARNar₁₁₁-AG1 y
 ARNar₁₁₁-AG4 (figura 24). Western blot mostró que ARNar₁₁₁, ARNar₁₁₁-AG1 y ARNar₁₁₁-AG4 podrían
 rescatar la producción de proteína p53 de longitud completa basada en los transcritos *TP53*^{W53X} en
 células HEK293T *TP53*^{-/-}, mientras que ARN₁₁₁ Ctrl no pudo (figura 23C).

Para verificar si las proteínas p53 reparadas son completamente funcionales,
 20 probamos la actividad reguladora de la transcripción de p53 con un sistema de notificación en *cis* de
p53-luciferasa^{47,48}. Las tres versiones de ARNar pudieron restaurar la actividad de p53, y la versión
 optimizada ARNar₁₁₁-AG1 tuvo el mejor rendimiento (figura 23D). En conclusión, demostramos que
 LEAPER es capaz de reparar el codón de terminación prematuro de *TP53* relevante para el cáncer y
 restaurar su función.

25

EJEMPLO 13

Correcciones de mutaciones patogénicas por LEAPER

A continuación, investigamos si LEAPER podría usarse para corregir más mutaciones
 30 patógenas. Apuntando a mutaciones clínicamente relevantes de seis genes patógenos, *COL3A1* del
 síndrome de Ehlers-Danlos, *BMPR2* de hipertensión pulmonar primaria, *AHI1* del síndrome de Joubert,
FANCC de anemia de Fanconi, *MYBPC3* de miocardiopatía hipertrófica familiar primaria e *IL2RG* de
 inmunodeficiencia combinada severa ligada al cromosoma X, se diseñó ARNar 111-nt para cada uno de
 estos genes portadores de mutaciones patogénicas correspondientes G>A (figura 25 y secuencias de
 35 ARNar y ARN de control usados en este estudio enumerados anteriormente, y los siguientes ADNc

relevantes para la enfermedad usados en este estudio).

ADNc relacionados con la enfermedad utilizados en este estudio

Candidato	Enfermedad	Adenosina mutante
NM_000090.3 (<i>COL3A1</i>)	Síndrome de Ehlers-Danlos, tipo 4	c.3833G>A (p.Trp1278Ter)
NM_001204.6 (<i>BMPR2</i>)	Hipertensión pulmonar primaria	c.893G>A (p.Trp298Ter)
NM_017651.4 (<i>AHI1</i>)	Síndrome de Joubert 3	c.2174G>A (p.Trp725Ter)
NM_000136.2 (<i>FANCC</i>)	Anemia de Fanconi, grupo de complementación C	c.1517G>A (p.Trp506Ter)
NM_000256.3 (<i>MYBPC3</i>)	Miocardiopatía hipertrófica familiar primaria	c.3293G>A (p.Trp1098Ter)
NM_000206.2 (<i>IL2RG</i>)	Inmunodeficiencia combinada severa ligada al cromosoma X	c.710G>A (p.Trp237Ter)

Al coexpresar pares de ARNc/ADNc en células HEK293T, identificamos cantidades significativas de transcritos objetivo con correcciones A>G en todas las pruebas (figura 24). Debido a que las mutaciones G>A representan casi la mitad de las mutaciones puntuales que causan enfermedades conocidas en humanos^{10,49}, la conversión A>G por LEAPER puede ofrecer inmensas oportunidades para la terapéutica.

10

EJEMPLO 14

Edición de ARN en múltiples células primarias humanas por LEAPER

Para explorar más a fondo la utilidad clínica de LEAPER, nos propusimos probar el método en múltiples células primarias humanas. Primero, probamos LEAPER en fibroblastos pulmonares primarios humanos y células epiteliales bronquiales primarias humanas con ARNar 151-nt (secuencias de ARNar y ARN de control usadas en este estudio enumeradas anteriormente) para editar el reportero-1 (figura 15A). El 35-45% de las células positivas para EGFP se pudieron obtener mediante LEAPER en ambas células primarias humanas (figura 27A). Luego probamos LEAPER en la edición del gen PPIB endógeno en estas dos células primarias y células T primarias humanas, y encontramos que ARNar₁₅₁-PPIB podría alcanzar >40%, >80% y >30% de las tasas de edición en fibroblastos pulmonares primarios humanos, bronquios primarios células epiteliales (figura 27B) y células T primarias (figura 27C), respectivamente. La alta eficiencia de edición de LEAPER en células primarias

humanas es particularmente alentadora por su potencial aplicación en terapéutica.

EJEMPLO 15

Edición eficiente mediante expresión lentiviral y síntesis química de ARNar

5

Luego investigamos si LEAPER pudiera administrarse mediante métodos más relevantes desde el punto de vista clínico. Primero probamos el efecto del ARNar a través de la expresión basada en lentivirus. ARNar₁₅₁ dirigido a reportero-1 indujo una fuerte expresión de EGFP en más del 40% de las células totales que albergan reportero-1 en células HEK293T 2 días después de la infección (dpi). A 8 dpi, la proporción de EGFP se mantuvo en un nivel comparable de ~38% (figura 28A y secuencias de ARNar y ARN de control utilizadas en este estudio enumeradas anteriormente), lo que sugiere que LEAPER podría adaptarse a las terapias que requieren una administración continua. Para la edición de genes nativos, entregamos ARNar₁₅₁ dirigido a *PP1B* a través de transducción lentiviral en células HEK293T y observamos más del 6% de edición objetivo a 6 dpi (figura 28B).

15

A continuación, probamos los oligonucleótidos de ARNar sintetizados y el método de administración de electroporación para LEAPER. El ARNar de 111-nt dirigido a los transcriptos de *PP1B*, así como el ARN Ctrl, se sintetizaron químicamente con modificaciones de enlace de 2'-O-metilo y fosforotioato en los primeros tres y últimos tres residuos de ARNar (figura 28C). Después de introducirse en las células T mediante electroporación, los oligonucleótidos ARNar₁₁₁-*PP1B* lograron ~20% de edición en los transcriptos de *PP1B* (figura 28D), lo que indica que LEAPER es prometedor para el desarrollo de fármacos oligonucleotídicos.

20

EJEMPLO 16

Restauración de la actividad de α -L-iduronidasa en fibroblasto primario derivado del paciente con síndrome de Hurler por LEAPER

25

Finalmente, examinamos el potencial de LEAPER en el tratamiento de una enfermedad monogénica - el síndrome de Hurler, el subtipo más grave de mucopolisacaridosis tipo I (MPS I) debido a la deficiencia de α -L-iduronidasa (IDUA), un metabolismo lisosómico enzima responsable de la degradación de mucopolisacáridos⁵⁰. Elegimos un fibroblasto primario GM06214 que se aisló originalmente de un paciente con síndrome de Hurler. Las células GM06214 contienen una mutación homocigótica TGG>TAG en el exón 9 del gen *IDUA*, lo que da como resultado una mutación Trp402Ter en la proteína. Diseñamos dos versiones de ARNar mediante oligonucleótidos de ARN sintetizados con modificaciones químicas, ARNar₁₁₁-IDUA-V1 y ARNar₁₁₁-IDUA-V2, dirigidas al ARNm maduro y al pre-ARNm de *IDUA*, respectivamente (figura 29A y secuencias de ARNar y ARN de control utilizado en

35

este estudio enumerado anteriormente). Después de la introducción de ARNar₁₁₁-IDUA-V1 o ARNar₁₁₁-IDUA-V2 en células GM06214 mediante electroporación, medimos las tasas de edición de ARN dirigido mediante análisis NGS y la actividad catalítica de α -L-iduronidasa con el sustrato 4-MU- α -L-iduronidasa en diferentes momentos. Tanto ARNar₁₁₁-IDUA-V1 como ARNar₁₁₁-IDUA-V2 restauraron significativamente la actividad catalítica de IDUA en células GM06214 deficientes en *IDUA* progresivamente con el tiempo después de la electroporación, y ARNar₁₁₁-IDUA-V2 funcionó mucho mejor que ARNar₁₁₁-IDUA-V1, mientras que no se pudo detectar actividad de la α -L-iduronidasa en tres grupos de control (figura 29B).

Para evaluar más a fondo hasta qué punto la actividad de IDUA restaurada en GM06214 por LEAPER alivia el síndrome de Hurler, examinamos la actividad de IDUA en células GM01323, otros fibroblastos primarios de pacientes con síndrome de Scheie, un subtipo de MPS I mucho más leve que el síndrome de Hurler debido a la actividad de IDUA remanente resultante del genotipo heterocigoto en el gen *IDUA*. Encontramos que la actividad catalítica de IDUA en células GM06214 que albergan ARNar₁₁₁-IDUA-V2 era mayor que la de las células GM01323 48 horas después de la electroporación (figura 29B). De acuerdo con estos resultados, el análisis NGS indicó que ARNar₁₁₁-IDUA-V2 convirtió casi el 30% de la edición de A en I, una tasa mucho más alta que ARNar₁₁₁-IDUA-V1 (figura 29C). Un análisis adicional reveló que se detectaron ediciones no deseadas mínimas dentro de las regiones cubiertas de ARNar de los transcritos de *IDUA* (figura 29D). Es importante destacar que LEAPER no desencadenó respuestas inmunes en células primarias, ya que demostramos que, a diferencia del ARN dúplex poli(I:C) que sirve como un control positivo, ni ARNar₁₁₁-IDUA-V1 ni ARNar₁₁₁-IDUA-V2 indujeron expresiones de un panel de genes implicados en el interferón de tipo I y las respuestas proinflamatorias (figura 29E). Estos resultados mostraron el potencial terapéutico de LEAPER para dirigir ciertas enfermedades monogénicas.

25

EJEMPLO 17

Restauración de la actividad de la α -L-iduronidasa en ratones con síndrome de Hurler por LEAPER

Como LEAPER brinda la posibilidad de realizar experimentos en un ratón, examinamos el potencial de LEAPER para tratar una enfermedad monogénica - el síndrome de Hurler *in vivo*. El virus adenoasociado (AAV) se selecciona como el sistema de administración. Las dos secuencias (Seq ID NO. 341 y Seq ID NO. 342) se insertaron en el vector plásmido AAV-U6-CMV-GFP y el plásmido se empaquetó en el virus AAV8 por PackGene Biotech (ver la figura 31). El título de virus fue el siguiente:

nombre	título	volumen	Cantidad total	Secuencia ARNar 5'-3
AAV8-IDUA -ARNad151 -KD2	9.61E+1 2GC/ml	590ul	5.67E+12GC	Seq ID NO. 341: gggtgatgggtgctggccaggacacccactgtatg attgctgtccaacacagccccagcctttgagacctct gccagagttgttctcatccaacagggccatgagc cccatgactgtgagtactggcttcgagcaactgca catggg
AAV8-Aleat orio151-KD 2	9.91E+12GC/ml	600ul	5.95E+12GC	Seq ID NO. 342: actacagttgctccgatatttaggctacgtcaata ggcactaacttattggcgctggtgaacggacttcc tctcgagtaccagaagatgactacaaaactcctt ccattgcgagtatcggagtctggtcagtttgcc aggaggcact

Se seleccionaron seis ratones macho MPSI (ratones idua W392X, B6.129S-Idua^{tm1.1Kmke/J}) (4-10 semanas de edad) y se dividieron en dos grupos, y a cada uno de los tres

5 ratones del primer grupo se le inyectó 1E12 GC de virus AAV8 empacado con plásmido AAV8-IDUA-aARNd151-KD2 a través de la vena de la cola, a cada uno de los otros tres ratones del segundo grupo se les inyectó 1E12 GC de virus AAV8 empacado con plásmido AAV8-Aleatorio151-KD2 a través de la vena de la cola, así como un control. Veintiocho días después de la inyección, los ratones se sacrificaron mediante dislocación cervical y se extrajeron hígados. . Medimos la actividad enzimática

10 de α -L-iduronidasa en células hepáticas de ratón con sustrato de 4-MU- α -L-iduronidasa para cada ratón en dos grupos. El hígado se divide en trozos pequeños, una parte después de la trituración se utiliza para medir la actividad enzimática (inmediatamente). La actividad enzimática relativa media del primer grupo es mayor que la del segundo grupo (Ver la figura 32A). La actividad enzimática relativa de cada ratón del primer grupo es mayor que la del segundo grupo (Ver la figura 32B, GM01323 como control y

15 tiene 0.3% de actividad de la enzima α -L-iduronidasa en comparación con las células de tipo salvaje). Puede verse por el resultado de la actividad enzimática que se restaura la actividad de la α -L-iduronidasa del primer grupo de ratones. Para confirmar que la mutación TGG>TAG se restaura y el LEAPER proporciona una edición de ARN eficaz, medimos las tasas de edición de ARN dirigido en las células hepáticas de ratón mediante un análisis de secuenciación de próxima generación para cada

20 ratón en dos grupos mediante el uso de otra parte del hígado triturado. La otra parte se agrega con trizol para NGS. La eficiencia de edición de ARN promedio del primer grupo es mayor que la del segundo grupo (Ver la figura 33A). La eficiencia de edición de ARN de cada ratón del primer grupo es casi mayor que la del segundo grupo (Ver la figura 33B). Se puede ver en el resultado de la eficiencia

de la edición de ARN que el método LEAPER proporciona una edición de ARN efectiva de hasta el 7%.

ADNc de ADAR1(p110)

5'-ATGGCCGAGATCAAGGAGAAAATCTGCGACTATCTCTTCAATGTGTCTGACTCCTCTGCCCTGAATTTGGCTA
 AAAATATTGGCCTTACCAAGGCCCGAGATATAAATGCTGTGCTAATTGACATGGAAAGGCAGGGGGATGTCTAT
 5 AGACAAGGGACAACCCCTCCCATATGGCATTGACAGACAAGAAGCGAGAGAGGATGCAAATCAAGAGAAATA
 CGAACAGTGTTCTGAAACCGCTCCAGCTGCAATCCCTGAGACCAAAAGAAACGCAGAGTTCCTCACCTGTAAT
 ATCCACATCAAATGCCTCAAATAACATGGTAACACAGAAAAAGTGGAGAATGGGCAGGAACCTGTCATAAA
 GTTAGAAAACAGGCAAGAGGCCAGACCAGAACCAGCAAGACTGAAACCACCTGTTTATTACAATGGCCCCCTCA
 AAAGCAGGGTATGTTGACTTTGAAAATGGCCAGTGGGCCACAGATGACATCCCAGATGACTTGAATAGTATCCG
 10 CGCAGCACCAGGTGAGTTTCGAGCCATCATGGAGATGCCCTCCTTCTACAGTCATGGCTTGCCACGGTGTTTAC
 CCTACAAGAACTGACAGAGTGCCAGCTGAAGAACCCCATCAGCGGGCTGTTAGAATATGCCAGTTCGCTAGT
 CAAACCTGTGAGTTCAACATGATAGAGCAGAGTGGACCACCCCATGAACCTCGATTTAAATTCCAGGTTGTCAT
 CAATGGCCGAGAGTTTCCCCAGCTGAAGCTGGAAGCAAGAAAGTGGCCAAGCAGGATGCAGCTATGAAAGCC
 ATGACAATTCTGCTAGAGGAAGCCAAAGCCAAGGACAGTGGAAAATCAGAAGAATCATCCCACTATTCCACAGA
 15 GAAAGAATCAGAGAAGACTGCAGAGTCCCAGACCCCCACCCCTTCAGCCACATCCTTCTTTTCTGGGAAGAGCC
 CCGTCACCACACTGCTTGAGTGTATGCACAAATTGGGGAACCTCTGCGAATTCCGTCTCCTGTCCAAAGAAGGC
 CCTGCCCATGAACCCAAGTTCCAATACTGTGTTGCAGTGGGAGCCCAAACCTTCCCCAGTGTGAGTGCTCCCAG
 CAAGAAAGTGGCAAAGCAGATGGCCGAGAGGAAGCCATGAAGGCCCTGCATGGGGAGGCGACCAACTCCAT
 GGCTTCTGATAACCAGCCTGAAGGTATGATCTCAGAGTCACTTGATAACTTGAATCCATGATGCCCAACAAGG
 20 TCAGGAAGATTGGCGAGCTCGTGAGATACCTGAACACCAACCCCTGTGGGTGGCCTTTTGGAGTACGCCCCGCTC
 CCATGGCTTTGCTGCTGAATTCAAGTTGGTTCGACCAAGTCCGGACCTCCTCACGAGCCCAAGTTCGTTTACCAAG
 CAAAAGTTGGGGGTGCTGGTTCAGCCGTCTGCGCACACAGCAAGAAGCAAGGCAAGCAGGAAGCAGCAG
 ATGCGGCTCTCCGTGTCTTGATTGGGGAGAACGAGAAGGCAGAACGCATGGGTTTCACAGAGGTAACCCCACT
 GACAGGGGCCAGTCTCAGAAGAACTATGCTCCTCCTCTCAAGGTCCCCAGAAGCACAGCCAAAGACACTCCCT
 25 CTCACTGGCAGCACCTTCCATGACCAGATAGCCATGCTGAGCCACCGGTGCTTCAACACTCTGACTAACAGCTT
 CCAGCCCTCCTTGCTCGGCCGCAAGATTCTGGCCGCCATCATTATGAAAAAAGACTCTGAGGACATGGGTGTGCG
 TCGTCAGCTTGGGAACAGGGAATCGCTGTGTAAAAGGAGATTCTCTCAGCCTAAAAGGAGAACTGTCAATGA
 CTGCCATGCAGAAATAATCTCCCGGAGAGGCTTCATCAGGTTTCTCTACAGTGAGTTAATGAAATACAACTCCCA
 GACTGCGAAGGATAGTATATTTGAACCTGCTAAGGGAGGAGAAAAAGCTCCAAATAAAAAAGACTGTGTCAATTCC
 30 ATCTGTATATCAGCACTGCTCCGTGTGGAGATGGCGCCCTCTTTGACAAGTCCTGCAGCGACCGTGCTATGGAA
 AGCACAGAATCCCGCCACTACCCTGTCTTCGAGAATCCCAAACAAGGAAAGCTCCGCACCAAGGTGGAGAACG
 GAGAAGGCACAATCCCTGTGGAATCCAGTGACATTGTGCCTACGTGGGATGGCATTGCGCTCGGGGAGAGACT
 CCGTACCATGTCCTGTAGTGACAAAATCCTACGCTGGAACGTGCTGGGCCTGCAAGGGGCACTGTTGACCCAC
 TTCCTGCAGCCATTTATCTCAAATCTGTCACATTGGGTTACCTTTTTCAGCCAAGGGCATCTGACCCGTGCTATT
 35 TGCTGTGCTGTGACAAGAGATGGGAGTGCATTTGAGGATGGACTACGACATCCCTTTATTGTCAACCACCCCAA

GGTTGGCAGAGTCAGCATATATGATTCCAAAAGGCAATCCGGGAAGACTAAGGAGACAAGCGTCAACTGGTGT
 CTGGCTGATGGCTATGACCTGGAGATCCTGGACGGTACCAGAGGCACTGTGGATGGGCCACGGAATGAATTGT
 CCCGGGTCTCCAAAAGAACATTTTTCTTCTATTTAAGAAGCTCTGCTCCTTCCGTTACCGCAGGGATCTACTGA
 GACTCTCCTATGGTGAGGCCAAGAAAGCTGCCCCGTGACTACGAGACGGCCAAGAACTACTTCAAAAAGGCCT
 5 GAAGGATATGGGCTATGGGAACTGGATTAGCAAACCCAGGAGGAAAAGAACTTTTATCTCTGCCAGTA
[GATTACAAGGATGACGACGATAAG\(etiqueta Flag\)](#) TAG-3' (SEQ ID NO:332)

ADNc de ADAR1(p150)

5'ATGAATCCGCGGCAGGGGTATTCCCTCAGCGGATACTACACCCATCCATTTCAGGCTATGAGCACAGACAGC
 TCAGATACCAGCAGCCTGGGCCAGGATCTTCCCCAGTAGTTTCTGCTTAAGCAAATAGAATTTCTCAAGGGG
 10 CAGCTCCCAGAAGCACCGGTGATTGGAAAAGCAGACACCGTCACTGCCACCTTCCCTCCCAGGACTCCGGCCAA
 GGTTCAGTACTACTTGCCTCCAGTACCAGAGGCAGGCAAGTGGACATCAGGGGTGTCCCAGGGGCGTGCA
 TCTCGGAAGTCAGGGGCTCCAGAGAGGGTTCCAGCATCCTTACCACGTGGCAGGAGTCTGCCACAGAGAGG
 TGTTGATTGCCTTTCCTCACATTTCCAGGAACTGAGTATCTACCAAGATCAGGAACAAAGGATCTTAAAGTTCCT
 GGAAGAGCTTGGGGAAGGGAAGGCCACCACAGCACATGATCTGTCTGGGAAACTTGGGACTCCGAAGAAAGA
 15 AATCAATCGAGTTTTATACTCCCTGGCAAAGAAGGGCAAGCTACAGAAAGAGGCAGGAACACCCCCTTTGTGGA
 AAATCGCGGTCTCCACTCAGGCTTGAACACAGCACAGCGGAGTGTAAGACCAGACGGTCATAGCCAAGGAGC
 CCCAACTCAGACCCGAGTTTGAACCGGAAGACAGAACTCCACATCTGTCTCAGAAGATCTTCTTGAGCCTT
 TTATTGCAGTCTCAGCTCAGGCTTGAACACAGCACAGCGGAGTGTAAGACCAGACAGTCATAGCCAAGGATC
 CCCAACTCAGACCCAGGTTTGAACCTGAAGACAGCAACTCCACATCTGCCTTGAAGATCCTCTTGAGTTTT
 20 TAGACATGGCCGAGATCAAGGAGAAAAATCTGCGACTATCTCTTCAATGTGTCTGACTCCTCTGCCCTGAATTTG
 GCTAAAAATATTGGCCTTACCAAGGCCCGAGATATAAATGCTGTGCTAATTGACATGGAAAGGCAGGGGGATGT
 CTATAGACAAGGGACAACCCCTCCCATATGGCATTGACAGACAAGAAGCGAGAGAGGATGCAATCAAGAGAA
 ATACGAACAGTGTTCTGAAACCGCTCCAGCTGCAATCCCTGAGACCAAAAAGAAACGCAGAGTTCCTCACCTGT
 AATATACCCACATCAAATGCCTCAAATAACATGGTAACCACAGAAAAAGTGGAGAATGGGCAGGAACCTGTCAT
 25 AAAGTTAGAAAACAGGCAAGAGGCCAGACCAGAACCAGCAAGACTGAAACCACCTGTTCAATACAATGGCCCC
 TCAAAAGCAGGGTATGTTGACTTTGAAAATGGCCAGTGGGCCACAGATGACATCCAGATGACTTGAATAGTAT
 CCGCGCAGCACCAGGTGAGTTTCGAGCCATCATGGAGATGCCCTCCTTCTACAGTCATGGCTTGCCACGGTGTT
 CACCCTACAAGAACTGACAGAGTGCCAGCTGAAGAACCCCATCAGCGGGCTGTTAGAATATGCCAGTTCGCT
 AGTCAAACCTGTGAGTTCAACATGATAGAGCAGAGTGGACCACCCCATGAACCTCGATTTAAATTCCAGGTTGT
 30 CATCAATGGCCGAGAGTTTCCCCAGCTGAAGCTGGAAGCAAGAAAGTGGCCAAGCAGGATGCAGCTATGAAA
 GCCATGACAATTCTGCTAGAGGAAGCCAAAGCCAAGGACAGTGGAAAATCAGAAGAATCATCCCACTATTCCAC
 AGAGAAAGAATCAGAGAAGACTGCAGAGTCCCAGACCCCCACCCCTTACGCCACATCCTTCTTTCTGGGAAGA
 GCCCCGTCAACCACTGCTTGAGTGTATGCACAAATTGGGGAACCTCTGCGAATTCCGTCTCCTGTCCAAAGAA
 GGCCCTGCCCATGAACCCAAGTTCCAATACTGTGTTGCAAGTGGGAGCCCAACTTTCCCAAGTGTGAGTGCTCC
 35 CAGCAAGAAAGTGGCAAAGCAGATGGCCGCAGAGGAAGCCATGAAGGCCCTGCATGGGGAGGCGACCAACTC

CATGGCTTCTGATAACCAGCCTGAAGGTATGATCTCAGAGTCACTTGATAACTTGGAATCCATGATGCCCAACAA
 GGTGAGGAAGATTGGCGAGCTCGTGAGATACCTGAACACCAACCCTGTGGGTGGCCTTTTGGAGTACGCCCCG
 TCCCATGGCTTTGCTGCTGAATTCAAGTTGGTCGACCAGTCCGGACCTCCTCAGAGCCCAAGTTCGTTTACCA
 AGCAAAAGTTGGGGGTCGCTGGTTCAGCCGTCTGCGCACACAGCAAGAAGCAAGGCAAGCAGGAAGCAGC
 5 AGATGCGGCTCTCCGTGTCTTGATTGGGGAGAACGAGAAGGCAGAACGCATGGGTTTCACAGAGGTAACCCCA
 GTGACAGGGGCCAGTCTCAGAAGAACTATGCTCCTCCTCTCAAGGTCCCCAGAAGCACAGCCAAAGACTCC
 CTCTACTGGCAGCACCTTCCATGACCAGATAGCCATGCTGAGCCACCGGTGCTTCAACACTCTGACTAACAGC
 TTCCAGCCCTCCTTGCTCGGCCGCAAGATTCTGGCCGCCATCATTATGAAAAAGACTCTGAGGACATGGGTGT
 CGTCGTCAGCTTGGGAACAGGGAATCGCTGTGTAAAAGGAGATTCTCTCAGCCTAAAAGGAGAACTGTCAAT
 10 GACTGCCATGCAGAAATAATCTCCCGGAGAGGCTTCATCAGGTTTCTCTACAGTGAGTTAATGAAATACAACCTCC
 CAGACTGCGAAGGATAGTATATTTGAACCTGCTAAGGGAGGAGAAAAAGTCCAAATAAAAAAGACTGTGTCATT
 CCATCTGTATATCAGCACTGCTCCGTGTGGAGATGGCGCCCTCTTTGACAAGTCCTGCAGCGACCGTGCTATGG
 AAAGCACAGAATCCCGCCACTACCCTGTCTTCGAGAATCCCAAACAAGGAAAGCTCCGCACCAAGGTGGAGAA
 CGGAGAAGGCACAATCCCTGTGGAATCCAGTGACATTGTGCCTACGTGGGATGGCATTTCGGCTCGGGGAGAGA
 15 CTCCGTACCATGTCCTGTAGTGACAAAATCCTACGCTGGAACGTGCTGGGCCTGCAAGGGGCACTGTTGACCC
 ACTTCCTGCAGCCATTTATCTCAAATCTGTCACATTGGGTACCTTTTCAGCCAAGGGCATCTGACCCGTGCTA
 TTTGCTGTGCTGTGACAAGAGATGGGAGTGCAATTTGAGGATGGACTACGACATCCCTTTATTGTCAACCACCCC
 AAGGTTGGCAGAGTCAGCATATATGATTCCAAAAGGCAATCCGGGAAGACTAAGGAGACAAGCGTCAACTGGT
 GTCTGGCTGATGGCTATGACCTGGAGATCCTGGACGGTACCAGAGGCACTGTGGATGGGCCACGGAATGAATT
 20 GTCCCGGGTCTCCAAAAAGAACATTTTTCTTCTATTTAAGAAGCTCTGCTCCTTCCGTTACCGCAGGGATCTACT
 GAGACTCTCCTATGGTGAGGCCAAGAAAGCTGCCCCTGACTACGAGACGGCCAAGAACTACTTCAAAAAAGGC
 CTGAAGGATATGGGCTATGGGAACTGGATTAGCAAACCCAGGAGGAAAAGAACTTTATCTCTGCCAGTA
[GATTACAAGGATGACGACGATAAG\(etiqueta Flag\)](#) TAG-3' (SEQ ID NO:333)

25 **ADNc de ADAR2**

5'-ATGGATATAGAAGATGAAGAAAACATGAGTTCCAGCAGCACTGATGTGAAGGAAAACCGCAATCTGGACAAC
 GTGTCCCCCAAGGATGGCAGCACACCTGGGCCTGGCGAGGGCTCTCAGCTCTCCAATGGGGGTGGTGGTGGC
 CCCGGCAGAAAGCGGCCCTGGAGGAGGGCAGCAATGGCCACTCCAAGTACCGCCTGAAGAAAAGGAGGAAA
 ACACCAGGGCCCGTCCTCCCCAAGAACGCCCTGATGCAGCTGAATGAGATCAAGCCTGGTTTGCAGTACACAC
 30 TCCTGTCCCAGACTGGGCCCGTGACGCGCCTTTGTTTGTGATGTCTGTGGAGGTGAATGGCCAGGTTTTTGA
 GGGCTCTGGTCCCACAAAGAAAAAGGCAAACTCCATGCTGCTGAGAAGGCCTTGAGGTCTTTCGTTCAGTTT
 CCTAATGCCTCTGAGGCCACCTGGCCATGGGGAGGACCCTGTCTGTCAACACGGACTTCACATCTGACCAGG
 CCGACTTCCCTGACACGCTCTTCAATGGTTTTGAACTCCTGACAAGGCGGAGCCTCCCTTTTACGTGGGCTCC
 AATGGGGATGACTCCTTCAGTTCCAGCGGGGACCTCAGTTGTCTGCTTCCCCGGTGCCTGCCAGCCTAGCCC
 35 AGCCTCCTCTCCCTGCCTTACCACCATTCCCACCCCCGAGTGGGAAGAATCCCGTGATGATCTTGAACGAAGTG

CGCCCAGGACTCAAGTATGACTTCCTCTCCGAGAGCGGGGAGAGCCATGCCAAGAGCTTCGTCATGTCTGTGG
 TCGTGGATGGTCAGTTCTTTGAAGGCTCGGGGAGAAACAAGAAGCTTGCCAAGGCCCGGGCTGCGCAGTCTG
 CCCTGGCCGCCATTTTTAACTTGCACTTGGATCAGACGCCATCTCGCCAGCCTATTCCAGTGAGGGTCTTCAG
 CTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCACGCCTGGTCCTGGGTAAGTTTGGTGACCTGACCGACAA
 5 CTTCTCCTCCCCTCACGCTCGCAGAAAAGTGCTGGCTGGAGTCGTCATGACAACAGGCACAGATGTTAAAGAT
 GCCAAGGTGATAAGTGTTTCTACAGGAACAAAATGTATTAATGGTGAATACATGAGTGATCGTGGCCTTGCAATTA
 AATGACTGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACACAACCTTGAGCTTTACTTAA
 ATAACAAAGATGATCAAAAAAGATCCATCTTTCAGAAATCAGAGCGAGGGGGGTTTAGGCTGAAGGAGAATGTC
 CAGTTTCATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAA
 10 GAACCAGCAGATAGACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACCAAATAGAGTCTGGTGAGGGGA
 CGATTCCAGTGCGCTCCAATGCGAGCATCCAAACGTGGGACGGGGTGCTGCAAGGGGAGCGGCTGCTACCA
 TGTCTCAGTGACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGGGATCCCTGCTCAGCATTTTCGTGGA
 GCCCATTACTTCTCGAGCATCATCTGGGCAGCCTTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAGC
 GGATCTCCAACATAGAGGACCTGCCACCTCTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCA
 15 GAAGCACGGCAGCCAGGGAAGGCCCCCACTTCAGTGTCAACTGGACGGTAGGCGACTCCGCTATTGAGGTCA
 TCAACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCCTGTGTAAGCACGCGTTGTACTGTGCGCT
 GGATGCGTGTGCACGGCAAGGTTCCCTCCCACTTACTACGCTCCAAGATTACCAAACCAACGTGTACCATGAG
 TCCAAGCTGGCGGCAAGGAGTACCAGGCCGCCAAGGCGCGTCTGTTTACAGCCTTCATCAAGGCGGGGCTG
 GGGGCCTGGGTGGAGAAGCCCACCGAGCAGGACCAGTTCTCACTCACGCCC
 20 [GATTACAAGGATGACGACGATAAG\(etiqueta Flag\)](#) TAG-3' (SEQ ID NO:334)

Secuencia de codificación (CDS) de los genes relevantes para la enfermedad

COL3A1

5'-atgatgagctttgtgcaaaaggggagctggctacttctcgctctgcttcatccactattattttggcacaacaggaagctgtgaaggaggat
 25 gttcccatcttggtcagtcctatgcggatagagatgtctggaagccagaaccatgccaaatatgtgtctgtgactcaggatccgttctctgcatg
 acataatatgtgacgatcaagaattagactgccccaacccagaaattccatttgagaatgttggtgagtttggccacagcctccaactgctccta
 ctgcccctctaattggtcaaggacctcaaggccccagggagatccaggccctcctggtattcctgggagaaatggtgaccttggtattccagg
 acaaccaggtcccctggttctctggtccccctggaatctgtgaatcatgcccactggtcctcagaactattctccccagtatgattcatatgatg
 tcaagtctggagtagcagtaggaggactcgcaggctatcctggaccagctggccccccagggcctcccgtccccctggtacatctggtcatcc
 30 tgggtcccctggtatccaggataccaaggacccctggtgaacctgggcaagctggtcctcaggccctccaggacctctggtgctataggtc
 catctggtcctgctggaaaagatggagaatcaggtagaccggacgacctggagagcgaggattgctggacctccaggatcaaaaggtcca
 gctgggatacctggattcccctggtatgaaaggacacagaggctcgtgagacgaaatggagaaaagggtagaacaggtgctctcgattaaa
 gggtagaaaatggtctccaggcgaaaatggagctcctggacctatgggtccaagaggggctcctggtgagcgaggacggccaggactcctg
 gggctgcaggtgctcggggaatgacggtgctcgaggcagtgatggtcaaccaggccctcctggtcctcctggaactgccgattcccctggatc
 35 ccctggtgctaagggtagagttgacctgcaggggtcctcctggttcaaatggtgccctggacaaagaggagaacctggacctcagggacacg

ctggtgctcaaggtcctcctggccctcctgggattaatggtagtcctggtggttaaaggcgaaatgggtcccgtggcattcctggagctcctgga
 ctgatgggagccccgggtcctccaggaccagccggtgctaattggtgctcctggactgaggtggtgagcctggaagaatggtgc
 caaaggagagccccgaccacgtggtgaacgcggtgaggctggtattccagggttccaggagctaaaggcgaaagatggcaaggatggatc
 acctggagaacctggtgcaaatgggttccaggagctgcaggagaaaaggggtcccctgggtccgaggacctgctggacaaatggcatcc
 5 caggagaaaagggctcctgctggagagcgtggtgctccaggccctgcagggcccagaggagctgctggagaacctggcagagatggcgctcc
 ctggaggctcaggaatgaggggcatgcccggaaagtccaggaggaccaggaagtgatgggaaaccagggcctcccgaagtcaaggagaa
 agtggctgaccaggctcctcctgggccatctggtccccgaggtcagcctggtgcatgggttccccgggtcctaaaggaaatgatggtgctcctgg
 taagaatggagaacgaggtggcctggaggacctggcctcagggctcctggaagaatggtgaaactggacctcagggacccccaggg
 cctactggcctggtggtgacaaaggagacacaggacccccctggtccacaaggattacaaggcttgctggtacaggtggtcctccaggaga
 10 aaatggaaaacctggggaaccaggtccaaaggggatgcccgtgcacctggagctccaggaggcaaggggatgctggtgcccctggtgaa
 cgtggacctcctggattggcaggggccccaggacttagaggtggagctggtccccctggtcccgaaggaggaaaggggtgctgctgctcct
 gggccacctggtgctgctggtactcctggtctgcaaggaatgctggagaaaaggagggtcttggaagtctggtccaaaggggtgacaaggg
 tgaaccaggcgggtccaggtgctgatggtgtccagggaagatggccaaggggtcctactggtcctattggtcctcctggcccagctggcca
 gcctggagataagggtaaggtggtgccccggacttccaggtatagctggacctcgtggtagccctggtgagagaggtgaaactggccctc
 15 caggacctgctggttccctggtgctcctggacagaatggtgaacctggtggttaaaggagaaagaggggtccgggtgagaaaggtgaagg
 aggcctcctggagttgaggacccccctggaggttctggacctgctggtcctcctggtcccaagggtgcaaggtgaacctggcagctcctggt
 ggacctggtgctgctggttccctggtgctggttctcctggtcctcctggtagtaatggttaaccaggacccccagggtccagcggttccca
 ggcaaggatgggccccaggctcgcggtaacactggtgctcctggcagccctggagtgctggacaaaagggtgatgctggccaaccagg
 agagaagggatcgctggtgccaggggcccaccaggagctccaggccactgggattgctgggatcactggagcacggggtcttgaggga
 20 ccaccaggcatgccaggtcctaggggaagccctggccctcaggggtgcaagggtaagtgaggaaaccaggagctaaccggtcagtgagg
 aacgtggtcccctggaccccagggtcttctggtgctggtgacagctggtgaacctggaagagatggaacctggtatcagatggtctcc
 aggccgagatggatcctcctggtggcaaggggtgatcggtgaaaatggctcctggtgcccctggcgctcctggtcatccaggcccacctggt
 cctgtcgggtccagctggaagagtggtgacagaggagaaagtggccctgctggcctgctggtgctcccgtcctggttcccagggtgctc
 ctggtcctcaaggcccacgtggtgacaaaggtgaaacaggtgaacctggagctgctggcatcaaaggacatcaggattcctggtgaatcca
 25 ggtgccccaggttctccaggccctgctggtcagcaggggtgcaatcggcagtcaggacctgcaggccccagaggacctgttgaccagtggtg
 acctcctggcaagatggaaccagtgacatccaggtccattggaccaccagggcctcaggtgaacagaggtgaaagaggatctgagggc
 tccccaggccaccagggaaccaggccctcctggacctcctggtgcccctggtcctgctggtggtggtggtgagccgctgccattgctgggat
 tggaggtgaaaaagctggcggttttggccgtattatggagatgaaccaatggatttcaaatcaacaccgatgagattatgacttcaactcaagt
 ctgtaatggacaaatagaaagcctcattagtcctgatggttctgtaaaaaccccgtagaaactgcagagacctgaaattctgccatcctgaa
 30 ctcaagagtgagaaatactgggtgacctaaccaaggatgcaaatggatgctatcaaggattctgtaatatgaaactggggaaacatgc
 ataagtccaatccttgaatgttccacggaacactggtggacagattctagtctgagaagaaacacgttggttgagagtcctatggatg
 gtggtttcagtttagctacggaatcctgaacttctgaagatgtccttgatgtgcagctggcattccttcgacttctcagccgagcttccag
 aacatcacatatcactgcaaaaatagcattgcatacatggatcaggccagtggaatgtaagaaggccctgaagctgatggggtcaaatga
 aggtgaattcaaggctgaaggaaatagcaaatcacctacacagttctggaggatggtgcacgaaacacactggggaatggagcaaaaca
 35 gtcttgaatatcgaacacgcaaggctgtgagactacctattgtagatattgcaccctatgacattggtggtcctgatcaagaatttgggtggac

gttggccctgtttgcttttataa-3' (SEQ ID NO:335)

BMPR2

5'-atgacttctcgtcgcagcggccctggcgggtgccctggctaccatggaccatcctgctggtcagcgtgcggctgcttcgcagaatcaaga
acggctatgtgcgtttaagatccgtatcagcaagaccttgggataggtgagagtagaatctctcatgaaaatgggacaatattatgctcga
5 ggtagcacctgctatggcctttgggagaaatcaaaagggacataaatctgtaaaacaaggatgttggtctcacattggagatccccaagagt
gtcactatgaagaatgtgtagtaactaccactcctccctcaattcagaatggaacataccgcttctgctgtttagcacagatttatgtaatgtcaac
ttactgagaattttccacctctgacacaacaccactcagtcacactcattcatttaaccgagatgagacaataatcattgctttggcatcagtctct
gtattagctgtttgatagttgccttatgctttggatacagaatgttgacaggagaccgtaaacaaggcttcacagtatgaacatgatggaggca
gcagcatccgaaccctcttgatctagataatctgaaactgttgagctgattggccgaggctgatatggagcagtatataaaggctccttgga
10 tgagcgtccagttgctgtaaaagtgtttcctttgcaaacctgcagaattttatcaacgaaaagaacattacagagtgcccttgatggaacatgac
aacattgcccgtttatagttggagatgagagagtcactgcagatggacgcatggaatattgcttgatggagtactatccaatggatcttta
tgcaagtatttaagtctccacacaagtgactgggtaagctcttgccgtctgtcattctgttactagaggactggcttatcttcacacagaattacc
acgaggagatcattataaacctgcaatttccatcgagatttaaacagcagaaatgtcctagtgaataatgatggaacctgtgttattagtgactt
tgagctgtccatgaggctgactggaaatagactggcgccaggagggaagataatgcagccataagcgagggtggcactatcagatata
15 tggcaccagaagtctagaaggagctgtgaactgagggactgtgaatcagctttgaacaagtagacatgtatgctcttgactaatctattg
ggagatatttatgagatgtacagaccttccagggggaatccgtaccagagtaccagatggctttcagacagagggtggaacatcccactt
ttgaggatgagcaggttctgtgtctagggaaaaacagagaccaagttccagaagcctggaaagaaatagcctggcagtgaggctactc
aaggagacaatcgaagactgttggaccaggatgcagaggctcggttactgcacagtgctgaggaaaggatggctgaacttatgatgat
ttgggaaagaaacaaatctgtgagcccaacagtcaatccaatgtctactgctatgcagaatgaacgaacctgtcacataataggcgtgtgcc
20 aaaattggtccttatccagattattcttctcctcatattgaagactctatccatcactgacagcatcgtgaagaatatttctctgagcattcta
tgtccagcacaccttgactataggggaaaaaacgaaattcaattaactatgaacgacagcaagcacaagctgaatcccagccctgaaa
caagtgtcaccagccttccaccaacaacaacacacacaggaactcagccaagtactggcatgactactatatctgagatgccata
cccagatgaaacaaatctgcataccacaaatgttgacagtcgaattgggccaacctgtctgcttacagctgacagaagaagacttggaacc
aacaagctagacccaaagaagtgtgataagaacctcaaggaaagctctgatgagaatctcatggagcactcttaaacagttcagtgccca
25 gaccactgagcagtagttagtctgctttaccactcataaaactgcagtagaagcaactggacagcaggacttcacacagactgcaa
atggccaagcatgtttgattcctgatgttctgctactcagatctatctctcccaagcagcagaacctcccaagagacctactagttgccttg
aacacaaaaattcaacaaagagccccggctaaaatttggcagcaagcacaatcaaactgaaacaagtcgaaactggagttgccaagat
gaataaatcaatgcagcagaacctcatgtgtgacagtcacatgaatggtgtggcaggtagaaccacagtgtaactcccatgctgccac
aaccaatatgccaatgggacagtactatctggccaacaaccaacatagtgacacatagggccaagaaatgtgcagaatcagttattggt
30 gaggacaccggctgaatattaattccagtcctgatgagcatgagcctttactgagacgagagcaacaagctggccatgatgaagggttctg
gatcgtctgtggacaggagggaacggccactagaagggtggccgaactaattccaataacaacaacagcaatccatgttcagaacaagatgtt
cttgacagggtgttcaagcagcagcagatcctgggcatcaaagccagaagagcacagaggcctaattctctggatctttcagccaca
aatgtcctggatggcagcagatatacagataggtgagtcacaacagatggcaaatcaggatcaggtgaaaagatcaagaacgtgtgaaaa
ctccctattcttaagcgggtggcgccctccacctgggtcatctccactgaatcgtggactgtgaagtcaacaataatggcagtaacagggca
35 gttcattccaaatccagcactgctgtttaccttgagaaggaggcactgtacaacatggtgtctaagatataggaatgaactgtctgtga-3'

(SEQ ID NO:336)

AHI1

5'-atgcctacagctgagagtgaagcaaaagtaaaaaccaaagttcgctttgaagaattgcttaagaccacagtgatctaatacggtgaaaaga
 aaaaactgaagaaaaaactgtcaggtctgaagaaaacatctcacctgacactattagaagcaatcttactatataagaaactacaagtg
 5 atgatcccgacactattagaagcaatcttcccatattaaagaaactacaagtgatgatgaagtgtgctgaactaacaacctaagaagag
 cacgagagtcactaaaaacaaattgaggaacacacagttagcaactgaaaatcctaattggtgatgtagtaggaagacaaacaagga
 aagccaaataaaaagggtgataaagacggtgcccagttgactacacaagacctgaaaccggaaactctgagaataagggtgattctacaca
 ccagaaaacacatacaaagccacagccaggcggtgatcatcagaaaagtgagaaggcaaatgagggagagaagagactgatttagaag
 aggatgaagaattgatgcaagcatatcagtgccatgtaactgaagaaatggcaaaggagattaagaggaaaataagaaagaaactgaaag
 10 aacagttgacttactttccctcagatactttatccatgatgacaaactaagcagtgaaaaaaggaaaaagaaaagggaagttccagttcttct
 aaagctgaaacaagtacattgaccatctctggtgacacagttgaaggtgaacaaaagaaagaatcttcagttagatcagtttctcagattctca
 tcaagatgatgaataagctaatggaacaaagcacagaagacagcatgcaagatgatataaaacctaaccacaaaaaaacaaaaaaga
 agactaaagcagttgcagataataatgaagatgttgatggtgatggtgtcatgaaataacaagccgagatagcccgggttatcccaaatgtttg
 cttgatgatgacctgtcttgggagttacattaccggaactgatagacttaagtcatgattttatgatttctacccaatggtaaaaattcatgtggtt
 15 gatgagcactactggtcaatatgtcaagaaagatgatagtgacgacctgttcatcttactatgaaaaagagaatgtggattatattctcttatta
 tgacccagccatgatgatttaaacagttaaatcaagacttcagagtgagggaagaacaaattgtatttaataaaaatttccctattgtctcgagg
 ctctgatgagagtcctaaagtcactctgttcttgagattcttgatttctaagcgtggatgaaattaagaataattctgaggtcaaaaccaagaat
 gtggctttcgaaaaattgctgggcatttctaagcttctgggagccaatggaaatgcaaacatcaactcaaaactcgcttgacgtatattacc
 cacctactaagcctcgatcccattaagtgtgttgaggcattgaatggtgtgcaaatgtccaagaaatcattacccatcaacactgtacgtaa
 20 ctgtaagaggactgaaagttcagactgtataagccatcttaccgctctatgatggctcttcaggaggaaaaaggtaaacagtgacattgtga
 acgtcacatgagtcaagctcagtagacacagaacctggattagaagagtcaaaggaagtaataaagtggaaacgactccctgggcaggctt
 gccgatcccaacaaacacctcttctactaaatgcaggagaacgaggatgttttcttctgatttctccacaatggaagaatattagcagcag
 cttgtgccagccgggatggatatccaatttttatgaaattccttctggacgttcatgagagaattgtgtggccacctcaatatcatttatgatc
 tttctggtcaaaagatgatcactacatcttactcatcatctgatggcactgccaggatatggaaaaatgaaataaacaatacaataactttca
 25 gagttttactcatccttctttgtttacacggctaaattccatccagctgtaagagagctagtagttacaggatgctatgattccatgatacggatat
ggaaagttgagatgagagaagattctgccatattggtccgacagtttgacgttcacaaaagttttatcaactcatttgtttgatactgaaggtca
 tcatatgtattcaggagattgtacaggggtgattgtgtttggaatacctatgtcaagattaatgatttgaacattcagtcaccactggactata
 aataaggaaattaaagaaactgagtttaagggaattccaataagttatttgagattcatcccaatggaaaacgtttgtaattccatacacaaga
 cagtactttgagaattatggatctccgatattagtagcaaggaagttttaggagcagcaaatatcgggagaagattcatagtactttgactcc
 30 atgtgggacttttctgttctggaagtgaggatggtatagtgatgtttggaaccagaaacaggagaacaagtagccatgtattctgacttgcc
 attcaagtcacccattcgagacatttctatcatccattgaaaaataggtgcattctgtgcatttgggcaaatgagccaattcttctgtatattac
 gatttccatgttggccagcaggaggtgaaatgttcaaacgctacaatggaacatttccattacctggaatacaccaaaagtcaagatgcctatg
 tactgtccaaaactacccatcaaggctctttcagattgatgaattgtccacactgaaagttcttcaacgaagatgcagctagtaaaacagag
 gcttgaaactgtcacagaggtgatacgttctgtgtgcaaaagtcaaaaaatctctatttacttaccaccagcagtttctcacaacagctct
 35 aagttaaagcagtcacacatgctgaccgctcaagagattctacatcagtttggttctactcagaccgggattatcagcatagaaagaaagccttg

taaccatcaggtagatacagcaccaacggtagtggtctttatgactacacagcgaatcgatcagatgaactaacatccatcgcgagacatt
atccgagtgTTTTCAAagataatgaagactggtggtatggcagcataggaaagggacaggaaggttatttccagctaatacatgtggtagtg
aaacactgtatcaagaactgcctcctgagataaaggagcgatccccctcttaagccctgaggaaaaactaaaaatagaaaaatctccagctcc
tcaaaagcaatcaataagaacaagtcccaggacttcagactaggctcagaatctatgacacattctgaaatgagaaaagaacagagcc
5 atgaggaccaaggacacataatggatacacggatgaggaagaacaagcaagcaggcagaaaagtcactctaatagagta-3' (SEQ ID
NO:337)

FANCC

5'-atggctcaagattcagtagatcttcttgattatcagtttggatgcagaagcttctgtatgggatcaggcttccatttgaaaccagca
agacacctgtcttcagtggtcagttccaggagttcctaaggaagatgtatgaagcctgaaagagatggattctaatacagtcattgaaagat
10 tccccacaattggtcaactgttgcaaaagcttggtggaatcctttatttagcatatgtatgaaagccaaaaattctaataatggtgcttatgtgtc
taattaacaaagaaccacagaattctggacaataaaaacttaactcctggatacaggggtgattatctcatatactttcagcactcagatttgataa
agaagtgctcttttactcaaggtcttggtatgcacatatagattactatcctggttgcttaaaaatggtttatcattagcgtctgaactcaga
gagaatcatctaatggatttaacactcaaaggcgaatggctcccgagcgagtggtcctgtcacgagtttggtcccactattaccctgaca
gatgttgacccctggtgaggtctcctcatctgtcatggacgtgaacctcaggaatcctccagccagagttctttgaggctgtaaacgaggc
15 cattttgctgaagaagatttctccccatgtcagctgtagtctgcctctggcttcggcacctcccagcctgaaaaagcaatgctgcatcttttga
aaagctaattccagtgagagaaattgtctgagaaggatcgaatgcttataaaagattcatcgctgcctcaagcagcctgccacctgccatat
tccgggtgttgatgagatgttcaggtgtgcactcctggaaaccgatggggcctggaaatcatagccactattcaggtgtttacgcagtgctttg
tagaagctctggagaaagcaagcagcgtcggtttgactcaagacctcttctacattctccatctcttgcatggtgtctgtgcaag
acctcaagatatccctcggggacactggctccagacactgaagcatatttctgaactgtcagagaagcagttgaagaccagactcatgggtc
20 ctgaggaggtcccttgagagctggttctgttcattcattcggaggatgggtgagatggtggcagagcaattactgatgtcggcagccgaa
ccccccacggcctgtgtggtcttggtcttactacggcccccgatgggaggcagcagagagcacagactatggtccaggtgaaggcc
gtgctggggcacctcctggcaatgtccagaagcagcagctctcagccaggacctgcagacggtagcaggacagggcacagacagacc
tcagagctctgcacaacagctgatcaggcaccttctcctaacttctgctctgggtcctggaggccacagatcgctgggatgtcatcaccc
tgatggctcacactgtgagataactcacgagatcattggcttcttgaccagacctgtacagatggaatcgcttggcattgaaagccctagat
25 cagaaaaactggccgagagctccttaaagagctgcgaactcaagtctag-3' (SEQ ID NO:338)

MYBPC3

5'-atgcctgagccggggaagaagccagctcagcttttagcaagaagccacggtcagtggaagtggccgaggcagccctgccgtgttcga
ggccgagacagagcgggaggtgaaggtgcgtggcagcgaggagcagtgacatcagcgccagcaacaagtagcgctggccaca
gagggcacacggcatagctgacagtgcggaagtgggccccctgccgaccagggatcttacgcagtcattgtggtcctccaaggtcaagttc
30 gacctcaaggtcatagaggcagagaaggcagagcccatgctggccccctgccccctgctgaggccactggagccccctggagaagccc
cgccccagccgtgagctgggagaaagtgtcccaagtccaaaggttaagctcagcagctctcaatggtcctaccctggagccccgat
gacccattggcctctctgtatgcgccacaggatggcgaggtgacctgggtggcagcatcaccttctcagcccggtggccggcgccagc
ctctgaagccgctgtgttaagtggtcaagggcaaatgggtggacctgagcagcaaggtgggcccagcacctgcagctgcacgacagcta
cgaccgcccagcaaggtctatctgttcgagctgcacatcaccgatgccagcctgccttactggcagctaccgctgtgaggtgtccaccaag
35 gacaaatttgactgtccaactcaatctactgtccacgaggccatgggacccggagacctggacctctatcagcctccgcccagcagcct

ggctggaggtggtcggcggatcagtgatagccatgaggacactgggattctggactcagctcactgctgaaaaagagagacagttccgga
 ccccgagggactcgaagctggaggcaccagcagaggagacgtgtgggagatcctacggcaggcaccatctgagtacgagcgcatcg
 cctccagtacggcgctcactgacctgcgcgcatgctaaagggtcaaggcgatgagcgcatgagaagaagagcacagccttcagaa
 gaagctggagccggcctaccaggtgagcaaaggccacaagatccggctgaccgtggaactggctgacctgacgtgaggtcaaattggctc
 5 aagaatggccaggagatccagatgagcggcagcaagtacatctttgagtcacgtgccaagcgtaccctgacctagaccagtgctcattg
 gcgagcagcagcctaccagtgctggtgggtggcgagaagtgtacacggagctctttgtgaaagagccccctgtgctcatcacgcgccc
 cttggaggaccagctggtgatggtggggcagcgggtggagtttgagtgtaagtatcgaggagggggcgcaagtcaaattggctgaagga
 cggggtggagctgaccgggaggagaccttcaataccgggttcaagaaggacgggagagacaccacctgatcatcaacgagggcatgctg
 gaggacgcggggcactatgcactgtgcactagcggggccaggcgctggctgagctcattgtgcaggaagaagctggaggtgtaccag
 10 agcatcgagacctgatggtggcgcaaaggaccaggcggtgttcaaatgtgaggtctcagatgagaatgttcggggtgtgtggctgaagaa
 tgggaaggagctggtgcccagacggcgataaagggtgtccacatcgggcggtccacaaactgaccattgacgacgtcacacctgccgacg
 aggtgactacagctttgtcccagagggttcgctgcaacctgtcagccaagctccattcatggaggtcaagattgacttcgtaccaggcag
 gaacctccaagatccacctggactgccaggccgcataccagacaccattgtggtttagctggaaataagctacgtctggacgtcctatctc
 tggggacctgtctccactgtgatctggcagaaggctatcacgcagggaataaggcccgaggccagccccagatgccccagaggac
 15 acaggtgacagcgatgagtggtgttgacaagaagctgctgtgtgagaccgagggcggtccgctggagaccaccaaggaccgcagc
 atcttcacggtcagggggcagagaaggaagatgagggcgtctacacggtcacagtgaagaacctgtggggcaggaccagggtcaacctc
 acagtcaaggctacgacgtgccagacgcacctgcggccccaagatcagcaacgtgggagaggactcctgcacagtacagtgggagccgc
 ctgcctacgatggcgggcagccatctgggtacatcctggagcgcaagaagaagagctaccggtggatggcggtgaacttcgacctg
 attcaggagctgagtcataagcgcgcatgatcagggcggtgtgtacagatgcgctctacgggtcaacgccatcgcatgtccag
 20 gcccagccctgctccagccctcatgctatcggtccccccagcgaacccaccacctggcagtagaggacgtctctgacaccaggtctccc
 tcaagtggcgccccagagcggtgggagcaggaggcctggatggctacagcgtggagtactgccagagggtgctcagagtgggtgg
 ctgcctgcagggggtgacagagcacatcgatactggtgaaggacctgccacggggggccggtgctttccgagtgcgggcacacaat
 atggcagggctggagccctgttaccaccacggagccggtgacagtgcaggagatcctgcaacggccacggcttcagctgccaggcacct
 gcgcagaccattcagaagaaggcgggagcctgtgaaccttctcatcccttcaggggcaagccccggcctcaggtgacctggaccaaaga
 25 ggggcagccctggcaggcgaggaggtgagcatccgaacagccccacagacaccatcctgttcacggggccgctcgccgcgtgcattcag
 gcacttaccaggtgacggtgacgattgagaacatggaggacaaggccacgctggtgctgcaggtgttgacaagccaagtcctcccaggatc
 tccgggtgactgacgctggggtcttaattgtggtctggagtgaagccacccaggatgtcggcaacacggagctctgggggtacacagtg
 cagaagccgacaagaagaccatggagtgttcacctgttgagcattaccgccgacccactgctggtgccagagctcatcattggcaat
 ggctactactccgctcttcagccagaatatggttggtttagtacagagcggccaccaccaaggagcccgctttatccccagaccaggcat
 30 cacctatgagccaccaactataaggccctggacttctccgagggcccaagcttcaccagccctggtgaacgctcggtcatcggggtac
 actgctatgctctgtgtgtccgggtagcccaagcccaagatttctggttcaagaatggcctggacctgggagaagacgcccgttccg
 catgttcagcaagcaggaggtgtgactctggagattagaaagccctgcccctttacgggggcatctatgtctgcagggccaccaacttacg
 ggcgaggcacggtgtgagtgccgctggaggtgcagtgctcagtga-3' (SEQ ID NO:339)

IL2RG

35 5'-atgtgaagccatcattaccattcacatccctcttattcctgcagctgccctgctgggagtggggtgaacacgacaattctgacgccaatg

ggaatgaagacaccacagctgatttctctgaccactatgccactgactccctcagtggttccactctgcccctcccagaggttcagtgtttg
 ttcaatgtcgagtacatgaattgcacttggaacagcagctctgagcccagcctaccaacctcactctgcattattggtacaagaactcggataat
 gataaagtcagaagtgcagccactatctattctctgaagaaatcacttctggctgtcagttgcaaaaaaaggagatccacctctaccaaacattt
 gttgttcagctccaggacccacgggaacccaggagacaggccacacagatgctaaaactgcagaatctggtgatcccctgggctccagagaa
 5 cctaacacttcacaaactgagtgaatcccagctagaactgaactggaacaacagattctgaaccactggttgagcacttggtgcagtaccgg
 actgactgggaccacagctggactgaacaatcagtggtattatagacataagttctccttgctagtgtggatgggcagaaacgctacagtttc
 gtgttcggagccgctttaaccactctgtggaagtgtcagcattggagtgaatggagccaccaatccactgggggagcaatacttcaaaag
 agaatctttcctggttgattggaagccgtggttatctgttggtccatgggattgattatcagccttctgtgtgtatttctggctggaacggac
 gatgccccgaattcccaccctgaagaacctagaggatctgttactgaataccacgggaactttcggcctggagtgtgtgtctaagggactgg
 10 ctgagagtctgcagccagactacagtgaaacgactctgcctcgtcagtgagattccccaaaaggaggggcccttggggaggggctggggcc
 tcccatgcaaccagcatagcccctactgggccccccatgttacaccctaaagcctgaaacctga-3' (SEQ ID NO:340)

Discusión

Las tecnologías de edición del genoma están revolucionando la investigación
 15 biomédica. Las nucleasas altamente activas, como las nucleasas con dedos de zinc (ZFN)¹, las
 nucleasas efectoras de tipo activador de la transcripción (TALEN)²⁻⁴ y las proteínas Cas del sistema
 CRISPR (repeticiones palindrómicas cortas agrupadas regularmente interespaciadas)⁵⁻⁷ se han
 diseñado con éxito para manipular genoma en una miríada de organismos. Recientemente, se han
 aprovechado las desaminasas para cambiar con precisión el código genético sin romper el ADN de
 20 doble cadena. Al acoplar una citidina o una adenosina desaminasa con el sistema CRISPR-Cas9, los
 investigadores crearon editores de base programables que permiten la conversión de C • G en T • A o
 A • T en G • C en el ADN genómico⁸⁻¹⁰, ofreciendo nuevas oportunidades para corregir mutaciones que
 causan enfermedades.

Aparte del ADN, el ARN es un objetivo atractivo para la corrección genética porque la
 25 modificación del ARN podría alterar la función de la proteína sin generar cambios permanentes en el
 genoma. Las adenosina desaminasas ADAR se aprovechan actualmente para lograr una edición de
 base precisa en los ARN. Se han identificado tres tipos de proteínas ADAR en mamíferos, ADAR1
 (isoformas p110 y p150), ADAR2 y ADAR3 (catalítico inactivo)^{11,12}, cuyos sustratos son ARN de cadena
 doble, en los que una adenosina (A) se empareja erróneamente con una citosina (C) se desamina
 30 preferentemente a inosina (I). Se cree que la inosina imita a la guanosina (G) durante la traducción^{13,14}.
 Para lograr la edición de ARN dirigida, la proteína ADAR o su dominio catalítico se fusionó con un
 péptido λ N¹⁵⁻¹⁷, una etiqueta SNAP¹⁸⁻²² o una proteína Cas (dCas13b)²³, y se diseñó un ARN guía para
 reclutar la proteína ADAR quimérica para el sitio específico. Alternativamente, también se informó que
 la sobreexpresión de las proteínas ADAR1 o ADAR2 junto con un ARN guía que lleva el motivo R/G
 35 permite la edición de ARN dirigida²⁴⁻²⁷.

Todos estos métodos de edición de ácidos nucleicos informados en el sistema de mamíferos se basan en la expresión ectópica de dos componentes: una enzima y un ARN guía. Aunque estos sistemas binarios funcionan de manera eficiente en la mayoría de los estudios, algunos obstáculos inherentes limitan sus amplias aplicaciones, especialmente en terapias. Debido a que la administración *in vivo* más eficaz para la terapia génica es a través de vectores virales²⁸, y los vectores de virus adenoasociados (AAV) altamente deseables están limitados con el tamaño de la carga (~4.5 kb), lo que dificulta la adaptación tanto de la proteína como del ARN guía^{29,30}. Recientemente se ha informado que la sobreexpresión de ADAR1 confiere oncogenicidad en mielomas múltiples debido a la hiperedición aberrante en los ARN³¹ y genera importantes ediciones globales fuera de la focalización³². Además, la expresión ectópica de proteínas o sus dominios de origen no humano tiene un riesgo potencial de provocar inmunogenicidad^{30,33}. Además, la inmunidad adaptativa preexistente y la respuesta al daño del ADN mediada por p53 pueden comprometer la eficacia de la proteína terapéutica, como Cas9³⁴⁻³⁸. Aunque se ha intentado utilizar un mecanismo endógeno para la edición de ARN, esto solo se intentó inyectando el transcripto objetivo preensamblado:ARN dúplex en embriones de *Xenopus*³⁹. Son muy necesarias tecnologías alternativas para la edición robusta de ácidos nucleicos que no dependan de la expresión ectópica de proteínas. Aquí, desarrollamos un enfoque novedoso que aprovecha la ADAR endógena para la edición de ARN. Demostramos que la expresión de un ARN guía diseñado deliberadamente permite una edición eficiente y precisa de ARN endógenos y corrige mutaciones patógenas. Esta plataforma de edición de ácido nucleico unario puede abrir nuevas vías para la terapéutica y la investigación.

En particular, mostramos que la expresión de un ARNar lineal con una longitud adecuada es capaz de guiar a las proteínas ADAR endógenas para editar adenosina en inosina en los transcriptos dirigidos. Este sistema, denominado LEAPER, utiliza proteínas ADAR endógenas para lograr la edición programable de ácidos nucleicos, por lo que posee ventajas sobre los enfoques existentes.

La rara cualidad de LEAPER es su simplicidad porque solo se basa en un tamaño pequeño de molécula de ARN para dirigir las proteínas endógenas para la edición de ARN. Esto recuerda al ARNi, en el que un pequeño ARNs podría invocar un mecanismo nativo para la degradación del ARN dirigido⁵¹. Debido al pequeño tamaño, el ARNar se podría administrar fácilmente mediante una variedad de vehículos virales y no virales. A diferencia del ARNi, LEAPER cataliza el cambio preciso de A a I sin generar corte o degradación de los transcriptos objetivo (figura 18A). Aunque el requisito de longitud para el ARNar es más largo que el ARNi, no induce efectos inmunoestimuladores a nivel celular (figura 22E, figura 22F y figura 29E) ni afecta la función de las proteínas ADAR endógenas (figuras 22A, 22B), lo que hace es una estrategia segura para la selección de ARN. Sorprendentemente, se ha informado que la expresión ectópica de las proteínas ADAR o sus

dominios catalíticos induce importantes ediciones globales fuera del objetivo³² y posiblemente desencadena el cáncer³¹.

Recientemente, varios grupos informaron que el editor de bases de citosina podría generar variantes sustanciales de un solo nucleótido fuera del objetivo en embriones de ratón, líneas celulares de arroz o humanas debido a la expresión de una proteína efectora, lo que ilustra la ventaja de LEAPER para una posible aplicación terapéutica⁵²⁻⁵⁴. Afortunadamente, LEAPER permite una edición eficiente mientras provoca una rara edición global fuera del objetivo (figura 20 y figura 21). Además, LEAPER podría minimizar la inmunogenicidad potencial o superar los obstáculos de administración que comúnmente comparten otros métodos que requieren la introducción de proteínas extrañas.

Para LEAPER, recomendamos usar ARNar con un tamaño mínimo superior a 70-nt para lograr la actividad deseada. En el contexto nativo, las proteínas ADAR editan de forma no específica las repeticiones Alu que tienen un dúplex de más de 300-nt⁵⁵. Es de destacar que las repeticiones Alu forman un dúplex intramolecular estable, mientras que LEAPER da como resultado un dúplex intermolecular entre ARNar y ARNm o pre-ARNm, que se supone que es menos estable y más difícil de formar. Por lo tanto, planteamos la hipótesis de que un dúplex de ARN de más de 70-nt es estequiométricamente importante para reclutar o acoplar proteínas ADAR para una edición eficaz. De hecho, un ARNar más largo dio como resultado un mayor rendimiento de edición tanto en los reporteros expresados ectópicamente como en los transcritos endógenos (figura 16D y figura 17B). Sin embargo, debido a que las proteínas ADAR desaminan promiscuamente la base de adenosina en el dúplex de ARN, los ARNar más largos pueden incurrir en más objetivos desviados dentro de la ventana de direccionamiento.

Si bien LEAPER podría apuntar eficazmente a los transcritos nativos, sus eficiencias de edición y las tasas fuera del objetivo variaron. Para el direccionamiento del transcripto de PPIB, podríamos convertir el 50% de la adenosina dirigida en inosina sin objetivos desviados evidentes dentro de las ventanas de cobertura (figuras 17B, 17F). Los objetivos desviados se volvieron más severos para otros transcritos. Hemos logrado reducir los objetivos desviados, como la introducción de desajustes A-G o desajustes consecutivos para reprimir la edición no deseada. Sin embargo, demasiados desajustes podrían disminuir la eficiencia en el objetivo. Ponderando la eficiencia y los posibles objetivos desviados, recomendamos ARNar con una longitud que va de 100 a 150-nt para la edición en transcritos endógenos. Si hay una opción, es mejor seleccionar regiones con menos adenosina para minimizar la posibilidad de ediciones no deseadas. De manera alentadora, no hemos detectado ningún objetivo desviado de los dúplex del transcripto dirigido a ARNar (figura 20).

Hemos optimizado el diseño del ARNar para lograr una mayor eficiencia de edición y hemos demostrado que LEAPER podría aprovecharse para manipular la función genética o corregir la mutación patógena. También hemos demostrado que LEAPER no se limita a trabajar solo en UAG, sino

que funciona con posiblemente cualquier adenosina independientemente de sus nucleótidos flanqueantes (figura 16F, G y figura 17C). Tal flexibilidad es ventajosa para la posible corrección terapéutica de enfermedades genéticas causadas por determinadas mutaciones puntuales. Curiosamente, al editar los transcritos de *IDUA*, el ARNar que se dirige al pre-ARNm es más eficaz que el que se dirige al ARN maduro, lo que indica que los núcleos son los principales sitios de acción de las proteínas ADAR y LEAPER podría aprovecharse para manipular el empalme modificando los sitios de empalme dentro de los pre-ARNm. Además, LEAPER ha demostrado una alta eficiencia para dirigir simultáneamente a múltiples transcritos de genes (figura 17D). Esta capacidad de multiplexación de LEAPER podría desarrollarse para curar ciertas enfermedades poligenéticas en el futuro.

Es beneficioso realizar una corrección genética a nivel de ARN. En primer lugar, la edición de transcritos dirigidos no cambiaría permanentemente el repertorio del genoma o transcriptoma, lo que hace que los enfoques de edición de ARN sean más seguros para la terapéutica que los medios de edición del genoma. Además, la edición transitoria es muy adecuada para el control temporal del tratamiento de enfermedades causadas por cambios ocasionales en un estado específico. En segundo lugar, LEAPER y otros métodos de edición de ARN no introducirían DSB en el genoma, evitando el riesgo de generar deleciones indeseables de grandes fragmentos de ADN³⁷. Los métodos de edición de bases de ADN que adoptan la nickasa Cas9 aún podrían generar indeles en el genoma⁸. Además, independientemente de la maquinaria de reparación del ADN nativo, LEAPER también debería funcionar en células post-mitosis como las células del cerebelo con alta expresión de ADAR2¹¹.

Hemos demostrado que LEAPER podría aplicarse a un amplio espectro de tipos de células, como líneas celulares humanas (figura 14C), líneas celulares de ratón (figura 14D) y células primarias humanas, incluidas las células T primarias (figura 27 y figura 28D). La edición eficiente a través de la administración lentiviral o el oligo sintetizado proporciona un mayor potencial para el desarrollo terapéutico (figura 28). Además, LEAPER podría producir cambios fenotípicos o fisiológicos en una variedad de aplicaciones, incluida la recuperación de la actividad reguladora transcripcional de p53 (figura 7), la corrección de mutaciones patogénicas (figura 26) y la restauración de la actividad de α -L-iduronidasa en fibroblastos primarios derivados de pacientes con síndrome de Hurler (figura 29). Por tanto, se puede considerar que LEAPER tiene un enorme potencial en el tratamiento de enfermedades.

Stafforst y sus colegas informaron de un método de edición de ARN nuevo y aparentemente similar, llamado RESTORE, que funciona mediante el reclutamiento de ADAR endógenas utilizando oligonucleótidos antisentido sintéticos⁵⁶. La diferencia fundamental entre RESTORE y LEAPER radica en la naturaleza distinta del ARN guía para reclutar ADAR endógenas. El ARN guía de RESTORE se limita a oligonucleótidos antisentido quimiosintéticos (ASO) dependiendo de la modificación química compleja, mientras que el ARNar de LEAPER se puede generar de diversas

formas, mediante síntesis química y expresión a partir de vectores virales o no virales (figura 28 y figura 29). Es importante destacar que, al estar muy modificados químicamente, los ASO están restringidos para actuar transitoriamente en el tratamiento de enfermedades. Por el contrario, el ARNar podría producirse a través de la expresión, una característica particularmente importante para el propósito de la edición constante.

Todavía quedan espacios para mejorar la eficiencia y la especificidad de LEAPER. Debido a que LEAPER se basa en la ADAR endógena, el nivel de expresión de las proteínas ADAR en las células objetivo es uno de los determinantes para una edición exitosa. Según el informe anterior⁵⁷ y nuestras observaciones (figuras 14A, 14B), la ADAR1^{p110} se expresa de forma ubicua en los tejidos, lo que garantiza la amplia aplicabilidad de LEAPER. La ADAR1^{p150} es una isoforma⁵⁸ inducible por interferón y ha demostrado ser funcional en LEAPER (figura 11E, figura 12B). Por tanto, la cotransfección de ARN estimuladores de interferón con ARNar podría mejorar aún más la eficiencia de edición en determinadas circunstancias. Alternativamente, dado que ADAR3 desempeña funciones inhibitoras, la inhibición de ADAR3 podría mejorar la eficiencia de edición en las células que expresan ADAR3. Además, la modificación adicional de ARNar podría aumentar su eficiencia de edición. Por ejemplo, el ARNar fusionado con cierto andamio de reclutamiento de ADAR puede aumentar la concentración de proteína ADAR local y, en consecuencia, aumentar el rendimiento de la edición. Hasta ahora, solo pudimos aprovechar las proteínas ADAR1/2 endógenas para la conversión de la base A a I. Es emocionante explorar si más mecanismos nativos podrían aprovecharse de manera similar para la modificación de elementos genéticos, especialmente para realizar una potente edición de ácidos nucleicos.

En conjunto, proporcionamos una prueba de principio de que la maquinaria endógena en las células podría cooptarse para editar los transcritos de ARN. Demostramos que LEAPER es un sistema simple, eficiente y seguro, que arroja luz sobre un camino novedoso para la investigación y la terapéutica basadas en la edición de genes.

Referencias

- 1 Porteus, M. H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nat Biotechnol* **23**, 967-973 (2005).
- 30 2 Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512 (2009).
- 3 Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
- 4 Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143-148 (2011).
- 35

- 5 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
- 6 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).
- 5 7 Mali, P. *et al.* ARN-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
- 8 Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
- 9 Ma, Y. *et al.* Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* **13**, 1029-1035 (2016).
- 10 Gaudelli, N. M. *et al.* Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471 (2017).
- 11 Tan, M. H. *et al.* Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249-254 (2017).
- 15 12 Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **79**, 321-349 (2010).
- 13 Bass, B. L. & Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**, 1089-1098 (1988).
- 14 Wong, S. K., Sato, S. & Lazinski, D. W. Substrate recognition by ADAR1 and ADAR2. *RNA* **7**, 846-858 (2001).
- 20 15 Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc Natl Acad Sci U S A* **110**, 18285-18290 (2013).
- 16 Sinnamon, J. R. *et al.* Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc Natl Acad Sci U S A* **114**, E9395-E9402 (2017).
- 25 17 Montiel-Gonzalez, M. F., Vallecillo-Viejo, I. C. & Rosenthal, J. J. An efficient system for selectively altering genetic information within mRNAs. *Nucleic Acids Res* **44**, e157 (2016).
- 18 Hanswillemenke, A., Kuzdere, T., Vogel, P., Jekely, G. & Stafforst, T. Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein. *J Am Chem Soc* **137**, 15875-15881 (2015).
- 30 19 Schneider, M. F., Wettengel, J., Hoffmann, P. C. & Stafforst, T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res* **42**, e87 (2014).
- 20 20 Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching Protein Localization by Site-Directed RNA Editing under Control of Light. *ACS synthetic biology* **6**, 1642-1649 (2017).
- 35

- 21 Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angewandte Chemie* **53**, 6267-6271 (2014).
- 22 Vogel, P. *et al.* Efficient and precise editing of endogenous transcripts with SNAP-tagged
5 ADARs. *Nat Methods* **15**, 535-538 (2018).
- 23 Cox, D. B. T. *et al.* RNA editing with CRISPR-Cas13. *Science* **358**, 1019-1027 (2017).
- 24 Fukuda, M. *et al.* Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Scientific reports* **7**, 41478 (2017).
- 25 Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. Harnessing human
10 ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res* **45**, 2797-2808 (2017).
- 26 Heep, M., Mach, P., Reautschnig, P., Wettengel, J. & Stafforst, T. Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing-G/C Substitution Stabilizes GuideRNAs against Editing. *Genes (Basel)* **8** (2017).
- 15 27 Katrekar, D. *et al.* In vivo RNA editing of point mutations via RNA-guided adenosine deaminases. *Nat Methods* **16**, 239-242 (2019).
- 28 Yin, H., Kauffman, K. J. & Anderson, D. G. Delivery technologies for genome editing. *Nat Rev Drug Discov* **16**, 387-399 (2017).
- 29 Platt, R. J. *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell*
20 **159**, 440-455 (2014).
- 30 Chew, W. L. *et al.* A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods* **13**, 868-874 (2016).
- 31 Teoh, P. J. *et al.* Aberrant hyperediting of the myeloma transcriptome by ADAR1 confers oncogenicity and is a marker of poor prognosis. *Blood* **132**, 1304-1317 (2018).
- 25 32 Vallecillo-Viejo, I. C., Liscovitch-Brauer, N., Montiel-Gonzalez, M. F., Eisenberg, E. & Rosenthal, J. J. C. Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA biology* **15**, 104-114 (2018).
- 33 Mays, L. E. & Wilson, J. M. The complex and evolving story of T cell activation to AAV vector-encoded transgene products. *Mol Ther* **19**, 16-27 (2011).
- 30 34 Wagner, D. L. *et al.* High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat Med* **25**, 242-248 (2019).
- 35 Simhadri, V. L. *et al.* Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. *Mol Ther Methods Clin Dev* **10**, 105-112 (2018).
- 36 Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med* **25**, 249-254 (2019).
- 35

- 37 Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med* **24**, 927-930 (2018).
- 38 Ihry, R. J. *et al.* p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med* **24**, 939-946 (2018).
- 5 39 Woolf, T. M., Chase, J. M. & Stinchcomb, D. T. Toward the therapeutic editing of mutated RNA sequences. *Proc Natl Acad Sci U S A* **92**, 8298-8302 (1995).
- 40 Zheng, Y., Lorenzo, C. & Beal, P. A. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res* **45**, 3369-3377 (2017).
- 41 Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting
10 CRISPR effector. *Science* **353**, aaf5573 (2016).
- 42 Daniel, C., Widmark, A., Rigardt, D. & Ohman, M. Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome. *Genome Biol* **18**, 195 (2017).
- 43 Chen, C. X. *et al.* A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **6**, 755-767
15 (2000).
- 44 Savva, Y. A., Rieder, L. E. & Reenan, R. A. The ADAR protein family. *Genome Biol* **13**, 252 (2012).
- 45 Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat Rev Mol Cell Biol* **17**, 83-96 (2016).
- 20 46 Floquet, C., Deforges, J., Rousset, J. P. & Bidou, L. Rescue of non-sense mutated p53 tumor suppressor gene by aminoglycosides. *Nucleic Acids Res* **39**, 3350-3362 (2011).
- 47 Kern, S. E. *et al.* Identification of p53 as a sequence-specific DNA-binding protein. *Science* **252**, 1708-1711 (1991).
- 48 Doubrovin, M. *et al.* Imaging transcriptional regulation of p53-dependent genes with positron emission tomography in vivo. *Proc Natl Acad Sci U S A* **98**, 9300-9305 (2001).
- 25 49 Landrum, M. J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* **44**, D862-868 (2016).
- 50 Ou, L. *et al.* ZFN-Mediated In Vivo Genome Editing Corrects Murine Hurler Syndrome. *Mol Ther* **27**, 178-187 (2019).
- 30 51 Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811 (1998).
- 52 Zuo, E. *et al.* Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* (2019).
- 53 Jin, S. *et al.* Cytosine, but not adenine, base editors induce genome-wide off-target mutations
35 in rice. *Science* (2019).

- 54 Kim, D., Kim, D. E., Lee, G., Cho, S. I. & Kim, J. S. Genome-wide target specificity of CRISPR
RNA-guided adenine base editors. *Nat Biotechnol* **37**, 430-435 (2019).
- 55 Levanon, E. Y. *et al.* Systematic identification of abundant A-to-I editing sites in the human
transcriptome. *Nat Biotechnol* **22**, 1001-1005 (2004).
- 5 56 Merkle, T. *et al.* Precise RNA editing by recruiting endogenous ADARs with antisense
oligonucleotides. *Nat Biotechnol* **37**, 133-138 (2019).
- 57 Wagner, R. W. *et al.* Double-stranded RNA unwinding and modifying activity is detected
ubiquitously in primary tissues and cell lines. *Mol Cell Biol* **10**, 5586-5590 (1990).
- 58 Patterson, J. B. & Samuel, C. E. Expression and regulation by interferon of a
10 double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms
of the deaminase. *Mol Cell Biol* **15**, 5376-5388 (1995).
- 59 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases.
Nat Methods **6**, 343-345 (2009).
- 60 Zhou, Y., Zhang, H. & Wei, W. Simultaneous generation of multi-gene knockouts in human
15 cells. *FEBS Lett* **590**, 4343-4353 (2016).
- 61 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
- 62 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome
Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-33 (2013).
- 63 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from
20 high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
- 64 Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes.
Nature **491**, 56-65 (2012).
- 65 Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression
analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* **11**,
25 1650-1667 (2016).