

## TRANSACTIVADORES DE DOMINIOS DE UNIÓN A ADN Y USOS DE ESTOS

### SOLICITUD RELACIONADA

- 5 La presente solicitud reivindica el beneficio de prioridad, de conformidad con el art. 119(e) del título 35 del USC, de la fecha de presentación de la solicitud provisional estadounidense con el n.º de serie 62/810,005, titulada «ZINC FINGER PROTEIN TRANSACTIVATORS AND USES THEREOF» y presentada el 25 de febrero de 2019; cuyo contenido se incorpora en su totalidad a la presente mediante referencia.

10

### ANTECEDENTES

- La regulación de la expresión de genes diana ha emergido como un área importante de la investigación biomédica. La regulación al alza de la expresión génica puede corregir condiciones de haploinsuficiencia como resultado de la disminución de la expresión génica. Normalmente, la haploinsuficiencia tiene lugar cuando una o más mutaciones de pérdida de función se encuentran presentes en al menos una copia de un gen. Las estrategias basadas en AAV de mejora génica para el tratamiento de enfermedades asociadas a la haploinsuficiencia son obstaculizadas por la capacidad de encapsulación de los vectores de rAAV tradicionales.

20

### COMPENDIO

- Algunos aspectos de la descripción están vinculados con ácidos nucleicos aislados y vectores de AAV recombinantes para el suministro de genes. La descripción se basa, en parte, en composiciones (p. ej., rAAV y vectores de rAAV) y métodos para regular la expresión de genes diana, donde el gen diana presenta haploinsuficiencia, tal como el *SCN1A*. En algunas modalidades, la descripción proporciona proteínas de fusión que comprenden un dominio de unión a ADN, tal como una proteína de dedos de cinc (ZFP, por sus siglas en inglés) Cys2-His2, y un dominio regulador de la transcripción. En algunas modalidades, las composiciones descritas en la descripción comprenden una proteína de fusión que comprende un dominio de unión a ADN (p. ej., una ZFP, un dominio efector de tipo activador de la transcripción (TALE, por sus siglas en inglés), etc.) que se fusiona con un dominio regulador de la transcripción. En algunas modalidades, las proteínas de fusión descritas en la descripción aumentan la expresión de un gen diana (p. ej., *SCN1A*) y, por lo tanto, son útiles para tratar enfermedades caracterizadas por la expresión deficiente del gen diana (p. ej., enfermedades asociadas a la

haploinsuficiencia de un gen diana) en una célula o un sujeto en comparación con una célula o un sujeto normal.

Por consiguiente, en algunos aspectos, la descripción proporciona un ácido nucleico aislado que comprende un transgén configurado para expresar al menos un dominio de unión a ADN fusionado con al menos un dominio regulador de la transcripción, donde el dominio de unión a ADN se une a un gen diana o una región reguladora (p. ej., una secuencia potenciadora, una secuencia promotora, una secuencia represora, etc.) de un gen diana (p. ej., en un sujeto o una célula), donde el gen diana codifica un canal de sodio dependiente de voltaje (p. ej.,  $Na_v1.1$ ). En algunas modalidades, un gen diana es un gen *SCN1A*. En algunas modalidades, un transgén está flanqueado por repeticiones terminales invertidas (ITR, por sus siglas en inglés) del virus adenoasociado (AAV, por sus siglas en inglés). En algunas modalidades, el o los dominios de unión a ADN se unen a un gen diana (p. ej., en un sujeto o una célula) y el dominio regulador de la transcripción modifica, p. ej., regula al alza, la expresión de un gen diana.

En algunos aspectos, la descripción proporciona un AAV recombinante (rAAV) que comprende: un ácido nucleico que comprende un transgén que codifica al menos un dominio de unión a ADN fusionado con al menos un dominio regulador de la transcripción, donde el dominio de unión a ADN se une a un gen diana o una región reguladora de un gen diana (p. ej., en un sujeto o una célula), donde el gen diana codifica un canal de sodio dependiente de voltaje (p. ej.,  $Na_v1.1$ ) y al menos una proteína de cápside. En algunas modalidades, un gen diana es un gen *SCN1A*. En algunas modalidades, un transgén está flanqueado por repeticiones terminales invertidas (ITR) de AAV.

En algunas modalidades, al menos un dominio de unión a ADN se une a un gen diana (p. ej., en un sujeto o una célula) y el dominio regulador de la transcripción modifica, p. ej., regula al alza, la expresión de un gen diana en el sujeto.

En algunas modalidades, al menos un dominio de unión a ADN se une a una región no traducida de un gen diana. En algunas modalidades, un dominio de unión a ADN se une a una región reguladora del gen diana, opcionalmente una secuencia potenciadora, una secuencia promotora y/o una secuencia represora.

En algunas modalidades, un dominio de unión a ADN se une entre 2 y 2000 pb en dirección 5' o 2 y 2000 pb en dirección 3' con respecto a una región reguladora (p. ej., una secuencia potenciadora, una secuencia promotora, una secuencia represora, etc.) de un gen diana.

En algunas modalidades, al menos un dominio de unión a ADN codifica una proteína de dedos de cinc (ZFP), un efector de tipo activador de la transcripción (TALE), una proteína dCas (p. ej., dCas9 o dCas12a) y/o un homeodominio. En algunas modalidades, al menos un dominio de unión a ADN se une a una secuencia de ácido nucleico indicada en cualquiera de las SEQ ID NO: 5-7. En algunas

modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento codificada por un ácido nucleico con una secuencia indicada en cualquiera de las SEQ ID NO: 11-16, 23-28 o 35-40. En algunas modalidades, al menos un dominio de unión a ADN es una proteína de dedos de cinc que comprende una secuencia de aminoácidos indicada en cualquiera de las SEQ ID NO: 17-22, 29-34 o 41-46.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 11, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 12, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 13, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 14, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 15 y/o una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 16. En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una secuencia de aminoácidos de la SEQ ID NO: 57. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 % o al menos un 99 % de identidad de secuencia con respecto a la secuencia de aminoácidos de la SEQ ID NO: 57.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 23, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 24, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 25, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 26, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 27 y/o una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 28. En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una secuencia de aminoácidos de la SEQ ID NO: 59. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 % o al menos un 99 % de identidad de secuencia con respecto a la secuencia de aminoácidos de la SEQ ID NO: 59.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 35, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO:

36, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 37, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 38, una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 39 y/o una hélice de reconocimiento codificada por un ácido nucleico que comprende la SEQ ID NO: 40. En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una secuencia de aminoácidos de la SEQ ID NO: 61. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 % o al menos un 99 % de identidad de secuencia con respecto a la secuencia de aminoácidos de la SEQ ID NO: 61.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 17; una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 18; una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 19; una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 20; una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 21 y/o una hélice de reconocimiento que comprende la secuencia de aminoácidos de la SEQ ID NO: 22.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento que comprende la SEQ ID NO: 29, una hélice de reconocimiento que comprende la SEQ ID NO: 30, una hélice de reconocimiento que comprende la SEQ ID NO: 31, una hélice de reconocimiento que comprende la SEQ ID NO: 32, una hélice de reconocimiento que comprende la SEQ ID NO: 33 y/o una hélice de reconocimiento que comprende la SEQ ID NO: 34.

En algunas modalidades, el o los dominios de unión a ADN son proteínas de dedos de cinc que comprenden una hélice de reconocimiento que comprende la SEQ ID NO: 41, una hélice de reconocimiento que comprende la SEQ ID NO: 42, una hélice de reconocimiento que comprende la SEQ ID NO: 43, una hélice de reconocimiento que comprende la SEQ ID NO: 44, una hélice de reconocimiento que comprende la SEQ ID NO: 45 y/o una hélice de reconocimiento que comprende la SEQ ID NO: 46.

En la presente, el o los dominios de unión a ADN son proteínas asociadas a CRISPR con inactividad catalítica (proteína Cas). En algunas modalidades, una proteína Cas con inactividad catalítica (o «proteína Cas muerta») es una proteína dCas9 o dCas12. En algunas modalidades, un ácido nucleico o rAAV comprende además al menos un ácido nucleico guía (p. ej., ARN guía o ARNg). En algunas modalidades, el ácido nucleico guía comprende una secuencia separadora que se dirige a *SCN1A*. En

algunas modalidades, el ácido nucleico guía comprende una secuencia separadora que tiene una secuencia de nucleótidos de cualquiera de las SEQ ID NO: 85, 86, 89, 90, 93 o 94. En algunas modalidades, el ácido nucleico guía comprende una secuencia de nucleótidos de cualquiera de las SEQ ID NO: 83-94. En algunas modalidades, el ácido nucleico guía es codificado por la secuencia de

5 ácido nucleico indicada en cualquiera de las SEQ ID NO: 83-94.

En algunas modalidades, al menos un dominio regulador de la transcripción es un dominio transactivador que comprende un dominio VP16, dominio VP64, dominio Rta, dominio p65, dominio Hsf1, o cualquier combinación de estos, tal como un dominio VPR (dominios VP64+p65+Rta1). En algunas modalidades, al menos un dominio regulador de la transcripción es codificado por una

10 secuencia de ácido nucleico tal como se indica en la SEQ ID NO: 47. En algunas modalidades, al menos un dominio de transactivación comprende la secuencia de aminoácidos indicada en la SEQ ID NO: 48.

En algunas modalidades, las ITR que flanquean el transgén comprenden una ITR que se selecciona dentro del grupo que consiste en: AAV1 ITR, AAV2 ITR, AAV3 ITR, AAV4 ITR, AAV5 ITR, AAV6 ITR,

15 AAV8 ITR, AAVrh8 ITR, AAV9 ITR, AAV10 ITR o AAVrh10 ITR. En algunas modalidades, la ITR es una  $\Delta$ TR o una mTR.

En algunas modalidades, un transgén de un ácido nucleico aislado se enlaza de manera operativa a un promotor. En algunas modalidades, un promotor es un promotor con especificidad de tejido. En algunas modalidades, un promotor con especificidad de tejido es un promotor neuronal, tal como SST,

20 NYP glutaminasa activada por fosfato (PAG, por sus siglas en inglés), transportador de glutamato vesicular-1 (VGLUT1), descarboxilasa del ácido glutámico 65 y 57 (GAD65, GAD67), Sinapsina I,  $\alpha$ -CamKII, Dock10, Prox1, Parvalbúmina (PV), Somatostatina (SST), Colecistocinina (CCK), Calretinina (CR) o Neuropeptido Y (NPY).

En algunas modalidades, un dominio de unión a ADN de un transgén se fusiona con un dominio regulador de la transcripción mediante un dominio enlazador. En algunas modalidades, un dominio enlazador es un enlazador flexible, por ejemplo, un enlazador rico en glicina o un enlazador de glicina-serina, o un enlazador escindible, tal como un enlazador fotoescindible o un enlazador escindible enzimático (p. ej., de proteasa).

25 En algunas modalidades, un dominio de unión a ADN de un transgén se fusiona con un dominio regulador de la transcripción mediante un dominio enlazador. En algunas modalidades, un dominio enlazador es un enlazador flexible, por ejemplo, un enlazador rico en glicina o un enlazador de glicina-serina, o un enlazador escindible, tal como un enlazador fotoescindible o un enlazador escindible enzimático (p. ej., de proteasa).

En algunas modalidades, un ácido nucleico aislado comprende un transgén que codifica múltiples dominios de unión a ADN, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 dominios de unión a ADN. En algunas modalidades, un ácido nucleico aislado comprende un transgén que codifica múltiples dominios reguladores de la transcripción, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 dominios reguladores de la transcripción.

30 En algunas modalidades, un ácido nucleico aislado comprende un transgén que codifica múltiples dominios de unión a ADN, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 dominios de unión a ADN. En algunas modalidades, un ácido nucleico aislado comprende un transgén que codifica múltiples dominios reguladores de la transcripción, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 dominios reguladores de la transcripción.

En algunas modalidades, un ácido nucleico aislado o un rAAV se expresa en una célula o un sujeto

que se caracteriza por la expresión anómala o haploinsuficiencia (p. ej., un aumento de la expresión o una disminución de la expresión) de un gen diana con respecto a una célula normal o un sujeto. En algunas modalidades, un ácido nucleico aislado o rAAV se expresa en una célula o un sujeto que se caracteriza por la expresión deficiente (p. ej., disminuida) de un gen diana con respecto a una célula o un sujeto normal. En algunas modalidades, un gen diana del ácido nucleico aislado o rAAV es *SCN1A*. En algunas modalidades, un serotipo de cápside de AAV se selecciona dentro del grupo que consiste en AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAV9, AAV10, AAVrh10 o AAV.PHPB.

En algunos aspectos, la descripción proporciona métodos de regulación de la expresión de un gen diana. En algunas modalidades, los métodos de la descripción comprenden administrar un ácido nucleico aislado o rAAV tal como se describe en la presente a una célula o un sujeto que expresa un gen diana, donde el sujeto presenta una haploinsuficiencia con respecto al gen diana (p. ej., haploinsuficiencia con respecto a *SCN1A*). Por ejemplo, en algunas modalidades, la expresión de un gen diana, tal como *SCN1A*, en la célula o el sujeto es deficiente (p. ej., menor) con respecto a la expresión del gen diana en una célula o un sujeto normal. En algunas modalidades, una célula a la que se le administra un ácido nucleico aislado o rAAV es una neurona. En algunas modalidades, una neurona es una neurona gabaérgica.

En algunas modalidades, la administración de un ácido nucleico aislado o rAAV provoca una expresión del gen diana (p. ej., una expresión de *SCN1A*) que aumenta al menos en 2 veces, al menos en 10 veces, al menos en 20 veces, al menos en 30 veces, al menos en 40 veces, al menos en 50 veces, al menos en 60 veces, al menos en 70 veces, al menos en 80 veces, al menos en 90 veces o al menos en 100 veces con respecto a un sujeto que no ha recibido la administración de un ácido nucleico aislado o rAAV. En algunas modalidades, la administración de un ácido nucleico aislado o rAAV provoca una expresión del gen diana (p. ej., una expresión de *SCN1A*) que aumenta al menos en 2 veces, al menos en 10 veces, al menos en 20 veces, al menos en 30 veces, al menos en 40 veces, al menos en 50 veces, al menos en 60 veces, al menos en 70 veces, al menos en 80 veces, al menos en 90 veces o al menos en 100 veces con respecto a la expresión del gen diana (p. ej., *SCN1A*) en el sujeto antes de la administración del ácido nucleico aislado o rAAV.

En algunos aspectos, la presente descripción proporciona un método de regulación de la expresión génica (p. ej., la expresión de *SCN1A*) en un sujeto, donde un ácido nucleico aislado o rAAV tal como se describe en la presente se administra a un sujeto que expresa un gen diana. En algunas modalidades, la expresión del gen diana en un sujeto es anómala (p. ej., mayor o menor) con respecto a la de un sujeto sano. En algunas modalidades, un sujeto tiene, o se presume que tiene, una haploinsuficiencia con respecto a la expresión de un gen diana con respecto a un sujeto sano.

En algunas modalidades, un sujeto tiene, o se presume que tiene, una enfermedad o afección provocada por la expresión haploinsuficiente de un gen diana. Por ejemplo, un sujeto con haploinsuficiencia con respecto a la expresión de *SCN1A*, en algunas modalidades, padece el síndrome de Dravet. En algunas modalidades, un ácido nucleico aislado o el rAAV se administra a un sujeto mediante inyección intravenosa, inyección intramuscular, inhalación, inyección subcutánea y/o inyección intracraneal.

En algunos aspectos, la descripción proporciona una composición que comprende el ácido nucleico aislado o el rAAV tal como se describe en la descripción. En algunas modalidades, una composición comprende un portador farmacéuticamente aceptable.

En algunos aspectos, la descripción proporciona un kit que comprende un recipiente que alberga un ácido nucleico aislado o el rAAV tal como se describe en la descripción. En algunas modalidades, un kit comprende un recipiente que alberga un portador farmacéuticamente aceptable. En algunas modalidades, un ácido nucleico aislado o rAAV y un portador farmacéuticamente aceptable se albergan en el mismo recipiente. En algunas modalidades, un recipiente es una jeringa.

En algunos aspectos, la descripción proporciona una célula hospedadora que comprende un ácido nucleico aislado o rAAV tal como se describe en la descripción. En algunas modalidades, una célula hospedadora es una célula eucariota. En algunas modalidades, una célula hospedadora es una célula de un mamífero. En algunas modalidades, una célula hospedadora es una célula humana, opcionalmente una neurona, por ejemplo, una neurona gabaérgica.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **FIG. 1** muestra datos de secuenciación por cromatografía que indican la conservación de la secuencia entre los genes *SCN1A* humanos (HEK) y de ratones (HEPG2) (Secuencia de consenso – SEQ ID NO: 98; Secuencia diana – SEQ ID NO: 99; Secuencia Hep-*SCN1A\_R4* (arriba) – SEQ ID NO: 100; Secuencia Hep-*SCN1A\_R4* (abajo) – SEQ ID NO: 101)

La **FIG. 2** muestra la alineación de secuencia de la región promotora proximal de los genes *SCN1A* de ser humano (SEQ ID NO: 1) y de ratón (SEQ ID NO: 2), donde se resalta una secuencia conservada. Dentro de esta secuencia conservada se encuentra una región diana de interés para la región de unión de la proteína de dedos de cinc (ZFP), que se encuentra en negrita (SEQ ID NO: 4).

La **FIG. 3** es un esquema que muestra la ubicación (SEQ ID NO: 3) de tres sitios de unión de ZFP diana en superposición (ZFP-1, ZFP-2, ZFP-3) (SEQ ID NO: 5-7) en la región promotora proximal del gen *SCN1A*.

Las **Figuras 4A-4D** muestran una alineación de seis secuencias de hélices de reconocimiento para los

dedos de cinc individuales (del Dedo 1 al Dedo 6; F1-F6) en la ZFP-1 que reconocerán regiones de tres bases individuales (los tripletes de ADN se indican en rojo separados por «•») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 2). La **FIG. 4A** resalta la secuencia de nucleótidos a la que se unirán los dedos de cinc 1 a 6 (F1-F6) de la ZFP-1 (SEQ ID NO: 3). La **FIG. 4B** muestra las tres secuencias de nucleótidos reconocidas por cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-1 (SEQ ID NO: 17-22). La **FIG. 4C** muestra las secuencias de aminoácidos de la ZFP-1, que contiene 6 dedos, uno en cada línea, donde los enlazadores entre los dedos están resaltadas para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP) (SEQ ID NO: 65-70). La **FIG. 4D** muestra la secuencia de nucleótidos de ZFP-1 (F1-F6) (SEQ ID NO: 102-107).

Las **Figuras 5A-5D** muestran una alineación de seis secuencias de hélices de reconocimiento para los dedos de cinc individuales (del Dedo 1 al Dedo 6; F1-F6) en la ZFP-2 que reconocerán regiones de tres bases individuales (los tripletes de ADN se indican en rojo separados por «\*») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 3). La **FIG. 5A** resalta la secuencia de nucleótidos (SEQ ID NO: 3) a la que se unirán los dedos de cinc 1 a 6 (F1-F6) de la ZFP-2. La **FIG. 5B** muestra los tres primeros nucleótidos reconocidos por cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-2 (SEQ ID NO: 29-34). La **FIG. 5C** muestra las secuencias de aminoácidos de la ZFP-2, que contiene 6 dedos, uno en cada línea (SEQ ID NO: 69-74), donde los enlazadores entre los dedos están resaltados para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP). La **FIG. 5D** muestra la secuencia de nucleótidos de ZFP-2 (F1-F6) (SEQ ID NO: 108-113).

Las **Figuras 6A-6D** muestran una alineación de seis secuencias de hélices de reconocimiento para los dedos de cinc individuales (del Dedo 1 al Dedo 6; F1-F6) en la ZFP-3 que reconocerán regiones de tres bases individuales (los tripletes de ADN se indican en rojo separados por «\*») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 4). La **FIG. 6A** resalta la secuencia de nucleótidos (SEQ ID NO: 3) a la que se unirán los dedos de cinc 1 a 6 (F1-F6) de la ZFP-3. La **FIG. 6B** muestra los tres primeros nucleótidos reconocidos por cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-3 (SEQ ID NO: 41-46). La **FIG. 6C** muestra las secuencias de aminoácidos de la ZFP-3, que contiene 6 dedos, uno en cada línea (SEQ ID NO: 75-80), donde los enlazadores entre los dedos están resaltados para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP). La **FIG. 6D** muestra la secuencia de nucleótidos de ZFP-3 (F1-F6) (SEQ ID NO: 114-119).

La **FIG. 7** muestra datos que indican que las ZFP que se unen a SCN1A que se describen en las **Figuras 4-6** aumentan la expresión del gen SCN1A en células HEK293T, tal como se mide mediante reacción en cadena de la polimerasa en tiempo real cualitativa (qRT-PCR, por sus siglas en inglés). Estas construcciones de expresión se suministraron a las células a través de la transfección transitoria

de plásmidos de expresión que codificaban los siguientes reguladores de la transcripción: *Streptococcus pyogenes* Cas9 + ARN guía de SCN1A (SpCas9 + Scn1a); Cas9 sin actividad de endonucleasa (dCas9); dominio de activación VPR + ARN guía de SCN1A (dCas9\_VPR + Scn1a); dominio de activación VPR + ZFP1 (VPR\_ZFP1); dominio de activación VPR + ZFP2 (VPR\_ZFP2); dominio de activación VPR + ZFP3 (VPR\_ZFP3); SpCas9 + ARN guía de ASCL1 (SpCas9 + Ascl1); tres VPR\_ZFP (VPR\_ZFP1 + VPR\_ZFP2 + VPR\_ZFP3). Los niveles de expresión se normalizaron con respecto a los niveles de expresión de TBP determinados mediante qRT-PCR en cada muestra.

La **FIG. 8** muestra datos que indican que las ZFP que se unen a SCN1A que se describen en las **Figuras 4-6** y Cas9+ARN guías de SCN1A aumentan la expresión del gen SCN1A en células HEK293T, tal como se mide mediante reacción en cadena de la polimerasa en tiempo real cualitativa (qRT-PCR).

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

Los aspectos de la descripción están vinculados con métodos y composiciones para modular (p. ej., aumentar) la expresión de un gen diana en una célula o un sujeto, donde hay haploinsuficiencia del gen diana (es decir, el gen diana comprende una copia funcional). En algunas modalidades, el gen diana es *SCN1A*.

En algunas modalidades, la descripción proporciona proteínas de fusión que comprenden un dominio de unión a ADN, tal como una ZFP, y un dominio regulador de la transcripción. En algunas modalidades, la descripción proporciona proteínas de fusión que comprenden un dominio de unión a ADN, tal como una ZFP, y un dominio transactivador (p. ej., un dominio VPR). En algunas modalidades, una proteína de unión a ADN se une a una secuencia de un gen diana o una región reguladora de un gen diana. En algunas modalidades, una región reguladora es una secuencia potenciadora, una secuencia promotora o una secuencia represora. En algunas modalidades, una secuencia promotora puede ser un promotor interno (p. ej., ubicado en un intrón de un gen diana) o un promotor externo (p. ej., en una ubicación retrógrada con respecto al sitio de inicio de la transcripción de un gen diana). En algunas modalidades, el dominio de unión a ADN de proteínas de fusión tal como se describe en la presente se une a una secuencia conservada en la región promotora de un gen diana (p. ej., *SCN1A*), a raíz de lo cual el dominio transactivador aumenta la expresión del gen.

En algunos aspectos, la descripción está vinculada con métodos para aumentar la expresión de un gen diana (p. ej., *SCN1A*) en una célula o un sujeto. En algunas modalidades, el gen diana contiene mutaciones que generan la haploinsuficiencia del gen diana en la célula o el sujeto. Por lo tanto, los métodos y las composiciones de la descripción pueden utilizarse, en algunas modalidades, para tratar

enfermedades y trastornos asociados a la haploinsuficiencia de un producto génico diana, por ejemplo, el síndrome de Dravet, que normalmente es el resultado de mutaciones en una copia del gen *SCN1A* que llevan a la haploinsuficiencia de la subunidad alfa de los canales de sodio dependientes de voltaje Nav1.1.

5

### *Proteínas de fusión de transactivación*

Algunos aspectos de la descripción están vinculados con proteínas de fusión que comprenden un dominio de unión a ADN (DBD, por sus siglas en inglés) y un dominio transactivador. Tal como se usa en la presente, una proteína de fusión comprende dos o más polipéptidos enlazados que son codificados por dos o más secuencias de aminoácidos separadas. Las proteínas quiméricas, tal como se usan en la presente, son proteínas de fusión donde los dos o más genes enlazados son de especies diferentes. Normalmente, las proteínas de fusión se producen de manera recombinante, donde los genes que codifican la proteína de fusión se encuentran en un sistema que respalda la expresión de los dos o más genes enlazados y la traducción de los ARNm resultantes para obtener proteínas recombinantes. En algunas modalidades, las proteínas de fusión se producen de manera recombinante en células procariotas o eucariotas. Las proteínas de fusión pueden configurarse en múltiples disposiciones. Por ejemplo, una proteína (Proteína A) tiene una ubicación retrógrada con respecto a una segunda proteína (Proteína B). En otras configuraciones de proteínas de fusión, la Proteína B tiene una ubicación retrógrada con respecto a la Proteína A. En algunas modalidades, una secuencia de ácido nucleico que codifica un dominio de unión a ADN tiene una ubicación retrógrada con respecto a una secuencia de ácido nucleico que codifica un dominio transactivador, y produce una proteína de fusión que comprende el DBD enlazado al transactivador. En algunas modalidades, una secuencia de ácido nucleico que codifica un dominio transactivador tiene una ubicación retrógrada con respecto a una secuencia de ácido nucleico que codifica un dominio de unión a ADN, y produce una proteína de fusión que comprende el dominio transactivador enlazado al dominio de unión a ADN. En algunas modalidades, una proteína de fusión comprende un dominio transactivador en una ubicación retrógrada con respecto a un dominio de unión a ADN. En algunas modalidades, una proteína de fusión comprende un dominio de unión a ADN en una ubicación retrógrada con respecto a un dominio transactivador.

En algunas modalidades, una proteína de fusión descrita en la descripción comprende un dominio de unión a ADN. Tal como se usa en la presente, un «dominio de unión a ADN (DBD)» hace referencia a una proteína plegada de manera independiente que comprende al menos un motivo estructural que reconoce ADN mono o bicatenario (ADNmc o ADNbc). Determinados DBD reconocen secuencias

específicas (motivo o secuencia de reconocimiento), mientras que otros tipos de DBD tienen una afinidad general con respecto al ADN. En algunas modalidades, una proteína de fusión descrita en la descripción comprende un DBD con especificidad de secuencia. En algunas modalidades, el DBD reconoce (p. ej., se une con especificidad a) una secuencia de ácido nucleico ubicada dentro o cerca del gen que codifica una proteína SCN1A (p. ej., Nav1.1). Normalmente, las proteínas que contienen DBD están involucradas en procesos celulares tales como la transcripción, la replicación, la reparación y el almacenamiento de ADN. Los DBD en los factores de transcripción reconocen secuencias de ADN específicas en la región promotora o en elementos potenciadores para promover la expresión génica. Se utilizan DBD de factores de transcripción como proteínas de fusión en la ingeniería genética para regular la expresión de los genes diana y pueden someterse a mutaciones para alterar la especificidad de unión del ADN o la afinidad de unión del ADN y, de esa manera, regular la expresión de un gen diana deseado. Los ejemplos de DBD incluyen, de modo no taxativo, un motivo hélice-giro-hélice, motivos de dedos de cinc (incluidos los dedos de cinc Cys2-His2), efectores de tipo activador de la transcripción (TALE), motivos de hélices aladas, cajas de HMG, proteínas dCas (p. ej., dCas9 o dCas12a), homeodominios y dominios de pliegue de OB.

En algunas modalidades, la descripción está vinculada con proteínas de fusión DBD de dedos de cinc. Tal como se usa en la presente, una «proteína de dedos de cinc (ZFP)» hace referencia a una proteína que contiene al menos un motivo estructural caracterizado por la coordinación de uno o más iones de cinc que estabilizan el pliegue de la proteína. Los dedos de cinc se encuentran entre los motivos estructurales más diversos que se encuentran en las proteínas, y hasta un 3 % de los genes humanos codifican dedos de cinc. La mayoría de las ZFP contiene múltiples dedos de cinc que establecen contactos en tándem con moléculas diana, que incluyen ADN, ARN y la proteína pequeña ubiquitina. Los motivos de dedos de cinc «clásicos» están compuestos por 2 aminoácidos de cisteína y 2 aminoácidos de histidina (C<sub>2</sub>H<sub>2</sub>) y se unen al ADN con especificidad de secuencia. Estas ZFP, que incluyen el factor de transcripción IIIIA (TFIIIA), están involucradas normalmente en la expresión génica. Múltiples motivos de dedos de cinc en las proteínas de unión a ADN se unen a y establecen una envoltura alrededor del exterior de una doble hélice de ADN. Debido a su tamaño relativamente pequeño (p. ej., cada dedo tiene alrededor de 25-40, habitualmente 27-35, aminoácidos), las proteínas de fusión de dedos de cinc se utilizan para crear DBD con una especificidad de unión de ADN novedosa. Estos DBD pueden suministrar otros dominios fusionados (p. ej., dominios de represión o activación de la transcripción o dominios de modificación de carácter epigenético) para alterar la regulación de transcripción de un gen diana. En algunas modalidades, las proteínas de dedos de cinc comprenden de 2 a 8 dedos, donde cada dedo contiene de 27 a 40 aminoácidos (p. ej., 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 aminoácidos).

En algunas modalidades, una ZFP comprende 1, 2, 3, 4, 5, 6, 7 u 8 dedos de cinc. Cada dedo de cinc puede comprender 25-40, 25-30, 30-35, 35-40 o 40-45 aminoácidos. En algunas modalidades, un dedo de cinc comprende 27-35 aminoácidos. En algunas modalidades, un dedo de cinc comprende 27, 28, 29, 30, 31, 32, 33, 34 o 35 aminoácidos. Un dedo de cinc puede reconocer o unirse con especificidad a una secuencia diana, p. ej., un gen diana o una región reguladora de un gen diana con haploinsuficiencia en un sujeto. En algunas modalidades, un dedo de cinc se une a una secuencia diana de un gen *SCN1A*, p. ej., un *SCN1A* humano, por ejemplo, tal como se indica en la SEQ ID NO: 49. En algunas modalidades, un dedo de cinc que se une a una secuencia diana de un gen *SCN1A* comprende una o más secuencias de aminoácidos de la SEQ ID NO: 63-80, o una combinación de estas. En algunas modalidades, un dedo de cinc reconoce o se une con especificidad a una secuencia diana que comprende una secuencia trinucleotídica.

En algunas modalidades, un dedo de cinc comprende una hélice de reconocimiento que reconoce o se une a una secuencia diana, p. ej., una secuencia diana que comprende una secuencia trinucleotídica. En algunas modalidades, una hélice de reconocimiento se une a un trinucleótido. En algunas modalidades, una hélice de reconocimiento comprende 4-10 aminoácidos. En algunas modalidades, una hélice de reconocimiento comprende 4, 6, 7, 8, 9 o 10 aminoácidos. En algunas modalidades, una hélice de reconocimiento se une a una secuencia trinucleotídica de un gen *SCN1A*. En algunas modalidades, una secuencia de reconocimiento que se une a un gen *SCN1A* comprende una secuencia de aminoácidos de cualquiera de las SEQ ID NO: 17-22, 29-34 o 41-46. En algunas modalidades, una secuencia de reconocimiento que se une a un gen *SCN1A* es codificada por cualquiera de las SEQ ID NO: 11-16, 23-28 o 35-40. En algunas modalidades, un dedo de cinc se une a la misma secuencia de nucleótidos que una hélice de reconocimiento que comprende una secuencia de aminoácidos de cualquiera de las SEQ ID NO: 17-22, 29-34 o 41-46.

En algunas modalidades, un dedo de cinc comprende una secuencia enlazadora en su extremo C que puede servir para enlazar o conectar dicho dedo de cinc con un dedo de cinc adicional. En algunas modalidades, una secuencia enlazadora puede ser un enlazador canónico, p. ej., que comprenda una secuencia de aminoácidos de TGEKP (SEQ ID NO: 120). En algunas modalidades, una secuencia enlazadora puede ser un enlazador no canónico, p. ej., que comprenda una secuencia de aminoácidos de TGSQKP (SEQ ID NO: 121). En algunas modalidades, una secuencia enlazadora puede tener 2-10 aminoácidos, p. ej., 2, 3, 4, 5, 6, 7, 8, 9 o 10 aminoácidos.

En algunas modalidades, una ZFP que se une a un gen diana, p. ej., un gen *SCN1A*, comprende seis dedos de cinc, cada uno de los cuales reconoce o se une a una secuencia trinucleotídica diferente del gen diana, p. ej., un gen *SCN1A*. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende una secuencia de aminoácidos de la SEQ ID NO: 57. En algunas modalidades, una ZFP

que se une a un gen *SCN1A* comprende dedos de cinc que comprenden secuencias de aminoácidos de las SEQ ID NO: 63, 64, 65, 66, 67 y/o 68. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende hélices de reconocimiento que comprenden secuencias de aminoácidos de las SEQ ID NO: 17, 18, 19, 20, 21 y/o 22. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende una secuencia de aminoácidos de la SEQ ID NO: 59. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende dedos de cinc que comprenden secuencias de aminoácidos de las SEQ ID NO: 69, 70, 71, 72, 73 y/o 74. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende hélices de reconocimiento que comprenden secuencias de aminoácidos de las SEQ ID NO: 29, 30, 31, 32, 33 y/o 34. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende una secuencia de aminoácidos de la SEQ ID NO: 61. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende dedos de cinc que comprenden secuencias de aminoácidos de las SEQ ID NO: 75, 76, 77, 78, 79 y/o 80. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende hélices de reconocimiento que comprenden secuencias de aminoácidos de las SEQ ID NO: 41, 42, 43, 44, 45 y/o 46. En algunas modalidades, una ZFP que se une a un gen *SCN1A* comprende al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 % o al menos un 99 % de identidad de secuencia con respecto a la SEQ ID NO: 57, 59 o 61, tal como se mostrará a continuación.

SEQ ID NO: 57 (secuencia de aminoácidos de la proteína ZFP1)

RPFQCRICMRNFSQRGNLVRHIRTHTGEKPFACDICGKKFALSFNLTRHTKIHTGSQKPFQCRICMRNFSRSDNLTRHIRTHTGEKPFACDICGKKFADRSHLARHTKIHTGSQKPFQCRICMRNFSQKAHLTAHIRTHTGEKPFACDICGRKFARSDNLTRHTKIHLRQKD

SEQ ID NO: 59 (secuencia de aminoácidos de la proteína ZFP2)

RPFQCRICMRNFSRSSNLTRHIRTHTGEKPFACDICGKKFADKRTLIRHTKIHTGSQKPFQCRICMRNFSQRGNLVRHIRTHTGEKPFACDICGKKFALSFNLTRHTKIHTGSQKPFQCRICMRNFSRSDNLTRHIRTHTGEKPFACDICGRKFADRSHLARHTKIHLRQKD

SEQ ID NO: 61 (secuencia de aminoácidos de la proteína ZFP3)

RPFQCRICMRNFSDRSALARHIRTHTGEKPFACDICGKKFARSDNLTRHTKIHTGSQKPFQCRICMRNFSQSGDLTRHIRTHTGEKPFACDICGKKFAVRQTLKQHTKIHTGSQKPFQCRICMRNFSAAGNLTRHIRTHTGEKPFACDICGRKFARSDNLTRHTKIHLRQKD

En algunas modalidades, los DBD son proteínas efectoras de tipo activador de la transcripción (TALE). Una TALE puede reconocer o unirse con especificidad a una secuencia diana, p. ej., un gen diana o una región reguladora de un gen diana. En algunas modalidades, un sujeto presenta haploinsuficiencia del gen diana. En algunas modalidades, una TALE se une a una secuencia diana de un gen *SCN1A*,

p. ej., un *SCN1A* humano, tal como se proporciona en la SEQ ID NO: 49. Las proteínas TALE son secretadas por bacterias y se unen a secuencias promotoras en una planta hospedadora para activar la expresión de genes de plantas que ayudan en la infección bacteriana. Normalmente, las proteínas TALE se manipulan para que se unan a nuevas secuencias de ADN, dado que reconocen secuencias diana a través de un dominio de repetición central que consiste en una cantidad variable de ~30-35 repeticiones de aminoácidos, donde cada repetición reconoce un solo par de bases dentro de la secuencia diana. Normalmente, se necesita un conjunto de estas repeticiones para reconocer una secuencia de ADN.

En algunas modalidades, los DBD son homeodominios. Un homeodominio puede reconocer o unirse con especificidad a una secuencia diana, p. ej., un gen diana o una región reguladora de un gen diana. En algunas modalidades, un sujeto presenta haploinsuficiencia del gen diana. En algunas modalidades, un homeodominio se une a una secuencia diana de un gen *SCN1A*, p. ej., un *SCN1A* humano, tal como se proporciona en la SEQ ID NO: 49. Los homeodominios son proteínas que contienen tres hélices alfa y un brazo de extremo N que son responsables del reconocimiento de una secuencia diana. Normalmente, un homeodominio reconoce una secuencia de ADN de tamaño pequeño (~4 a 8 pares de bases); sin embargo, estos dominios pueden fusionarse en tándem con otros dominios de unión a ADN (otros homeodominios o proteínas de dedos de cinc) para reconocer secuencias de mayor extensión (de 12 a 24 pares de bases). Por lo tanto, los homeodominios pueden ser componentes de DBD que reconozcan secuencias particulares dentro del genoma humano.

En la presente, el o los dominios de unión a ADN son proteínas asociadas a CRISPR con inactividad catalítica (proteína Cas). Una proteína Cas con inactividad catalítica (también conocida como dCas o «proteína Cas muerta») es una proteína Cas que se ha modificado o mutado de forma tal que tiene una disminución en la actividad de nucleasa (p. ej., actividad de endonucleasa) o carece totalmente de actividad de nucleasa (p. ej., actividad de endonucleasa). En algunas modalidades, una proteína Cas con inactividad catalítica es una dCas9 o dCas12. En algunas modalidades, los DBD son proteínas dCas (que también se conocen como «Cas muertas») tales como dCas9 o dCas12a. Las proteínas dCas son variantes mutantes de proteínas asociadas a CRISPR (Cas, p. ej., Cas9 o Cas12a) que se han sometido a mutaciones de forma tal que presenten inactividad catalítica, es decir, la incapacidad de escindir nucleótidos. Una dCas puede reconocer o unirse con especificidad a una secuencia diana, p. ej., un gen diana o una región reguladora de un gen diana. Un complejo que comprende una proteína dCas y un ácido nucleico guía (p. ej., ARNg) puede dirigirse y/o unirse a genes o secuencias de nucleótidos de carácter específico que sirvan como complemento del ácido nucleico guía. En algunas modalidades, un sujeto presenta haploinsuficiencia del gen diana. En algunas modalidades, una dCas se une a una secuencia diana de un gen *SCN1A*, p. ej., un *SCN1A* humano, tal como se

proporciona en la SEQ ID NO: 49. Sin embargo, las proteínas dCas conservan su capacidad de reconocer y unirse a secuencias de ADN diana cuando se unen a un ácido nucleico guía (p. ej., un ARN guía, ARNg o ARNGs) complementario o parcialmente complementario con respecto a dicha secuencia de ADN diana. En algunas modalidades, un ácido nucleico guía para dirigir dCas (p. ej., dCas9) a SCN1A comprende una secuencia separadora con cualquiera de las SEQ ID NO: 85, 86, 89, 90, 93 o 94. En algunas modalidades, un ácido nucleico guía para dirigir dCas (p. ej., dCas9) a SCN1A comprende una secuencia separadora con al menos 15 (p. ej., al menos 16, 17, 18, 19 o 20) nucleótidos consecutivos de cualquiera de las SEQ ID NO: 85, 86, 89, 90, 93 o 94. En algunas modalidades, un ácido nucleico guía para dirigir dCas (p. ej., dCas9) a SCN1A comprende cualquiera de las SEQ ID NO: 83, 84, 87, 88, 91 o 92. En algunas modalidades, un ácido nucleico guía para dirigir dCas (p. ej., dCas9) a SCN1A comprende o consiste en cualquiera de las SEQ ID NO: 83-94. Por lo tanto, las endonucleasas dCas pueden ser componentes de DBD que reconozcan secuencias particulares dentro del genoma humano. En algunas modalidades, una proteína de fusión comprende una proteína dCas9 y un dominio de transactivación (p. ej., un dominio VPR).

La descripción está vinculada, en algunos aspectos, a dominios de unión a ADN que se unen a un gen que codifica un canal de sodio dependiente de voltaje (p. ej.,  $Na_v1.1$ ). En algunas modalidades, un gen que codifica un canal de sodio dependiente de voltaje es un gen *SCN1A* y comprende la secuencia indicada en la SEQ ID NO: 49. En algunas modalidades, un dominio de unión a ADN se une a una región no traducida de un gen diana, tal como una región no traducida 3' (3'UTR, por sus siglas en inglés) o una región no traducida 5' (5'UTR, por sus siglas en inglés). En algunas modalidades, una región no traducida comprende una secuencia reguladora, por ejemplo, una secuencia potenciadora, promotora, intrónica o represora. En algunas modalidades, un dominio de unión a ADN es una proteína de dedos de cinc que comprende las secuencias indicadas en las SEQ ID NO: 57-62. En algunas modalidades, un dominio de unión a ADN se une a una secuencia de ácido nucleico indicada en cualquiera de las SEQ ID NO: 5-7.

La cantidad de dominios de unión a ADN codificados por un transgén puede variar. En algunas modalidades, un transgén codifica un dominio de unión a ADN. En algunas modalidades, un transgén codifica 2 dominios de unión a ADN. En algunas modalidades, un transgén codifica 3 dominios de unión a ADN. En algunas modalidades, un transgén codifica 4 dominios de unión a ADN. En algunas modalidades, un transgén codifica 5 dominios de unión a ADN. En algunas modalidades, un transgén codifica 6 dominios de unión a ADN. En algunas modalidades, un transgén codifica 7 dominios de unión a ADN. En algunas modalidades, un transgén codifica 8 dominios de unión a ADN. En algunas modalidades, un transgén codifica 9 dominios de unión a ADN. En algunas modalidades, un transgén codifica 10 dominios de unión a ADN. En algunas modalidades, un transgén codifica más de 10 (p. ej.,

20, 30, 50, 100, etc.) dominios de unión a ADN. Los dominios de unión a ADN pueden ser el mismo dominio de unión a ADN (p. ej., múltiples copias del mismo DBD), diferentes dominios de unión a ADN (p. ej., cada DBD se une a una secuencia particular), o una combinación de estos.

La descripción está vinculada, en algunos aspectos, a proteínas de fusión que comprenden un dominio transactivador. Tal como se usa en la presente, un «dominio de transactivación» hace referencia a un dominio de supercóntigo en un factor de transcripción que contiene sitios de unión para otras proteínas que regulan la expresión génica, tales como correguladores de la transcripción. En algunas modalidades, un dominio de transactivación (también conocido como «dominio de activación de la transcripción») actúa junto con un DBD para activar la transcripción desde un promotor o potenciador, directamente a través del contacto con factores de transcripción, o indirectamente a través de proteínas coactivadoras. A los dominios de transactivación (TAD, por sus siglas en inglés) comúnmente se los designa por sus composiciones de aminoácidos, donde los aminoácidos son esenciales para la actividad o son los más abundantes en los TAD. Se utilizan TAD como proteínas de fusión en la ingeniería genética para regular la expresión de los genes diana y pueden someterse a mutaciones para alterar el nivel de activación de la transcripción y, por lo tanto, la expresión del gen diana. Los ejemplos de dominios de transactivación incluyen, de modo no taxativo, GAL4, HAP1, VP16, P65, RTA y GCN4.

En algunas modalidades, un dominio transactivador comprende un dominio VP64. El VP64 es un TAD ácido compuesto por cuatro copias en tándem de proteína VP16, que el virus del herpes simple expresa de manera natural. Cuando se fusiona con un DBD que se une en el promotor de un gen o cerca de este, el VP64 actúa como un activador potente de la transcripción y, por lo tanto, puede utilizarse para regular la expresión de un gen diana (p. ej., *SCN1A*). Normalmente, el dominio VP64 consiste en una repetición tetramérica del dominio de activación de carácter mínimo de la proteína del herpes simple VP16. En algunas modalidades, el dominio VP64 comprende cuatro repeticiones de los residuos de aminoácidos 437-448 en VP16. En algunas modalidades, una proteína VP16 es codificada por el gen del virus del herpes simple 2 UL48, que comprende la secuencia indicada en el n.º de registro de sec. de ref. del NCBI: NC\_001798.2. En algunas modalidades, un gen VP16 comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en el n.º de registro de sec. de ref. del NCBI: YP\_009137200.1. En algunas modalidades, una proteína VP16 comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en el n.º de registro de sec. de ref. del NCBI: Q69113-1. En algunas modalidades, un gen VP16 comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica,

90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en la SEQ ID NO: 51. En algunas modalidades, una proteína VP16 comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en la SEQ ID NO: 52.

En algunas modalidades, un dominio transactivador comprende un dominio de activación de P65. P65 es una subunidad del factor de transcripción NF- $\kappa$ B que contiene dos TAD ácidos adyacentes en su extremo C. Cuando se fusiona con un DBD que se une al promotor de un gen o cerca de este, la proteína p65 actúa como un activador fuerte de la transcripción y, por lo tanto, puede utilizarse para regular la expresión de un gen diana, por ejemplo, tal como se describe en Urlinger, et ál. «The p65 domain from NF-kappaB is an efficient human activator in the tetracycline-regulatable gene expression system», *Gene*, **2000**. En algunas modalidades, una proteína p65 es codificada por un gen *RELA* humano, que comprende la secuencia indicada en el n.º de registro de sec. de ref. del NCBI: NM\_001145138.1, NM\_001243984.1, NM\_001243985.1 o NM\_021975.3. En algunas modalidades, un gen *RELA* comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en cualquiera de los n.º de id. de sec. de ref. del NCBI: NM\_001145138.1, NM\_001243984.1, NM\_001243985.1 o NM\_021975.3. En algunas modalidades, una proteína p65 comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en NP\_001138610.1, NP\_001230913.1, NP\_001230914.1 y NP\_068110.3. En algunas modalidades, un gen *RELA* comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en la SEQ ID NO: 53. En algunas modalidades, una proteína p65 comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en la SEQ ID NO: 54.

En algunas modalidades, un dominio transactivador comprende un dominio RTA. RTA es un TAD hidrófobo derivado del virus de Epstein-Barr que es un dominio de transactivación potente que se une a la región potenciadora para promover la expresión de varios genes virales. Cuando se fusiona con un DBD que se une en el promotor de un gen o cerca de este, la proteína RTA actúa como un activador fuerte de la transcripción y, por lo tanto, puede utilizarse para regular la expresión de un gen diana, por ejemplo, tal como se describe en Miyazawa, et ál., «IL-10 promoter transactivation by the viral K-RTA protein involves the host-cell transcription factors, specificity proteins 1 and 3», *Journal of*

*Biological Chemistry*, **2018**. En algunas modalidades, una proteína RTA es codificada por un gen del virus de Epstein-Barr BRLF1, que comprende la secuencia indicada en el n.º de registro de sec. de ref. del NCBI: YP\_041674.1. En algunas modalidades, un gen BRLF1 comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en cualquiera de los n.º de id. de sec. de ref. del NCBI: YP\_041674.1. En algunas modalidades, una proteína RTA comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en YP\_041674.1. En algunas modalidades, un gen BRLF1 comprende una secuencia de nucleótidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en la SEQ ID NO: 55. En algunas modalidades, una proteína RTA comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos indicada en la SEQ ID NO: 56.

La descripción se basa, en parte, en proteínas de fusión que comprenden un dominio transactivador híbrido. Un «dominio transactivador híbrido», tal como se usa en la presente, hace referencia a una proteína de fusión que comprende más de una proteína activadora de la transcripción o partes de esta (p. ej., 2, 3, 4, 5 o más proteínas activadoras de la transcripción, o partes de estas). Se utilizan dominios de transactivación híbridos en la ingeniería genética para aumentar la expresión de genes diana. En algunas modalidades de la descripción, se utiliza un dominio de transactivación híbrido tripartito que comprende la secuencia de nucleótidos para VP64-P65-RTA (VPR), tal como se describe en Chavez, et ál. «Highly efficient Cas9-mediated transcriptional programming», *Nat Methods*, **2015**, (SEQ ID NO: 47) para aumentar la expresión del gen diana (p. ej., *SCN1A*).

En algunas modalidades, las proteínas de fusión descritas en la presente pueden comprender un DBD (p. ej., una ZFP) y una proteína represora de la transcripción. En algunos aspectos, la descripción está vinculada con proteínas de fusión que comprenden un dominio represor de la transcripción. Una proteína «represora de la transcripción», tal como se usa en la presente, hace referencia en términos generales a un polipéptido que regula a la baja la expresión de un gen diana. Los ejemplos de represores de la transcripción incluyen, de modo no taxativo, KRAB, SMRT/TRAC-2 y NCoR/RIP-13.

En algunas modalidades, esas proteínas de fusión represoras de la transcripción son útiles para reducir el nivel de expresión de un gen diana (p. ej., un gen que se expresa por demás en una enfermedad de ganancia de función).

Una secuencia de ácido nucleico aislado hace referencia a una secuencia de ADN o ARN. En algunas modalidades, las proteínas y los ácidos nucleicos de la descripción están aislados. Tal como se usa en la presente, el término «aislado/a» significa producido/a artificialmente. Tal como se usa en la presente con respecto a los ácidos nucleicos, el término «aislado» significa: (i) amplificado in vitro, por ejemplo, mediante una reacción en cadena de polimerasa (PCR, por sus siglas en inglés); (ii) producido de manera recombinante mediante clonación; (iii) purificado, por ejemplo, mediante escisión y separación en gel; o (iv) sintetizado, por ejemplo, mediante síntesis química. Un ácido nucleico aislado es aquel que puede manipularse fácilmente mediante procedimientos de ADN recombinante conocidos en la técnica. Por lo tanto, una secuencia de nucleótidos contenida en un vector donde se conocen sitios de restricción 5' y 3' o donde se han descrito secuencias cebadoras de la reacción en cadena de la polimerasa (PCR) se considera aislada, pero no así una secuencia de ácido nucleico que se encuentra en su estado natural en su hospedador natural. Un ácido nucleico aislado puede estar sustancialmente purificado, aunque no necesariamente. Por ejemplo, un ácido nucleico aislado dentro de un vector de clonación o expresión no es puro en cuanto a que puede comprender solamente un porcentaje ínfimo del material en la célula en la que se encuentra. Sin embargo, un ácido nucleico de ese tipo está aislado, tal como se usa el término en la presente, porque puede manipularse fácilmente mediante procedimientos convencionales conocidos por los expertos en la técnica. Tal como se usa en la presente con respecto a proteínas o péptidos, el término «aislado/a» hace referencia a una proteína o un péptido que se ha aislado de su entorno natural o que se produce artificialmente (p. ej., mediante síntesis química, mediante tecnología de ADN recombinante, etc.).

En algunos aspectos, la descripción está vinculada con ácidos nucleicos aislados (p. ej., construcciones de expresión, tales como vectores de rAAV) configurados para expresar una o más proteínas de fusión de dominios de transactivación-ZFP. En algunas modalidades, una proteína de fusión comprende entre 1 y 10 (p. ej., 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10) DBD y/o entre 1 y 10 (p. ej., 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10) dominios transactivadores. En algunas modalidades, una proteína de fusión comprende más de 10 DBS y/o más de 10 dominios transactivadores.

En algunos aspectos de la descripción, un dominio de unión a ADN se fusiona con un dominio regulador de la transcripción indirectamente a través de un enlazador. Tal como se usa en la presente, un «enlazador» es en términos generales una extensión de polipéptidos que une estructuralmente dos polipéptidos distintos dentro de un solo transgén. En algunas modalidades, un enlazador es flexible para permitir el movimiento de los distintos polipéptidos. En algunas modalidades, un enlazador flexible comprende residuos de glicina. En algunas modalidades, un enlazador flexible comprende una mezcla de residuos de glicina y serina. En algunas modalidades, un enlazador es escindible, lo cual permite la

separación de los polipéptidos. En algunas modalidades, un enlazador escindible es cortado por una proteasa. En algunas modalidades, la proteasa es tripsina o el Factor X.

En algunas modalidades, un enlazador comprende entre 5 y 30 aminoácidos (p. ej., 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 aminoácidos). En algunas modalidades, un enlazador comprende entre 3 y 30 aminoácidos (p. ej., 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 aminoácidos). En algunas modalidades, un enlazador comprende entre 3 y 20 aminoácidos (p. ej., 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 aminoácidos).

La descripción se basa, en parte, en proteínas de fusión que se modifican para aumentar la expresión de un gen que codifica una proteína de subunidad de canales de iones de sodio dependientes de voltaje (a la que también se hace referencia como «proteína SCN»), por ejemplo, *SCN1A*. Tal como se usa en la presente, «una proteína SCN» hace referencia a una proteína de canales de iones de sodio que media la permeabilidad a los iones de sodio dependientes de voltaje de las membranas excitables, lo cual permite que los iones de sodio pasen a través de la membrana. Los ejemplos de proteínas SCN en seres humanos incluyen, de modo no taxativo, *SCN1A*, *SCN3A*, *SCN5A*, *SCN10A* y *SCN11A*. En algunas modalidades, una proteína SCN es *SCN1A* (que también se denomina «Nav1.1»), que codifica una subunidad de canales de iones  $\alpha_1$  de Tipo 1. En algunas modalidades, una proteína SCN es una proteína *SCN1B* que codifica una subunidad de canales de iones  $\beta_1$  de Tipo 1 o una proteína *SCN1C*. En algunas modalidades, una proteína SCN es una combinación de proteínas *SCN1A*, *SCN1B* y/o *SCN1C*. Tal como se describe en la presente, una proteína SCN puede ser una parte o un fragmento de una proteína SCN. En algunas modalidades, una proteína SCN tal como se describe en la presente es una variante de una proteína SCN, tal como un mutante puntual o un mutante truncado. En los seres humanos, la *SCN1A* es codificada por el gen *SCN1A* (N.º de id. del gen: 6323, humano), que se conserva en los chimpancés, los macacos de la India, los perros, las vacas, los ratones, las ratas y los pollos. El gen *SCN1A* en el ser humano se expresa principalmente en el cerebro, los pulmones y los testículos. En algunas modalidades, las proteínas *SCN1A* comprenden cinco repeticiones estructurales (I, II, III, IV, Q).

En algunas modalidades, una proteína *SCN1A* es codificada por un gen *SCN1A* humano, que comprende la secuencia indicada en el n.º de id. de sec. de ref. del NCBI: NM\_001165963.2, NM\_00165964.2, NM\_001202435.2, NM\_001353948.1, NM\_001353949.1, NM\_001353950.1, NM\_00135395.1, NM\_001353952.1, NM\_001353954.1, NM\_00353955.1, NM\_001353957.1, NM\_001353958.1, NM\_001353960.1, NM\_001353961.1 o NM\_006920.5. En algunas modalidades, una proteína *SCN1A* es codificada por un gen *SCN1A* de ratón, que comprende la secuencia indicada en el n.º de id. de sec. de ref. del NCBI: NM\_001313997.1 o NM\_018733.2. En algunas modalidades,

- una proteína SCN1A comprende una secuencia de nucleótidos aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en el n.º de id. de sec. de ref. del NCBI: NG\_011906.1, NM\_001313997.1 o NM\_018733.2. En algunas modalidades, un gen SCN1A
- 5 comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia indicada en la SEQ ID NO: 50. En algunas modalidades, una proteína SCN1A humana comprende la secuencia indicada en el n.º de id. de sec. de ref. del NCBI: NP\_001159435.1, NP\_0011159436.1, NP\_001189364.1, NP\_001340877.1, NP\_001340878.1, NP\_001340879.1, NP\_001340880.1, NP\_001340881.1, NP\_001340883.1,
- 10 NP\_001340884.1, NP\_001340886.1, NP\_001340887.1, NP\_001340889.1, NP\_001340890.1, NP\_00851.3. En algunas modalidades, una proteína SCN1A comprende una secuencia de nucleótidos aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de aminoácidos codificada por la secuencia de ácido nucleico indicada en el n.º de id. de sec. de ref. del NCBI: NG\_011906.1, NM\_001313997.1 o NM\_018733.2. En algunas
- 15 modalidades, una proteína SCN1A de ratón comprende la secuencia indicada en el n.º de id. de sec. de ref. del NCBI: NP\_001300926.1 o NP\_061203.2. En algunas modalidades, una proteína SCN1A humana comprende una secuencia de aminoácidos 99 % idéntica, 95 % idéntica, 90 % idéntica, 80 % idéntica, 70 % idéntica, 60 % idéntica o 50 % idéntica a la secuencia de ácido nucleico indicada en la SEQ ID NO: 49.
- 20 Los ácidos nucleicos aislados de la descripción pueden ser vectores de virus adenoasociados (AAV) recombinantes (vectores rAAV). En algunas modalidades, un ácido nucleico aislado tal como se describe en la descripción comprende una región (p. ej., una primera región) que comprende una primera repetición de terminal invertida (ITR) de virus adenoasociado (AAV), o una variante de esta. El ácido nucleico aislado (p. ej., el vector de AAV recombinante) puede incluirse en una proteína de
- 25 cápside y administrarse a un sujeto y/o suministrarse a una célula diana seleccionada. Normalmente, los «vectores de AAV recombinantes (rAAV)» están compuestos, como mínimo, por un transgén y sus secuencias reguladoras, y repeticiones terminales invertidas (ITR) de AAV 5' y 3'. El transgén puede comprender una región que codifique, por ejemplo, una proteína y/o una secuencia de control de la expresión (p. ej., una cola de poli-A), tal como se describe en otras partes de la descripción.
- 30 Por lo general, las secuencias de ITR tienen alrededor de 145 pb de longitud. Preferentemente, se usa prácticamente la totalidad de las secuencias que codifican las ITR en la molécula, aunque se puede permitir cierto grado de modificaciones leves de estas secuencias. La capacidad de modificar estas secuencias de ITR se encuentra dentro de los conocimientos de la técnica. (Véanse, p. ej., textos tales como Sambrook et al, «Molecular Cloning. A Laboratory Manual», 2.<sup>a</sup> ed., Cold Spring Harbor

Laboratory, Nueva York (1989); y K. Fisher et ál., J Virol., 70:520 532 (1996)). Un ejemplo de una molécula de ese tipo empleada en la presente invención es un plásmido «de acción en cis» que contiene el transgén, en el que la secuencia de transgén seleccionada y los elementos reguladores asociados se encuentran flanqueados por las secuencias de ITR de AAV 5' y 3'. Las secuencias de ITR de AAV pueden obtenerse a partir de cualquier AAV conocido, lo cual incluye los tipos de AAV de mamíferos identificados en la presente. En algunas modalidades, el ácido nucleico aislado comprende además una región (p. ej., una segunda región, una tercera región, una cuarta región, etc.) que comprende una segunda ITR de AAV.

Además de los elementos principales identificados anteriormente para el vector de AAV recombinante, el vector también incluye elementos de control convencionales, que se enlazan de manera operativa con elementos del transgén, de forma tal que se permite su transcripción, traducción y/o expresión en una célula transfectada con el vector o infectada con el virus producido mediante la descripción. Tal como se usa en la presente, las secuencias «enlazadas de manera operativa» incluyen las secuencias de control de la expresión que son contiguas con el gen de interés y las secuencias de control de la expresión que actúan en trans o a cierta distancia para controlar el gen de interés. Las secuencias de control de la expresión incluyen secuencias adecuadas de inicio y terminación de la transcripción, promotoras y potenciadoras; señales de procesamiento de ARN eficientes tales como señales de corte y empalme y de poliadenilación (poliA); secuencias que estabilizan el ARNm citoplasmático; secuencias que potencian la eficiencia de la traducción (p. ej., secuencias de consenso de Kozak); secuencias que potencian la estabilidad de la proteína; y, cuando así se desea, secuencias que potencian la secreción del producto codificado. En la técnica se conoce una cantidad de secuencias de control de la expresión que pueden utilizarse, que incluyen promotores naturales, constitutivos, inducibles y/o con especificidad de tejido.

Tal como se usa en la presente, se dice que una secuencia de ácido nucleico (p. ej., una secuencia de codificación) y las secuencias reguladoras están enlazadas de manera operativa cuando están unidas mediante un enlace covalente de tal manera que la expresión o transcripción de la secuencia de ácido nucleico quede bajo la influencia o el control de las secuencias reguladoras. Si se desea que las secuencias de ácidos nucleicos se traduzcan para obtener una proteína funcional, se dice que dos secuencias de ADN están enlazadas de manera operativa si la inducción de un promotor en las secuencias reguladoras 5' produce la transcripción de la secuencia de codificación y si la naturaleza del enlace entre las dos secuencias de ADN no (1) produce la introducción de una mutación de cambio de marco, (2) interfiere con la capacidad de la región promotora de dirigir la transcripción de las secuencias de codificación, ni (3) interfiere con la capacidad de la transcripción de ARN correspondiente de traducirse en una proteína. Por lo tanto, una región promotora se enlazaría de

manera operativa a una secuencia de ácido nucleico si la región promotora fuera capaz de efectuar la transcripción de esa secuencia de ADN de forma tal que la transcripción resultante se pudiera traducir en la proteína o el polipéptido deseado. De manera similar, dos o más regiones codificantes están enlazadas de manera operativa cuando se enlazan de tal manera que su transcripción a partir de un promotor común produzca la expresión de dos o más proteínas que se hayan traducido en el marco. En algunas modalidades, las secuencias codificantes enlazadas de manera operativa producen una proteína de fusión.

Una región que comprende un transgén (p. ej., que comprende una proteína de fusión, etc.) puede posicionarse en cualquier ubicación adecuada del ácido nucleico aislado que permita la expresión de la proteína de fusión.

Se debe apreciar que, en los casos en los que un transgén codifica más de un polipéptido, cada polipéptido puede posicionarse en cualquier ubicación adecuada dentro del transgén. Por ejemplo, un ácido nucleico que codifica un primer polipéptido puede posicionarse en un intrón del transgén y una secuencia de ácido nucleico que codifica un segundo polipéptido puede posicionarse en otra región no traducida (p. ej., entre el último codón de una secuencia de codificación de la proteína y la primera base de la señal de poli-A del transgén).

Un «promotor» hace referencia a una secuencia de ADN reconocida por la maquinaria sintética de la célula, o la maquinaria sintética introducida, necesaria para iniciar la transcripción específica de un gen. Las expresiones «enlazado de manera operativa», «posicionado de manera operativa», «bajo control» o «bajo control de transcripción» significan que el promotor se encuentra en la ubicación y orientación correcta en relación con el ácido nucleico para controlar el inicio de la polimerasa de ARN y la expresión del gen.

Para los ácidos nucleicos que codifican proteínas, por lo general una secuencia de poliadenilación se inserta luego de las secuencias de transgenes y antes de la secuencia de ITR de AAV 3'. Una construcción de rAAV útil en la presente invención también puede contener un intrón, ubicado convenientemente entre la secuencia promotora/potenciadora y el transgén. Una posible secuencia intrónica deriva del SV-40 y se denomina la «secuencia intrónica de SV-40 T». Otro elemento vectorial que se puede usar es un sitio interno de entrada al ribosoma (IRES, por sus siglas en inglés). Se usa una secuencia de IRES para producir más de un polipéptido a partir de la transcripción de un solo gen.

Se usaría una secuencia de IRES para producir una proteína que contuviera más de una cadena polipeptídica. La selección de estos y otros elementos de vectores comunes es convencional y existen muchas secuencias de ese tipo disponibles [véase, p. ej., Sambrook et ál., y las referencias allí citadas, por ejemplo, en las páginas 3.18 3.26 y 16.17 16.27 y Ausubel et ál., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, 1989]. En algunas modalidades, se incluye una

secuencia del virus de la fiebre aftosa 2A en la poliproteína; este es un péptido pequeño (de aproximadamente 18 aminoácidos de longitud) que se ha demostrado que media la escisión de las poliproteínas (Ryan, M D et ál., EMBO, 1994; 4: 928-933; Mattion, N M et ál., J Virology, November 1996; págs. 8124-8127; Furler, S et ál., Gene Therapy, 2001; 8: 864-873; y Halpin, C et ál., The Plant Journal, 1999; 4: 453-459). La actividad de escisión de la secuencia 2A se ha demostrado anteriormente en sistemas artificiales que incluyen plásmidos y vectores para terapias génicas (AAV y retrovirus) (Ryan, M D et ál., EMBO, 1994; 4: 928-933; Mattion, N M et ál., J Virology, November 1996; págs. 8124-8127; Furler, S et ál., Gene Therapy, 2001; 8: 864-873; y Halpin, C et ál., The Plant Journal, 1999; 4: 453-459; de Felipe, P et ál., Gene Therapy, 1999; 6: 198-208; de Felipe, P et ál., Human Gene Therapy, 2000; 11: 1921-1931.; y Klump, H et ál., Gene Therapy, 2001; 8: 811-817).

Algunos ejemplos de promotores constitutivos incluyen, de modo no taxativo, el promotor de LTR del virus del sarcoma de Rous (RSV, por sus siglas en inglés) retroviral (opcionalmente con el potenciador del RSV), el promotor del citomegalovirus (CMV) (opcionalmente con el potenciador del CMV) [véase, p. ej., Boshart et ál., Cell, 41: 521-530 (1985)], el promotor del SV40, el promotor de la dihidrofolato reductasa, el promotor de la  $\beta$ -actina, el promotor de la fosfoglicerol cinasa (PGK, por sus siglas en inglés) y el promotor de EF1 $\alpha$  [Invitrogen]. En algunas modalidades, un promotor es un promotor de P2. En algunas modalidades, un promotor es un promotor de la  $\beta$ -actina de pollo (CBA, por sus siglas en inglés). En algunas modalidades, un promotor consiste en dos promotores de la CBA. En algunas modalidades, un promotor consiste en dos promotores de la CBA separados por un promotor del CMV. En algunas modalidades, un promotor es un promotor de CAG.

Los promotores inducibles permiten la regulación de la expresión génica y pueden ser regulados por compuestos suministrados de manera exógena, factores del entorno, tales como la temperatura, o la presencia de un estado fisiológico específico, p. ej., una fase aguda, un estado de diferenciación particular de la célula o solamente en las células de replicación. Los promotores inducibles y sistemas inducibles están disponibles a través de una serie de fuentes comerciales, que incluyen, de modo no taxativo, Invitrogen, Clontech y Ariad. Se han descrito muchos otros sistemas que pueden ser seleccionados fácilmente por el experto en la técnica. Los ejemplos de promotores inducibles regulados por promotores suministrados de manera exógena incluyen el promotor de la metalotionina (MT) de la oveja inducible mediante cinc, el promotor del virus del tumor mamario en ratones (MMTV, por sus siglas en inglés) inducible mediante dexametasona (Dex), el sistema promotor de la polimerasa T7 (WO 98/10088); el promotor de insectos de ecdisona (No et ál., Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), el sistema pasible de represión mediante tetraciclina (Gossen et ál., Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), el sistema inducible mediante tetraciclina (Gossen et ál., Science, 268:1766-1769 (1995), véase también Harvey et ál., Curr. Opin. Chem. Biol., 2:512-518

(1998)), el sistema inducible por RU486 (Wang et ál., *Nat. Biotech.*, 15:239-243 (1997) y Wang et ál., *Gene Ther.*, 4:432-441 (1997)) y el sistema inducible por rapamicina (Magari et ál., *J. Clin. Invest.*, 100:2865-2872 (1997)). Incluso otros tipos de promotores inducibles que pueden ser útiles en este contexto son aquellos que son regulados por un estado fisiológico específico, p. ej., la temperatura, la fase aguda, un estado de diferenciación particular de la célula o solamente en las células de replicación.

En otra modalidad, se usará el promotor natural para el transgén. Se puede preferir el promotor natural cuando se desee que la expresión del transgén imite la expresión natural. El promotor natural puede usarse cuando la expresión del transgén deba regularse temporalmente o a nivel del desarrollo, o con especificidad del tejido, o en respuesta a estímulos de transcripción específicos. En una modalidad adicional, otros elementos de control de la expresión naturales, tales como los elementos potenciadores, los sitios de poliadenilación o las secuencias de consenso de Kozak también pueden usarse para imitar la expresión natural.

En algunas modalidades, las secuencias reguladoras confieren capacidades de expresión génica con especificidad de tejido. En algunos casos, las secuencias reguladoras con especificidad de tejido se unen a factores de transcripción con especificidad de tejido que inducen la transcripción con especificidad de tejido. Esas secuencias reguladoras con especificidad de tejido (p. ej., promotores, potenciadores, etc.) son muy conocidas en la técnica. Los ejemplos de secuencias reguladoras con especificidad de tejido incluyen, de modo no taxativo, los siguientes promotores con especificidad de tejido: un promotor de globulina que se une a tiroxina (TBG, por sus siglas en inglés) específico del hígado, un promotor de la insulina, un promotor del glucagón, un promotor de la somatostatina, un promotor de los polipéptidos pancreáticos (PPY), un promotor de la sinapsina-1 (Syn), un promotor de la creatina cinasa (MCK), un promotor de la desmina (DES) de mamíferos, un promotor de la cadena pesada de  $\alpha$ -miosina ( $\alpha$ -MHC) y un promotor de la Troponina T cardíaca (cTnT). Otros ejemplos de promotores incluyen el promotor de la Beta-actina, el promotor del núcleo del virus de la hepatitis B, Sandig et ál., *Gene Ther.*, 3:1002-9 (1996); el promotor de la alfa-fetoproteína (AFP), Arbuthnot et ál., *Hum. Gene Ther.*, 7:1503-14 (1996)), el promotor de la osteocalcina de los huesos (Stein et ál., *Mol. Biol. Rep.*, 24:185-96 (1997)); el promotor de la sialoproteína de los huesos (Chen et ál., *J. Bone Miner. Res.*, 11:654-64 (1996)), el promotor de CD2 (Hansal et ál., *J. Immunol.*, 161:1063-8 (1998); el promotor de la cadena pesada de la inmunoglobulina; el promotor de la cadena  $\alpha$  de los receptores de linfocitos T, promotores neuronales tales como el de la enolasa específica de las neuronas (NSE, por sus siglas en inglés) (Andersen et ál., *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), el promotor del gen de cadena ligera de los neurofilamentos (Piccioli et ál., *Proc. Natl. Acad. Sci. USA*, 88:5611-5 (1991)), y el promotor del gen *vgf* específico de las neuronas (Piccioli et ál., *Neuron*, 15:373-84 (1995)), entre otros

que serán evidentes para el experto en la técnica.

En algunas modalidades, un transgén que codifica una proteína de fusión que comprende un DBD y un transactivador está enlazado de manera operativa a un promotor. En algunas modalidades, el promotor es un promotor constitutivo. En algunas modalidades, el promotor es un promotor inducible.

- 5 En algunas modalidades, el promotor es un promotor con especificidad de tejido. En algunas modalidades, el promotor es específico del tejido nervioso. En algunas modalidades, el promotor es un promotor de SST o NPY.

- Algunos aspectos de la descripción están vinculados con un ácido nucleico aislado que comprende más de un promotor (p. ej., 2, 3, 4, 5 o más promotores). Por ejemplo, en el contexto de una construcción con un transgén que comprende una primera región que codifica una proteína y una  
10 segunda región que codifica una proteína, puede ser conveniente impulsar la expresión de la primera región codificante de la proteína mediante el uso de una primera secuencia promotora (p. ej., una primera secuencia promotora enlazada de manera operativa a la región codificante de la proteína), e impulsar la expresión de la segunda región codificante de la proteína con una segunda secuencia promotora (p. ej., una segunda secuencia promotora enlazada de manera operativa a la segunda  
15 región codificante de la proteína). Por lo general, la primera secuencia promotora y la segunda secuencia promotora pueden ser la misma secuencia promotora o secuencias promotoras diferentes. En algunas modalidades, la primera secuencia promotora (p. ej., el promotor que impulsa la expresión de la región codificante de la proteína) es una secuencia promotora de ARN polimerasa III (pol III). Los  
20 ejemplos no taxativos de secuencias promotoras de pol III incluyen secuencias promotoras de H1 y U6. En algunas modalidades, la segunda secuencia promotora (p. ej., la secuencia promotora que impulsa la expresión de la segunda proteína) es una secuencia promotora de ARN polimerasa II (pol II). Los ejemplos no taxativos de secuencias promotoras de pol III incluyen secuencias promotoras de T7, T3, SP6, RSV y citomegalovirus. En algunas modalidades, una secuencia promotora de pol III  
25 impulsa la expresión de la primera región codificante de la proteína. En algunas modalidades, una secuencia promotora de pol II impulsa la expresión de la segunda región codificante de la proteína.

#### *Virus adenoasociados recombinantes (rAAV)*

- 30 En algunos aspectos, la descripción proporciona virus adenoasociados (AAV) aislados. Tal como se usa en la presente con respecto a los AAV, el término «aislado» hace referencia a un AAV que se ha producido u obtenido artificialmente. Los AAV aislados pueden producirse mediante el uso de métodos recombinantes. En la presente se hace referencia a esos AAV como «AAV recombinantes». Los AAV recombinantes (rAAV) tienen preferentemente capacidades de direccionamiento con especificidad de

tejido, de forma tal que se suministre una nucleasa y/o un transgén del rAAV específicamente a uno o más tejidos predeterminados. La cápside del AAV es un elemento importante para determinar estas capacidades de direccionamiento con especificidad de tejido. Por lo tanto, se puede seleccionar un rAAV que tenga una cápside adecuada para el tejido diana.

- 5 En la técnica se conocen muy bien los métodos para obtener AAV recombinantes con una proteína de cápside deseada. (Véase, por ejemplo, US 2003/0138772), cuyo contenido se incorpora en su totalidad a la presente mediante referencia). Normalmente, los métodos implican cultivar una célula hospedadora que contiene una secuencia de ácido nucleico que codifica una proteína de cápside de AAV; un gen rep funcional; un vector de AAV recombinante compuesto por repeticiones terminales invertidas (ITR) de AAV y un transgén; y funciones auxiliares suficientes para permitir la inclusión del vector de AAV recombinante en las proteínas de cápside de AAV. En algunas modalidades, las proteínas de cápside son proteínas estructurales codificadas por el gen cap de un AAV. Los AAV comprenden tres proteínas de cápside, las proteínas de viriones 1 a 3 (denominadas VP1, VP2 y VP3), todas las cuales se transcriben a partir de un solo gen cap a través de un corte y empalme alternativo.
- 10 En algunas modalidades, los pesos moleculares de VP1, VP2 y VP3 son, respectivamente, de alrededor de 87 kDa, alrededor de 72 kDa y alrededor de 62 kDa. En algunas modalidades, a partir de la traducción, las proteínas de cápside forman una proteína esférica de 60 unidades alrededor del genoma viral. En algunas modalidades, las funciones de las proteínas de cápside son proteger el genoma viral, suministrar el genoma e interactuar con el hospedador. En algunos aspectos, las proteínas de cápside suministran el genoma viral a un hospedador con especificidad del tejido.
- 15 En algunas modalidades, una proteína de cápside de AAV pertenece a un serotipo de AAV que se selecciona dentro del grupo que consiste en AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAV9, AAV10, AAVrh10 y AAV.PHP.B. En algunas modalidades, una proteína de cápside de AAV pertenece a un serotipo derivado de un primate no humano, por ejemplo, un serotipo AAVrh8. En algunas modalidades, una proteína de cápside de AAV pertenece a un serotipo obtenido para la transducción amplia y eficiente del SNC, por ejemplo, AAV.PHP.B. En algunas modalidades, la proteína de cápside pertenece al serotipo 9 del AAV.
- 20

- Los componentes que se han de cultivar en la célula hospedadora para incluir un vector de rAAV en una cápside de AAV pueden proporcionarse a la célula hospedadora en trans. De manera alternativa, uno o más cualesquiera de los componentes necesarios (p. ej., vector de AAV recombinante, secuencias *rep*, secuencias *cap* y/o funciones auxiliares) pueden ser proporcionados por una célula hospedadora estable que se haya modificado genéticamente para que contenga uno o más de los componentes necesarios, mediante el uso de métodos conocidos por los expertos en la técnica. Más adecuadamente, una célula hospedadora estable de ese tipo contendrá el o los componentes
- 30

necesarios bajo el control de un promotor inducible. Sin embargo, el o los componentes necesarios pueden estar bajo el control de un promotor constitutivo. En la presente, se proporcionan ejemplos de promotores inducibles y constitutivos adecuados, en la descripción de elementos reguladores adecuados para el uso con el transgén. En incluso otra alternativa, una célula hospedadora estable seleccionada puede contener uno o más componentes seleccionados bajo el control de un promotor constitutivo y otro u otros componentes seleccionados bajo el control de uno o más promotores inducibles. Por ejemplo, se puede generar una célula hospedadora estable que derive de células 293 (que contengan funciones auxiliares E1 bajo el control de un promotor constitutivo), pero que contenga las proteínas *rep* y/o *cap* bajo el control de promotores inducibles. Incluso otras células hospedadoras estables pueden ser generadas por el experto en la técnica.

En algunas modalidades, la presente descripción está vinculada con una célula hospedadora que contiene un ácido nucleico que comprende una secuencia de codificación que codifica un transgén (p. ej., un dominio de unión a ADN fusionado con un dominio regulador de la transcripción). En algunas modalidades, la célula hospedadora es una célula de mamífero, una célula de levadura, una célula bacteriana, una célula de insecto, una célula de planta o una célula de levadura.

El vector de AAV recombinante, las secuencias *rep*, las secuencias *cap* y las funciones auxiliares que se necesitan para producir el rAAV de la descripción pueden suministrarse a la célula hospedadora encapsuladora mediante el uso de cualquier elemento genético adecuado (vector). El elemento genético seleccionado puede suministrarse a través de cualquier método adecuado, incluidos los que se describen en la presente. Los métodos usados para construir cualquier modalidad de la presente descripción son conocidos por los expertos en la manipulación de ácidos nucleicos e incluyen la modificación genética, modificación recombinante y técnicas sintéticas. Véase, p. ej., Sambrook et ál., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY. De manera similar, los métodos de generación de viriones de rAAV son muy conocidos y la selección de un método adecuado no implica una limitación de la presente descripción. Véase, p. ej., K. Fisher et ál., *J. Virol.*, 70:520-532 (1993) y la pat. estadounidense n.º 5,478,745.

En algunas modalidades, los AAV recombinantes pueden producirse mediante el uso del método de transfección triple (descrito en detalle en la pat. estadounidense n.º 6,001,650). Normalmente, los AAV recombinantes se producen mediante la transfección de una célula hospedadora con un vector de AAV (que comprende un transgén flanqueado por elementos de ITR) que se ha de incluir en partículas de AAV, un vector de función auxiliar de AAV y un vector de función accesoria. Un vector de función auxiliar de AAV codifica las secuencias «de función auxiliar de AAV» (p. ej., *rep* y *cap*), que funcionan en trans para la replicación y encapsidación de AAV de carácter productivo. Preferentemente, el vector de función auxiliar de AAV respalda la producción del vector de AAV con eficiencia sin generar viriones

de AAV naturales detectables (p. ej., viriones de AAV que contengan genes rep y cap funcionales). Los ejemplos no taxativos de vectores adecuados para el uso con la presente descripción incluyen pHLP19, que se describe en la pat. estadounidense n.º 6,001,650 y el vector pRep6cap6, descrito en la pat. estadounidense n.º 6,156,303, que se incorporan a la presente mediante referencia en ambos

5 casos. El vector de función accesoria codifica las secuencias de nucleótidos para las funciones celulares y/o virales no derivadas de AAV de las cuales depende el AAV para la replicación (p. ej., «funciones accesorias»). Las funciones accesorias incluyen aquellas funciones necesarias para la replicación del AAV, lo cual incluye, de modo no taxativo, aquellos restos involucrados en la activación de la transcripción génica de AAV, el corte y empalme de ARNm de AAV de etapa específica, la

10 replicación del ADN de AAV, la síntesis de los productos de expresión de cap y el ensamblaje de la cápside de AAV. Las funciones accesorias de base viral pueden derivar de cualquiera de los virus auxiliares conocidos, tales como el adenovirus, el virus del herpes (diferente del virus del herpes simple tipo 1) y el virus vacuna.

En algunos aspectos, la descripción proporciona células hospedadoras transfectadas. El término

15 «transfección» se usa para hacer referencia a la absorción de ADN externo por parte de una célula, y una célula se ha «transfectado» cuando se ha introducido ADN exógeno dentro de la membrana celular. En la técnica, se conoce una cantidad de procedimientos de transfección. Véase, p. ej., Graham et ál. (1973) *Virology*, 52:456, Sambrook et ál. (1989) *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratories, Nueva York, Davis et ál. (1986) *Basic Methods in Molecular Biology*,

20 Elsevier y Chu et ál. (1981) *Gene* 13:197. Esas técnicas pueden usarse para introducir uno o más ácidos nucleicos exógenos, tales como un vector de integración de nucleótidos y otras moléculas de ácido nucleico, en células hospedadoras adecuadas.

Una «célula hospedadora» hace referencia a cualquier célula que albergue, o que sea capaz de albergar, una sustancia de interés. A menudo, una célula hospedadora es una célula de mamífero. En

25 algunas modalidades, una célula hospedadora es una neurona, opcionalmente una neurona gabaérgica. Una «neurona gabaérgica», tal como se usa en la presente, es una célula neuronal que genera ácido gamma aminobutírico (GABA). En los mamíferos, el GABA es un neurotransmisor que se distribuye de manera extendida en el sistema nervioso y se une y reprime las neuronas a las que se une. De esa manera, el GABA está involucrado en numerosos trastornos que afectan el sistema

30 nervioso, incluida la epilepsia, el autismo y la ansiedad. En los estudios en ratones con inactivación y hemiciotos de *SCN1A* se ha observado un déficit profundo de circulación de sodio en las neuronas gabaérgicas del cerebro. Una célula hospedadora puede usarse como receptora de una construcción auxiliar de AAV, un plásmido minigénico de AAV, un vector de función accesoria u otro ADN de transferencia asociado a la producción de AAV recombinantes. La expresión incluye la progenie de la

célula original que se ha transfectado. Por lo tanto, una «célula hospedadora», tal como se usa en la presente, puede hacer referencia a una célula que se haya transfectado con una secuencia de ADN exógena. Se entiende que la progenie de una célula original individual puede no necesariamente ser completamente idéntica en morfología o en el complemento de ADN total o genómico al progenitor original, debido a mutaciones naturales, accidentales o deliberadas.

Tal como se usa en la presente, la expresión «línea celular» hace referencia a una población de células capaces de un crecimiento continuo o prolongado y de división in vitro. A menudo, las líneas celulares son poblaciones de clones derivadas de una sola célula progenitora. En la técnica se sabe además que se pueden producir cambios espontáneos o inducidos en el cariotipo durante el almacenamiento o la transferencia de esas poblaciones de clones. Por lo tanto, las células derivadas de la línea celular mencionada pueden no ser precisamente idénticas a las células o los cultivos ancestrales, y la línea celular mencionada incluye esas variantes.

Tal como se usa en la presente, la expresión «célula recombinante» hace referencia a una célula en la cual se ha introducido un segmento de ADN exógeno, tal como un segmento de ADN que lleva a la transcripción de un polipéptido biológicamente activo o la producción de un ácido nucleico biológicamente activo, tal como un ARN.

Tal como se usa en la presente, el término «vector» incluye cualquier elemento genético, tal como un plásmido, fago, transposón, cósmido, cromosoma, cromosoma artificial, virus, virión, etc., capaz de replicación cuando está asociado a los elementos de control adecuados y que puede transferir secuencias génicas entre las células. En algunas modalidades, un vector es un vector viral, tal como un vector de rAAV, un vector lentiviral, un vector adenoviral, un vector retroviral, etc. Por lo tanto, el término incluye vehículos de clonación y expresión, así como vectores virales. En algunas modalidades, se contempla que los vectores útiles son aquellos vectores en los cuales el segmento de ácido nucleico que se ha de transcribir se ubica bajo el control de transcripción de un promotor.

Un «promotor» hace referencia a una secuencia de ADN reconocida por la maquinaria sintética de la célula, o la maquinaria sintética introducida, necesaria para iniciar la transcripción específica de un gen. Las expresiones «enlazado de manera operativa», «posicionado de manera operativa», «bajo control» o «bajo control de transcripción» significan que el promotor se encuentra en la ubicación y orientación correcta en relación con el ácido nucleico para controlar el inicio de la polimerasa de ARN y la expresión del gen. La expresión «construcción o vector de expresión» significa cualquier tipo de construcción genética que contenga un ácido nucleico, donde la totalidad o parte de la secuencia que codifica el ácido nucleico sea pasible de transcripción. En algunas modalidades, la expresión incluye la transcripción del ácido nucleico, por ejemplo, para generar un producto polipeptídico biológicamente activo a partir de un gen transcrito. No se pretende que los métodos que anteceden para incluir

vectores recombinantes en cápsides de AAV deseadas para producir los rAAV de la descripción sean taxativos y otros métodos adecuados serán evidentes para el experto en la técnica.

*Métodos para regular la expresión del gen diana*

5

La descripción proporciona métodos para regular la expresión génica en una célula o un sujeto. Normalmente, los métodos implican administrar a una célula o un sujeto un ácido nucleico aislado o rAAV que comprenda un transgén que codifique una proteína de fusión que comprenda un dominio de unión a ADN (p. ej., un dominio de ZFP) y un dominio de transactivación. En algunas modalidades, una  
10 proteína de fusión comprende ZFP y transactivador VP64. En algunas modalidades, una proteína de fusión comprende ZFP y transactivador p65. En algunas modalidades, una proteína de fusión comprende ZFP y transactivador RTA. En algunas modalidades, una proteína de fusión comprende ZFP y transactivador VPR. En algunas modalidades, el método implica administrar a una célula o un sujeto una proteína dCas9 y al menos un ácido nucleico guía que se dirija al *SCN1A* (p. ej., un ácido  
15 nucleico guía que comprenda cualquiera de las SEQ ID NO: 83-94 o codificado por cualquiera de las SEQ ID NO: 83-94).

Administrar un ácido nucleico aislado o un rAAV que codifique la proteína de fusión (p. ej., una proteína de fusión que comprenda un transactivador) a una célula o un sujeto, en algunas modalidades, produce un aumento de la expresión de un gen diana (p. ej., *SCN1A*). Por lo tanto, en  
20 algunas modalidades, las composiciones y los métodos descritos en la descripción son útiles para tratar afecciones que son el resultado de una haploinsuficiencia de un gen diana, tal como el síndrome de Dravet que es el resultado de la haploinsuficiencia del gen *SCN1A*.

Tal como se usa en la presente, una «haploinsuficiencia» hace referencia a una condición genética donde una copia de un gen (p. ej., *SCN1A*) se inactiva, p. ej., mediante mutación genética, o se  
25 elimina, y la copia funcional restante del gen no es adecuada para producir una cantidad suficiente de producto génico como para preservar la función normal del gen.

El síndrome de Dravet, también conocido como la epilepsia mioclónica grave de la infancia, es una forma de epilepsia poco frecuente y crónica que se manifiesta normalmente en los primeros tres años de vida. El síndrome de Dravet se caracteriza por convulsiones prolongadas y frecuentes, retrasos del  
30 comportamiento y el desarrollo, problemas de movimiento y equilibrio, retraso en el lenguaje y problemas con el habla y alteraciones del sistema nervioso autónomo. En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de Dravet, tal como una copia del gen *SCN1A* que se muta, lo cual da como resultado una reducción de la proteína *SCN1A* en una célula o un sujeto. La mayoría de los pacientes con síndrome de Dravet porta mutaciones de *SCN1A* que se traducen en

proteínas truncadas; otras mutaciones de *SCN1A* asociadas al síndrome de Dravet incluyen mutaciones del sitio de corte y empalme e interruptoras, así como mutaciones distribuidas aleatoriamente en todo el gen *SCN1A*. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *SCN1A* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *SCNA1* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *SCN1A*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de haploinsuficiencia de *MED13L*, donde el sujeto solamente tiene una única copia funcional del gen *MED13L*. Los sujetos que padecen el síndrome de haploinsuficiencia de *MED13L* tienen normalmente una mutación en su segunda copia no funcional del gen *MED13L*. El síndrome de haploinsuficiencia de *MED13L* se caracteriza por la discapacidad intelectual, problemas del habla, rasgos faciales distintivos y retrasos en el desarrollo. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *MED13L* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *MED13L* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *MED13L*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada a síndromes mielodisplásicos. Los sujetos que padecen un síndrome mielodisplásico normalmente tienen una mutación en una copia de los genes isocitrato deshidrogenasa 1 (*IDH1*), isocitrato deshidrogenasa 2 (*IDH2*) y/o *GATA2*. Los síndromes mielodisplásicos son un grupo de cánceres en los cuales los hematocitos inmaduros en la médula ósea no maduran para convertirse en hematocitos sanos. Ocasionalmente, este síndrome puede conducir a una leucemia mieloide aguda. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad a un gen *IDH1* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *IDH1* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *IDH1*. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad a un gen *IDH2* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *IDH2* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige

(p. ej., se une) con especificidad al gen *IDH2*. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad a un gen *GATA2* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *GATA2* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *GATA2*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de DiGeorge. Los sujetos que padecen un síndrome de DiGeorge normalmente tienen una eliminación de 30 a 40 genes en el medio del cromosoma 22 en una ubicación conocida como 22q11.2. En particular, la enfermedad puede caracterizarse por la haploinsuficiencia del gen *TBX*. El síndrome de DiGeorge se caracteriza por problemas cardíacos congénitos, rasgos faciales específicos, infecciones frecuentes, retrasos en el desarrollo, problemas de aprendizaje y paladar hendido. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *TBX* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *TBX* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad a un gen *TBX*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome CHARGE. En la mayoría de los casos, los sujetos que padecen el síndrome CHARGE tienen una haploinsuficiencia del gen *CHD7*. El síndrome CHARGE se caracteriza por colobomas oculares, defectos del corazón, atresia de las coanas, retrasos del crecimiento y/o desarrollo, anomalías genitales y/o urinarias y anomalías en los oídos y sordera. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *CHD7* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *CHD7* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *CHD7*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de Ehlers-Danlos. Los sujetos que padecen el síndrome de Ehlers-Danlos pueden tener una haploinsuficiencia de los genes *COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *COL5A2*, *TNXB*, *ADAMTS2*, *PLOD1*, *B4GALT7*, *DSE* y/o *D4ST1/CHST14*. El síndrome de Ehlers-Danlos se caracteriza por la hiperelasticidad de la piel y puede provocar disección aórtica, escoliosis y osteoartritis de inicio temprano. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad a cualquiera de los genes *COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*,

*COL5A2*, *TNXB*, *ADAMTS2*, *PLOD1*, *B4GALT7*, *DSE* o *D4ST1/CHST14* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse a cualquiera de *COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *COL5A2*, *TNXB*, *ADAMTS2*, *PLOD1*, *B4GALT7*, *DSE* o *D4ST1/CHST14* comprende (i) una proteína de fusión que comprende una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad a cualquiera de los genes *COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *COL5A2*, *TNXB*, *ADAMTS2*, *PLOD1*, *B4GALT7*, *DSE* o *D4ST1/CHST14*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada a demencias frontotemporales (FTD, por sus siglas en inglés). Los sujetos que padecen FTD tienen una haploinsuficiencia del gen *MAPT*, que codifica la proteína Tau, y/o del gen *GRN*. La FTD se caracteriza por la pérdida de memoria, la falta de habilidades sociales, el poco control de los impulsos y dificultades en el habla. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *MAPT* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *MAPT* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad a un gen *MAPT*. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *GRN* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *GRN* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad a un gen *GRN*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de Holt-Oram. Los sujetos que padecen el síndrome de Holt-Oram tienen una haploinsuficiencia del gen *TBX5*. El síndrome de Holt-Oram se caracteriza por complicaciones cardíacas, que incluyen defectos cardíacos congénitos y la enfermedad de la conducción cardíaca. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *TBX5* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *TBX5* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *TBX5*.

En algunas modalidades, un sujeto tiene una haploinsuficiencia asociada al síndrome de Marfan. Los sujetos que padecen el síndrome de Marfan normalmente tienen una haploinsuficiencia del gen *FBN1*, que codifica la proteína fibrilina-1. El síndrome de Marfan se caracteriza por longitudes desproporcionadas de las extremidades, la artritis de inicio temprano, complicaciones cardíacas y/o la

disfunción del sistema nervioso autónomo. En algunas modalidades, una proteína de fusión de la descripción comprende un dominio de ZFP que se dirige (p. ej., se une) con especificidad al gen *FBN1* y un dominio de transactivación. En algunas modalidades, una composición para dirigirse al *FBN1* comprende (i) una proteína de fusión que comprende: una proteína dCas y un dominio de transactivación, y (ii) un ácido nucleico guía (p. ej., un ARNg) que se dirige (p. ej., se une) con especificidad al gen *FBN1*.

La descripción se basa, en parte, en métodos de administración de una proteína de fusión tal como se describe en la presente a un sujeto. En algunas modalidades, la proteína de fusión comprende un DBD y un activador de la transcripción. En algunas modalidades, el DBD es un ZNF, una TALE, una proteína dCas (p. ej., dCas9 o dCas12a), o un homeodominio que se une a un gen *SCN1A*. En algunas modalidades, el activador de la transcripción es VP64, p65, RTA o un activador de la transcripción de carácter tripartito que comprende VP64-p65-RTA (VPR). En algunas modalidades, la proteína de fusión está flanqueada por secuencias de repeticiones terminales invertidas (ITR) de AAV. En algunas modalidades, la proteína de fusión está enlazada de manera operativa a un promotor. En algunas modalidades, el sujeto tiene, o se presume que tiene, mutaciones en el *SCN1A* que producen una haploinsuficiencia de la proteína *SCN1A*. En algunas modalidades, el sujeto tiene, o se presume que tiene, el síndrome de Dravet.

En algunos aspectos, la descripción proporciona métodos de modulación (p. ej., aumento, disminución, etc.) de la expresión de un gen diana en una célula. En algunas modalidades, la descripción proporciona métodos de aumento de la expresión de un gen diana (p. ej., *SCN1A*) en una célula. En algunas modalidades, una célula es una célula de un mamífero. En algunas modalidades, la célula está en un sujeto (p. ej., *in vivo*). En algunas modalidades, un sujeto es un sujeto mamífero, por ejemplo, un ser humano. En algunas modalidades, una célula es una célula del sistema nervioso (célula del sistema nervioso central o célula del sistema nervioso periférico), por ejemplo, una neurona (p. ej., neuronas gabaérgicas, neuronas unipolares, neuronas bipolares, células en cesta, células de Betz, células de Lugaro, neuronas espinosas, células de Purkinje, células piramidales, células de Renshaw, células granulosas, neuronas motoras, células fusiformes, etc.) o células gliales (p. ej., astrocitos, oligodendrocitos, células endoteliales, células gliales radiales, células de Schwann, células satélite, etc.).

En una célula o un sujeto «normal», la expresión de un gen diana (p. ej., *SCN1A*) es suficiente de forma tal que la célula o el sujeto no tiene haploinsuficiencia con respecto al gen diana (p. ej., *SCN1A*). En algunas modalidades, la «mejora» o el «aumento» de la expresión o actividad de un transgén se mide con respecto a la expresión o actividad de ese transgén en una célula o un sujeto que no ha recibido la administración de uno o más ácidos nucleicos aislados, rAAV o composiciones tal como se

describe en la presente. En algunas modalidades, la «mejora» o el «aumento» de la expresión o actividad de un transgén se mide con respecto a la expresión o actividad de ese transgén en el sujeto después de que el sujeto ha recibido la administración (p. ej., la expresión génica se mide antes y después de la administración de) uno o más ácidos nucleicos aislados, rAAV o composiciones tal como se describe en la presente. Por ejemplo, en algunas modalidades, la «mejora» o el «aumento» de la expresión de *SCN1A* en una célula o un sujeto se mide con respecto a una célula o un sujeto que no ha recibido la administración de un transgén que codifica un transactivador-ZFP de fusión. En algunas modalidades, los métodos descritos en la descripción producen un aumento de la expresión y/o actividad de *SCN1A* en sujeto de entre 2 veces y 100 veces (p. ej., 2 veces, 5 veces, 10 veces, 50 veces, 100 veces, etc.) con respecto a la expresión y/o actividad de *SCN1A* de un sujeto que no ha recibido la administración de una o más composiciones descritas en la descripción.

Tal como se usa en la presente, los términos «tratamiento», «tratar» y «terapia» hacen referencia al tratamiento terapéutico y a manipulaciones profilácticas o preventivas. Los términos incluyen además mejorar los síntomas existentes, prevenir síntomas adicionales, mejorar o prevenir las causas subyacentes de los síntomas, prevenir o revertir las causas de los síntomas, por ejemplo, síntomas asociados a un gen haploinsuficiente, p. ej., un gen *SCN1A* haploinsuficiente. Por lo tanto, estos términos indican que se ha conferido un resultado beneficioso a un sujeto con un trastorno (p. ej., una enfermedad o afección asociada a un gen haploinsuficiente, p. ej., el síndrome de Dravet), o con el potencial de desarrollar un trastorno de ese tipo. Además, el término «tratamiento» también incluye la aplicación o administración de un agente (p. ej., agente terapéutico o una composición terapéutica, p. ej., un ácido nucleico aislado o rAAV que se dirige o se une a un gen diana o una región reguladora de un gen diana) a un sujeto, o una línea celular o un tejido aislado de un sujeto, que puede tener una enfermedad, un síntoma de una enfermedad o predisposición a tener una enfermedad, con el fin de curar, solucionar, aliviar, paliar, alterar, remediar, mejorar o tener un efecto sobre la enfermedad, los síntomas de la enfermedad o la predisposición a tener la enfermedad.

Los agentes terapéuticos o las composiciones terapéuticas pueden incluir un compuesto en una forma farmacéuticamente aceptable que prevenga y/o reduzca los síntomas de una enfermedad particular (p. ej., una enfermedad o afección asociada a un gen haploinsuficiente, p. ej., el síndrome de Dravet). Por ejemplo, una composición terapéutica puede ser una composición farmacéutica que prevenga y/o reduzca los síntomas de una enfermedad asociada a un gen haploinsuficiente, p. ej., el síndrome de Dravet. Se contempla que la composición terapéutica de la presente invención se proporcione en cualquier forma adecuada. La forma de la composición terapéutica dependerá de una cantidad de factores, que incluyen el modo de administración tal como se describe en la presente. La composición terapéutica puede contener diluyentes, adyuvantes y excipientes, entre otros ingredientes tal como se

describe en la presente.

### *Modos de administración*

5 Los ácidos nucleicos aislados, los rAAV y las composiciones de la descripción pueden suministrarse a un sujeto en composiciones de acuerdo con cualquier método adecuado conocido en la técnica. Por ejemplo, un rAAV, preferentemente suspendido en un portador fisiológicamente compatible (p. ej., en una composición), puede administrarse a un sujeto, es decir, un animal hospedador, tal como un ser humano, un ratón, una rata, un gato, un perro, una oveja, un conejo, un caballo, una vaca, una cabra,  
10 un cerdo, un cobayo, un hámster, un pollo, un pavo o un primate no humano (p. ej., un macaco). En algunas modalidades, un animal hospedador no incluye un ser humano.

El suministro de los rAAV a un sujeto mamífero puede ser, por ejemplo, mediante inyección intramuscular o mediante la administración en el torrente sanguíneo del sujeto mamífero. La administración en el torrente sanguíneo puede ser mediante inyección en una vena, una arteria o  
15 cualquier otro conducto vascular. En algunas modalidades, los rAAV se administran en el torrente sanguíneo por medio de una perfusión en una extremidad aislada, una técnica muy conocida en la técnica quirúrgica, donde el método permite esencialmente que el experto aíse una extremidad de la circulación sistémica antes de la administración de los viriones de rAAV. Una variante de la técnica de perfusión en una extremidad aislada descrita en la pat. estadounidense n.º 6,177,403 también puede  
20 ser empleada por el experto en la técnica para administrar los viriones a la vasculatura de una extremidad aislada para poder llegar a mejorar la transducción en las células o el tejido de los músculos. Además, en determinados casos, puede ser conveniente suministrar los viriones al SNC de un sujeto. El «SNC» abarca todas las células y el tejido del cerebro y la médula espinal de un vertebrado. Por lo tanto, el término incluye, de modo no taxativo, las células neuronales, las células  
25 gliales, los astrocitos, el líquido cefalorraquídeo (LCR), los espacios intersticiales, los huesos, los cartílagos y similares. Los AAV recombinantes pueden suministrarse directamente al SNC o cerebro mediante inyección, p. ej., en la región ventricular, así como en el cuerpo estriado (p. ej., el núcleo caudado o putamen del cuerpo estriado), el tálamo, la médula espinal y la unión neuromuscular, o el lóbulo cerebeloso, con una aguja, un catéter o un dispositivo relacionado, mediante el uso de  
30 procedimientos neuroquirúrgicos conocidos en la técnica, por ejemplo, mediante inyección estereotáctica (véase, p. ej., Stein et ál., J Virol 73:3424-3429, 1999; Davidson et ál., PNAS 97:3428-3432, 2000; Davidson et ál., Nat. Genet. 3:219-223, 1993; y Alisky y Davidson, Hum. Gene Ther. 11:2315-2329, 2000). En algunas modalidades, un rAAV tal como se describe en la descripción se administra mediante inyección intravenosa. En algunas modalidades, los rAAV se administran

mediante inyección intracerebral. En algunas modalidades, los rAAV se administran mediante inyección intratecal. En algunas modalidades, los rAAV se administran mediante inyección en el cuerpo estriado. En algunas modalidades, los rAAV se suministran mediante inyección intracraneal. En algunas modalidades, los rAAV se suministran mediante inyección en la cisterna magna. En algunas modalidades, los rAAV se suministran mediante inyección en el ventrículo lateral cerebral.

Algunos aspectos de la presente descripción están vinculados con composiciones que comprenden un AAV recombinante que comprende una proteína de cápside y un ácido nucleico que codifica un transgén, donde el transgén comprende una secuencia de ácido nucleico que codifica una o más proteínas. En algunas modalidades, el ácido nucleico comprende además ITR de AAV. En algunas modalidades, una composición comprende además un portador farmacéuticamente aceptable.

Las composiciones de la descripción pueden comprender un rAAV solo, o en combinación con uno o más virus diferentes (p. ej., un segundo rAAV codificante que tiene uno o más transgenes diferentes). En algunas modalidades, una composición comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más rAAV diferentes, cada uno de los cuales tiene uno o más transgenes diferentes.

El experto en la técnica puede seleccionar fácilmente portadores adecuados en virtud de la indicación a la cual se dirige el rAAV. Por ejemplo, un portador adecuado incluye solución salina, la cual se puede formular con una variedad de soluciones amortiguadoras (p. ej., solución salina amortiguada con fosfato). Otros ejemplos de portadores incluyen solución salina estéril, lactosa, sacarosa, fosfato de calcio, gelatina, dextrano, agar, pectina, aceite de maní, aceite de sésamo y agua. La selección del portador no limita la presente descripción.

Opcionalmente, las composiciones de la descripción pueden contener, además del rAAV y el o los portadores, otros ingredientes farmacéuticos convencionales, tales como conservantes o estabilizadores químicos. Los ejemplos de conservantes adecuados incluyen clorobutanol, sorbato de potasio, ácido sórbico, dióxido de azufre, galato de propilo, los parabenos, etil vainillina, glicerina, fenol, paraclorofenol y poloxámeros (tensioactivos no iónicos) tales como Pluronic® F-68. Los estabilizadores químicos adecuados incluyen gelatina y albúmina.

Los rAAV se administran en cantidades suficientes para transfectar las células de un tejido deseado y para proporcionar niveles suficientes de transferencia y expresión génica sin efectos adversos indebidos. Las vías de administración convencionales y farmacéuticamente aceptables incluyen, de modo no taxativo, el suministro directo al órgano seleccionado (p. ej., suministro intraportal al hígado), la vía oral, por inhalación (que incluye el suministro intranasal e intratraqueal), intraocular, intravenosa, intramuscular, subcutánea, intradérmica, intratumoral y otras vías de administración parentales. Si se desea, las vías de administración pueden combinarse.

La dosis de viriones de rAAV necesaria para lograr un «efecto terapéutico» particular, es decir, las

unidades de dosis en copias de genoma/por kilogramo de peso corporal (CG/kg), variará en función de varios factores que incluyen, de modo no taxativo: la vía de administración de los viriones de rAAV, el nivel de expresión del gen o ARN necesario para lograr un efecto terapéutico, la enfermedad o el trastorno específico en tratamiento y la estabilidad del producto de ARN o gen. El experto en la técnica puede determinar fácilmente un intervalo de dosis de viriones de rAAV para tratar a un paciente con una enfermedad o un trastorno particular en función de los factores mencionados anteriormente, así como otros factores conocidos en la técnica.

Una cantidad eficaz de un rAAV es una cantidad suficiente para dirigirse a un animal e infectarlo, dirigirse a un tejido deseado. En algunas modalidades, se administra una cantidad eficaz de un rAAV al sujeto durante una etapa presintomática de la enfermedad de almacenamiento lisosómico. En algunas modalidades, la etapa presintomática de la enfermedad de almacenamiento lisosómico tiene lugar entre el nacimiento (p. ej., es perinatal) y las 4 semanas de edad.

En algunas modalidades, las composiciones de rAAV se formulan para reducir la acumulación de partículas de AAV en la composición, particularmente cuando hay concentraciones altas de rAAV presentes (p. ej.,  $\sim 10^{13}$  CG/mL o más). Los métodos para reducir la acumulación de rAAV son muy conocidos en la técnica e incluyen, por ejemplo, la adición de tensioactivos, el ajuste del pH, el ajuste de la concentración de sal, etc. (Véase, p. ej., Wright F R, et ál., *Molecular Therapy* (2005) 12, 171-178, cuyo contenido se incorpora a la presente mediante referencia).

La formulación de excipientes y soluciones portadoras farmacéuticamente aceptables es muy conocida por los expertos en la técnica, así como lo es el desarrollo de regímenes de dosificación y tratamiento adecuados para usar las composiciones particulares descritas en la presente en una variedad de regímenes de tratamiento.

Normalmente, estas formulaciones pueden contener al menos alrededor de 0,1 % del compuesto activo o más, aunque el porcentaje del o de los ingredientes activos, desde luego, puede variar y puede ser convenientemente entre alrededor de 1 o 2 % y alrededor de 70 % u 80 % o más del peso o volumen de la formulación total. Naturalmente, la cantidad de compuesto activo en cada composición con utilidad terapéutica puede prepararse de tal manera que se obtenga una dosificación adecuada en cualquier dosis unitaria determinada del compuesto. Los factores tales como la solubilidad, biodisponibilidad, semivida biológica, vía de administración, vida útil del producto, así como otras consideraciones farmacológicas, serán contemplados por el experto en la técnica de preparación de esas formulaciones farmacéuticas, y de esa manera, pueden ser convenientes diversas dosificaciones y regímenes de tratamiento.

En determinadas circunstancias, será conveniente suministrar las construcciones terapéuticas basadas en rAAV en las composiciones farmacéuticas formuladas de manera adecuada descritas en la

presente, ya sea por vía subcutánea, intrapancreática, intranasal, parenteral, intravenosa, intramuscular, intratecal u oral, intraperitoneal o mediante inhalación. En algunas modalidades, las modalidades de administración tal como se describe en las pat. estadounidenses n.º 5,543,158; 5,641,515 y 5,399,363 (cada una de las cuales se incorpora en su totalidad a la presente mediante referencia) pueden usarse para suministrar los rAAV. En algunas modalidades, un modo preferido de administración es mediante inyección en la vena porta.

Las formas farmacéuticas adecuadas para el uso en inyectables incluyen soluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles. También pueden prepararse dispersiones en glicerol, polietilenglicoles líquidos y mezclas de estos, así como en aceites. En condiciones de almacenamiento y uso normales, estas preparaciones contienen un conservante para evitar el crecimiento de microorganismos. En muchos casos, la forma es estéril y fluida en la medida en que exista la posibilidad de incluirla en una jeringa fácilmente. Debe ser estable en las condiciones de fabricación y almacenamiento y se debe conservar contra la acción contaminante de microorganismos, tales como bacterias y hongos. El portador puede ser un solvente o medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (p. ej., glicerol, propilenglicol y polietilenglicol líquido. y similares), mezclas adecuadas de estos y/o aceites vegetales. La fluidez adecuada puede mantenerse, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, mediante el mantenimiento del tamaño de partícula necesario en el caso de la dispersión y mediante el uso de tensioactivos. La prevención de la acción de los microorganismos puede ser provocada por diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro de sodio. La absorción prolongada de las composiciones inyectables puede ser provocada por el uso en las composiciones de agentes que retrasen la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Para la administración de una solución acuosa inyectable, por ejemplo, la solución puede amortiguarse de manera adecuada, de ser necesario, y el diluyente líquido puede volverse isotónico en primer lugar con una cantidad suficiente de solución salina o glucosa. Estas soluciones acuosas particulares son especialmente adecuadas para la administración por vía intravenosa, intramuscular, subcutánea e intraperitoneal. Con relación a esto, los expertos en la técnica conocerán un medio acuoso estéril que se puede emplear. Por ejemplo, se puede disolver una dosificación en 1 mL de solución de NaCl isotónica y se puede añadir a 1000 mL de fluido de hipodermocclisis o inyectar en el sitio de infusión propuesto (véase, por ejemplo, «Remington's Pharmaceutical Sciences», 15.<sup>a</sup> edición, páginas 1035-1038 y 1570-1580). Se producirá necesariamente algo de variación en la dosificación según la condición del hospedador. La persona responsable de la administración puede, en cualquier caso,

determinar la dosis adecuada para el hospedador individual.

Las soluciones inyectables estériles se preparan mediante la incorporación de rAAV activos en la cantidad necesaria en un solvente adecuado con varios de los ingredientes enumerados en la presente, según sea necesario, tras lo cual sigue la esterilización filtrada. Por lo general, las dispersiones se preparan mediante la incorporación de los diversos ingredientes activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico y los demás ingredientes necesarios de los enumerados anteriormente. En el caso de los polvos estériles para la preparación de soluciones inyectables estériles, los métodos de preparación preferidos son las técnicas de secado al vacío y liofilización que proporcionan un polvo del ingrediente activo y cualquier ingrediente adicional deseado a partir de una solución filtrada previamente en condiciones estériles de estos.

Las composiciones de rAAV descritas en la presente también pueden formularse en una forma neutra o de sal. Las sales farmacéuticamente aceptables incluyen las sales de adición de ácido (formadas con los grupos amino libres de la proteína) y que se forman con ácidos inorgánicos tales como, por ejemplo, ácidos clorhídricos o fosfóricos, o ácidos orgánicos tales como acético, oxálico, tartárico, mandélico, y similares. Las sales formadas con los grupos carboxilo libres también pueden derivar de bases inorgánicas tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio o férricos, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína y similares. A partir de la formulación, las soluciones se administrarán de manera compatible con la formulación de dosificación y en una cantidad que presente eficacia terapéutica. Las formulaciones se administran fácilmente en diversas formas de dosificación tales como soluciones inyectables, cápsulas de liberación de fármaco, y similares.

Tal como se usa en la presente, «portador» incluye todos y cada uno de los solventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antifúngicos y antibacterianos, agentes de retraso de la absorción e isotónicos, agentes de amortiguación, soluciones portadoras, suspensiones, coloides, y similares. El uso de esos medios y agentes para las sustancias activas farmacéuticas es muy conocido en la técnica. También se pueden incorporar en las composiciones ingredientes activos complementarios. La expresión «farmacéuticamente aceptable» hace referencia a entidades moleculares y composiciones que no producen una reacción alérgica o adversa similar cuando se administran a un hospedador.

Los vehículos de suministro tales como liposomas, nanocápsulas, micropartículas, microesferas, partículas lipídicas, vesículas, y similares, pueden usarse para la introducción de las composiciones de la presente descripción en células hospedadoras adecuadas. En particular, los transgenes suministrados mediante el vector rAAV pueden formularse para el suministro encapsulados en una partícula lipídica, un liposoma, una vesícula, una nanoesfera o una nanopartícula o similar.

Esas formulaciones pueden preferirse para la introducción de formulaciones farmacéuticamente aceptables de los ácidos nucleicos o las construcciones de rAAV descritas en la presente. La formación y el uso de liposomas son conocidos en términos generales por los expertos en la técnica. Recientemente, se desarrollaron liposomas con estabilidad sérica y semividas en circulación mejoradas (pat. estadounidense n.º 5,741,516). Además, se han descrito diversos métodos de preparaciones de liposomas y similares a liposomas como posibles portadores de fármacos (pat. estadounidenses n.º 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5,795,587).

Los liposomas se han usado con éxito con una serie de tipos celulares que normalmente son resistentes a la transfección mediante otros procedimientos. Además, los liposomas carecen de las restricciones de longitud de ADN que son típicas de los sistemas de suministro de base viral. Los liposomas se han usado con eficacia para introducir genes, fármacos, agentes radioterapéuticos, virus, factores de transcripción y efectores alostéricos en una variedad de líneas celulares cultivadas y animales. Además, se han completado diversos ensayos clínicos exitosos que analizaron la eficacia del suministro de fármacos mediada por liposomas.

Los liposomas se forman a partir de fosfolípidos que se dispersan en un medio acuoso y forman de manera espontánea vesículas de bicapas concéntricas multilamelares (también denominadas «vesículas multilamelares» (MLV, por sus siglas en inglés)). Por lo general, las MLV tienen diámetros de entre 25 nm y 4 µm. La sonicación de las MLV produce la formación de vesículas unilamelares pequeñas (SUV, por sus siglas en inglés) con diámetros en el intervalo de entre 200 y 500 Å, que contienen una solución acuosa en el núcleo.

De manera alternativa, se pueden usar formulaciones de nanocápsulas del rAAV. Por lo general, las nanocápsulas pueden captar sustancias de manera estable y reproducible. Para evitar efectos secundarios ocasionados por una sobrecarga polimérica intracelular, se deben diseñar partículas ultrafinas (con un tamaño de alrededor de 0,1 µm) mediante el uso de polímeros pasibles de degradación in vivo. Las nanopartículas biodegradables de polialquil-cianoacrilato que cumplen con estos requisitos se consideran para el uso.

Además de los métodos de suministro descritos anteriormente, también se contemplan las técnicas que siguen como métodos alternativos de suministro de las composiciones de rAAV a un hospedador. La sonoforesis (es decir, ultrasonido) se ha usado y descrito en la pat. estadounidense n.º 5,656,016 como un dispositivo para mejorar la velocidad y eficacia de la permeación del fármaco hacia y a través del sistema circulatorio. Otras alternativas de suministro de fármacos contempladas son la inyección intraósea (pat. estadounidense n.º 5,779,708), dispositivos de microchip (pat. estadounidense n.º 5,797,898), formulaciones oftálmicas (Bourlais et ál., 1998), matrices transdérmicas (pat. estadounidenses n.º 5,770,219 y 5,783,208) y el suministro con control de respuesta (pat.

estadounidense n.º 5,697,899).

## Ejemplos

### *Ejemplo 1. Diseño de proteínas de dedos de cinc para regular al alza la expresión del gen SCN1A*

5

Se identificaron regiones homólogas entre las secuencias promotoras de SCN1A de seres humanos (células HEK293T) y de ratones (células HEPG2) mediante la alineación de secuencias alrededor de los dos sitios prominentes de inicio de la transcripción identificados en el conjunto de datos de sec. RIKEN CAGE para cada especie (**FIG. 1**). Existe una secuencia altamente conservada entre los seres humanos (HEK) y ratones (HEPG2) en la región promotora proximal del SCN1A (**FIG. 2**). Se diseñaron tres ZFP consistieron en seis dedos para la unión a regiones de 15-22 nucleótidos en superposición en la región promotora proximal del SCN1A a través del ensamblaje de módulos de uno y dos dedos con una especificidad de unión a ADN predefinida (**FIG. 3**). Se diseñaron tres ZFP (ZFP1-ZFP3) que consistieron en seis dedos, en cada caso diseñadas para unirse a las secuencias altamente conservadas en superposición identificadas en la **FIG. 3**. Cada dedo se diseña para unirse a una región de tres bases (triplete) en la región altamente conservada del promotor proximal del SCN1A.

15

La ZFP-1 reconoce regiones de tres bases individuales (tripletes de ADN indicados en rojo separados por «•») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 2), tal como se muestra en la **FIG. 4A**. Cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-1 se une a secuencias de tres nucleótidos, tal como se muestra en la **FIG. 4B**. Las secuencias de aminoácidos de los seis dedos de ZFP-1 (SEQ ID NO: 17-22) se muestran en la **FIG. 4C**; los enlazadores entre los dedos están resaltados para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP). Las secuencias de nucleótidos de los seis dedos de ZFP-1 (SEQ ID NO: 11-16) se muestran en la **FIG. 4D**.

20

25 **Tabla 1.** Hélices de reconocimiento de la ZFP-1 que se dirige al SCN1A

|                                        | Secuencia de aminoácidos   | Secuencia de nucleótidos                 |
|----------------------------------------|----------------------------|------------------------------------------|
| Hélice de reconocimiento 1 de la ZFP-1 | QRGNLVR<br>(SEQ ID NO: 17) | CAGCGGGGAAACCTGGTGAGG<br>(SEQ ID NO: 11) |
| Hélice de reconocimiento 2 de la ZFP-1 | LSFNLTR<br>(SEQ ID NO: 18) | CTGAGCTTCAATCTAACCAGA<br>(SEQ ID NO: 12) |
| Hélice de reconocimiento 3 de la ZFP-1 | RSDNLTR<br>(SEQ ID NO: 19) | CGGAGTGACAACTTAACGCGG<br>(SEQ ID NO: 13) |

|                                        | Secuencia de aminoácidos   | Secuencia de nucleótidos                 |
|----------------------------------------|----------------------------|------------------------------------------|
| Hélice de reconocimiento 4 de la ZFP-1 | DRSHLAR<br>(SEQ ID NO: 20) | GACCGGTCTCACCTTGCCCGA<br>(SEQ ID NO: 14) |
| Hélice de reconocimiento 5 de la ZFP-1 | QKAHLTA<br>(SEQ ID NO: 21) | CAGAAGGCCCATTTGACTGCC<br>(SEQ ID NO: 15) |
| Hélice de reconocimiento 6 de la ZFP-1 | RSDNLTR<br>(SEQ ID NO: 22) | CGGTCGGACAACCTCACACGC<br>(SEQ ID NO: 16) |

La ZFP-2 reconoce regiones de tres bases individuales (tripletes de ADN indicados en rojo separados por «•») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 3), tal como se muestra en la **FIG. 5A**. Cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-2 se une a secuencias de tres nucleótidos, tal como se muestra en la **FIG. 5B**. Las secuencias de aminoácidos de los seis dedos de ZFP-2 (SEQ ID NO: 29-34) se muestran en la **FIG. 5C**; los enlazadores entre los dedos están resaltados para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP). Las secuencias de nucleótidos de los seis dedos de ZFP-1 (SEQ ID NO: 23-28) se muestran en la **FIG. 5D**.

**Tabla 2.** Hélices de reconocimiento de la ZFP-2 que se dirige al SCN1A

|                                        | Secuencia de aminoácidos   | Secuencia de nucleótidos                 |
|----------------------------------------|----------------------------|------------------------------------------|
| Hélice de reconocimiento 1 de la ZFP-2 | RSSNLTR<br>(SEQ ID NO: 29) | CGAAGTTCCAACCTGACACGG<br>(SEQ ID NO: 23) |
| Hélice de reconocimiento 2 de la ZFP-2 | DKRTLIR<br>(SEQ ID NO: 30) | GACAAGCGGACCTTAATCCGC<br>(SEQ ID NO: 24) |
| Hélice de reconocimiento 3 de la ZFP-2 | QRGNLVR<br>(SEQ ID NO: 31) | CAGCGGGGAAATCTAGTGCGA<br>(SEQ ID NO: 25) |
| Hélice de reconocimiento 4 de la ZFP-2 | LSFNLTR<br>(SEQ ID NO: 32) | CTGAGCTTCAACTTGACTCGT<br>(SEQ ID NO: 26) |
| Hélice de reconocimiento 5 de la ZFP-2 | RSDNLTR<br>(SEQ ID NO: 33) | CGGAGTGACAATCTTACGAGA<br>(SEQ ID NO: 27) |
| Hélice de reconocimiento 6 de la ZFP-2 | DRSHLAR<br>(SEQ ID NO: 34) | GACCGGAGCCACTTAGCCAGG<br>(SEQ ID NO: 28) |

La ZFP-3 reconoce regiones de tres bases individuales (tripletes de ADN indicados en rojo separados por «•») dentro de la región promotora proximal del gen SCN1A (SEQ ID NO: 4), tal como se muestra en la **FIG. 6A**. Cada hélice de reconocimiento (siete aminoácidos) de los dedos 1 a 6 para la ZFP-3 se une a secuencias de tres nucleótidos, tal como se muestra en la **FIG. 6B**. Las secuencias de

aminoácidos de los seis dedos de ZFP-3 (SEQ ID NO: 41-46) se muestran en la **FIG. 6C**; los enlazadores entre los dedos están resaltados para designar secuencias enlazadoras canónicas (TGEKP) y no canónicas (TGSQKP). Las secuencias de nucleótidos de los seis dedos de ZFP-1 (SEQ ID NO: 35-40) se muestran en la **FIG. 6D**.

5 **Tabla 3.** Hélices de reconocimiento de la ZFP-3 que se dirige al SCN1A

|                                        | Secuencia de aminoácidos   | Secuencia de nucleótidos                 |
|----------------------------------------|----------------------------|------------------------------------------|
| Hélice de reconocimiento 1 de la ZFP-3 | DRSALAR<br>(SEQ ID NO: 41) | GACCGGAGCGCGCTGGCACGG<br>(SEQ ID NO: 35) |
| Hélice de reconocimiento 2 de la ZFP-3 | RSDNLTR<br>(SEQ ID NO: 42) | CGAAGTGACAACTTAACGCGC<br>(SEQ ID NO: 36) |
| Hélice de reconocimiento 3 de la ZFP-3 | QSGDLTR<br>(SEQ ID NO: 43) | CAGTCAGGGGACCTCACTCGT<br>(SEQ ID NO: 37) |
| Hélice de reconocimiento 4 de la ZFP-3 | VRQTLKQ<br>(SEQ ID NO: 44) | GTACGACAGACGCTTAAACAA (SEQ ID NO: 38)    |
| Hélice de reconocimiento 5 de la ZFP-3 | AAGNLTR<br>(SEQ ID NO: 45) | GCCGCTGGTAACTTGACACGA<br>(SEQ ID NO: 39) |
| Hélice de reconocimiento 6 de la ZFP-3 | RSDNLTR<br>(SEQ ID NO: 46) | AGATCTGATAATCTAACGCGT (SEQ ID NO: 40)    |

Otras ZFP diseñadas para dirigirse a secuencias conservadas en la región promotora proximal del gen SCN1A comprenderán cinco o seis dominios de dedos cada una y se unirán a regiones de 15-22 nucleótidos altamente conservadas entre el SCN1A de seres humanos y ratones.

**Tabla 4.** Proteínas de dedos de cinc dirigidas al SCN1A

|       | Secuencia de aminoácidos                                                                                                                                                                                                        | Secuencia de nucleótidos                                                                                                                                                                                                                                                                                                                                                                 |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ZFP-1 | RPFQCRICMRNFSQRGNLVR<br>HIRTHTGEKPFACDICGKKFA<br>LSFNLTRHTKIHTGSQKPFQC<br>RICMRNFSRSDNLTRHIRTHT<br>GEKPFACDICGKKFADRSHLA<br>RHTKIHTGSQKPFQCRICMRN<br>FSQKAHLTAHIRTHTGEKPFA<br>CDICGRKFARSDNLTRHTKIH<br>LRQKD<br>(SEQ ID NO: 57) | CGACCATTCCAGTGTCTGAATCTGCATGCGCAACT<br>TCAGCCAGCGGGGAAACCTGGTGAGGCATATCC<br>GCACCCACACGGGAGAGAAGCCTTTTGCCTGCG<br>ATATTTGTGGAAAGAAGTTTGCTCTGAGCTTCAAT<br>CTAACCAGACACACCAAGATTCATACTGGGTCCC<br>AGAAACCGTTCCAGTGTAGGATATGCATGAGGAA<br>TTTCTCTCGGAGTGACAACTTAACGCGGCATATA<br>AGGACGCACACAGGTGAAAAACCATTTGCATGCG<br>ACATCTGTGGCAAAAAGTTTGCGGACCGGTCTCA<br>CCTTGCCCGACACACAAAAATCCATACCGGCAGT |

|       | Secuencia de aminoácidos                                                                                                                                                                                                           | Secuencia de nucleótidos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|       |                                                                                                                                                                                                                                    | CAAAAGCCCTTTCAATGTGCGATTTGCATGCGAA<br>ACTTCTCACAGAAGGCCCATTTGACTGCCCATAT<br>TCGTACTCATACTGGCGAGAAACCTTTGCTTGC<br>GATATATGTGGTCGTAAGTTTGCACGGTCGGACA<br>ACCTCACACGCCACACTAAGATACACCTGCGGCA<br>GAAGGAC<br>(SEQ ID NO: 58)                                                                                                                                                                                                                                                                                                                                                                                            |
| ZFP-2 | RPFQCRICMRNFSRSSLNLR<br>HIRHTHTGEKPFACDICGKKFA<br>DKRTLIRHTKIHTGSQKPFQC<br>RICMRNFSQRGNLVRHIRHTHT<br>GEKPFACDICGKKFALSFNLT<br>RHTKIHTGSQKPFQCRICMRN<br>FSRSDNLTRHIRHTHTGEKPFA<br>CDICGRKFADRSHLARHTKIH<br>LRQKD<br>(SEQ ID NO: 59) | CGACCATTCCAGTGTGCAATCTGCATGCGCAACT<br>TCAGCCGAAGTTCCAACCTGACACGGCATATCCG<br>CACCCACACGGGAGAGAAGCCTTTTGCCTGCGA<br>TATTTGTGGAAAGAAGTTTGTCTGACAAGCGGACC<br>TTAATCCGCCACACCAAGATTCATACTGGGTCCC<br>AGAAACCGTTCCAGTGTAGGATATGCATGAGGAA<br>TTTCTCTCAGCGGGGAAATCTAGTGCGACATATA<br>AGGACGCACACAGGTGAAAAACCATTTGCATGCG<br>ACATCTGTGGCAAAAAGTTTGCCTGAGCTTCAA<br>CTTGACTCGTCACACAAAAATCCATACCGGCAGT<br>CAAAAGCCCTTTCAATGTGCGATTTGCATGCGAA<br>ACTTCTCACGGAGTGACAATCTTACGAGACATAT<br>TCGTACTCATACTGGCGAGAAACCTTTGCTTGC<br>GATATATGTGGTCGTAAGTTTGCAGACCGGAGCC<br>ACTTAGCCAGGCACACTAAGATACACCTGCGGCA<br>GAAGGAC<br>(SEQ ID NO: 60) |
| ZFP-3 | RPFQCRICMRNFSDRSALAR<br>HIRHTHTGEKPFACDICGKKFA<br>RSDNLTRHTKIHTGSQKPFQC<br>RICMRNFSQSGDLTRHIRHTHT<br>GEKPFACDICGKKFAVRQTLK<br>QHTKIHTGSQKPFQCRICMR<br>NFSAAAGNLTRHIRHTHTGEKPF<br>ACDICGRKFARSDNLTRHTKI<br>HLRQKD                   | CGACCATTCCAGTGTGCAATCTGCATGCGCAACT<br>TCAGCGACCGGAGCGCGCTGGCACGGCATATCC<br>GCACCCACACGGGAGAGAAGCCTTTTGCCTGCG<br>ATATTTGTGGAAAGAAGTTTGTCTCGAAGTGACAA<br>CTTAACGCGCCACACCAAGATTCATACTGGGTCC<br>CAGAAACCGTTCCAGTGTAGGATATGCATGAGGA<br>ATTTCTCTCAGTCAGGGGACCTCACTCGTCATAT<br>AAGGACGCACACAGGTGAAAAACCATTTGCATGC<br>GACATCTGTGGCAAAAAGTTTGCGGTACGACAGA                                                                                                                                                                                                                                                                  |

|  | Secuencia de aminoácidos | Secuencia de nucleótidos                                                                                                                                                                                                                                      |
|--|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | (SEQ ID NO: 61)          | CGCTTAAACAACACACAAAAATCCATACCGGCAG<br>TCAAAGCCCTTTCAATGTCGCATTTGCATGCGA<br>AACTTCTCAGCCGCTGGTAACTTGACACGACATA<br>TTCGTACTCATACTGGCGAGAAACCTTTGCTTG<br>CGATATATGTGGTCGTAAGTTTGCAAGATCTGAT<br>AATCTAACGCGTCACACTAAGATACACCTGCGGC<br>AGAAGGAC<br>(SEQ ID NO: 62) |

*Ejemplo 2. Las ZFP aumentan la expresión del gen SCN1A en células humanas*

Para examinar la capacidad de ZFP1-ZFP3 de regular al alza la transcripción de SCN1A, los dominios de unión a ADN de ZFP1-ZFP3 se fusionan con un dominio activador de la transcripción, tripartito, potente e híbrido de VP64, p53 y RTA (VPR) para formar un transactivador quimérico. El dominio activador de la fusión VPR actúa para captar complejos reguladores de la transcripción y aumentar la capacidad de acceder a la cromatina y ayuda a lograr niveles altos de expresión del gen. Por lo tanto, el dominio de ZFP dirigirá el activador de VPR a la secuencia altamente conservada en la región promotora proximal para aumentar la expresión del gen SCN1A.

Se transfectaron plásmidos de expresión que codificaban proteínas de fusión VPR-ZFP1, VPR-ZFP2 y/o VPR-ZFP3 a través de la transfección transitoria en células HEK293 y se midió la expresión del gen SCN1A mediante qRT-PCR (mediante el uso de la expresión de TBP como referencia para la normalización). Las fusiones VPR-ZFP comprenden la ZFP1, ZFP2 y/o ZFP3 fusionadas con VPR. La transfección de tres construcciones para la regulación múltiple, que contenían dominios de unión a ADN de ZFP1, ZFP2 y ZFP3 fusionados cada uno a VPR, produjo un aumento de 45 veces en la expresión del gen SCN1A en relación con las células no transfectadas, lo cual indica que los transactivadores quiméricos VPR-ZFP son capaces de aumentar la expresión del gen SCN1A mediante la unión en la región promotora proximal del gen (**FIG. 7**).

Se transfectan proteínas de fusión VPR-[ZFP1-ZFP3] —así como proteínas de fusión VPR-ZFP en las que el dominio de unión a ADN de la ZFP se está diseñando actualmente— en células HeLa y HEPG2, que tienen en ambos casos niveles bajos de expresión de SCN1A. Las proteínas de fusión VPR-ZFP contienen versiones simples así como combinaciones de múltiples dominios de unión a ADN de ZFP que se fusionan con el transactivador VPR. La expresión del gen SCN1A se mide mediante qRT-PCR para determinar si estas fusiones VPR-ZFP son capaces de aumentar la expresión del gen. Los

candidatos más auspiciosos como fusiones VPR-ZFP se evalúan en neuronas corticales primarias de ratones luego del suministro mediante virus adenoasociados (AAV) de las proteínas de fusión para determinar su capacidad de aumentar la expresión de *SCN1A*.

La especificidad de los dominios de ZFP se optimiza de manera adicional mediante el uso de un sistema de selección de un híbrido bacteriano (véase, p. ej., Meng, et ál., «Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases», *Nat Biotechnol*, 2008) para identificar ZFP ideales dentro de un banco aleatorizado en el que los residuos importantes en la unión a ADN se someten a variaciones. Las ZFP seleccionadas se fusionarán con dominios transactivadores VPR, tanto de manera individual como en combinaciones de múltiples ZFP y se transfectarán en células HEK293, HeLa y HEPG2, así como neuronas corticales primarias de ratones para identificar los dominios ZFP candidatos que más aumentan la expresión del gen *SCN1A* luego del análisis de qRT-PCR.

### *Ejemplo 3. Generación de series de transactivadores ZFP<sup>SCN1A</sup> con distintas potencias*

Las ZFP más eficaces en la regulación al alza de la expresión del gen *SCN1A* del Ejemplo 2 se fusionan con una serie de dominios de transactivación humanos (p. ej., Rta, p65, Hsf1, etc.) con un gradiente de potencias anticipadas para identificar una configuración que logre una regulación al alza de 2 veces de la expresión del gen *SCN1A* con un intervalo de multiplicidades de infección (MOI, por sus siglas en inglés) de los AAV. Se infectaron neuronas corticales primarias de ratones normales y *SCN1A*<sup>+/-</sup> con vectores de AAV que expresaban transactivadores de fusión ZFP<sup>SCN1A</sup>. Los niveles de expresión de la proteína Na<sub>v</sub>1.1 se evalúan mediante inmunoelectrotransferencia mediante el uso de qPCR. Se usan neuronas primarias tratadas durante 8 horas con TGF-α como control positivo dado que este tratamiento aumenta la expresión de la proteína Na<sub>v</sub>1.1 en ~6 a 8 veces (Chen et ál., 2015, *Neuroinflammation* 12: 126). También se evalúan los cambios en los niveles de expresión de otros genes de subunidad alfa de Na<sub>v</sub> para mostrar la especificidad de la transactivación de ZFP<sup>SCN1A</sup>. Se usa la inmunofluorescencia para determinar si la expresión de Na<sub>v</sub>1.1 permanece restringida para las interneuronas gabaérgicas a través de una doble tinción inmunofluorescente con anticuerpos para ZFP<sup>SCN1A</sup> (etiqueta de HA) y marcadores específicos para neuronas gabaérgicas (p. ej., parvalbúmina<sup>+</sup> o somatostatina<sup>+</sup>) o marcadores neuronales universales (p. ej., NeuN, TUBIII y/o Map2). La especificidad de ZFP<sup>SCN1A</sup> para la transactivación del gen *SCN1A* también se evalúa mediante ChIP-Seq y RNA-Seq para localizar los sitios de unión del genoma y el perfil transcriptómico resultante generado a partir de la transferencia génica.

*Ejemplo 4. Organización de las histonas y panorama epigenómico del promotor del SCN1A en inhibidores gabaérgicos para guiar el diseño de transactivadores SCN1A-ZFP dependientes de la actividad del promotor*

5 La capacidad de las ZFP de unirse a dianas genómicas depende de la accesibilidad de la secuencia diana (p. ej., la presencia de una región sin nucleosoma). Esta necesidad de accesibilidad del ADN se aprovecha para diseñar transactivadores de ZFP que solamente sean funcionales en un subconjunto de tipos de células en función de la presencia de accesibilidad de las secuencias diana de ADN. Se logra una restricción adicional en la actividad del tipo celular a través del uso de promotores con especificidad de tejido para la expresión de los transactivadores de ZFP. Se ha demostrado que promotores pequeños de los genes neuropéptido Y y somatostatina del pez globo (*Takifugu rubripes*) impulsan la expresión transgénica altamente específica en interneuronas inhibitoras corticales y del hipocampo en el contexto de vectores de AAV y lentivirus. En algunas modalidades, la combinación de la restricción de la transcripción basada en AAV de las ZFP específicas de SCN1A que son sensibles a la accesibilidad del ADN produce una regulación al alza altamente específica en la expresión de la proteína Nav1.1 en las interneuronas inhibitoras en todo el cerebro. Esta estrategia reguladora doble minimizará los efectos secundarios que puedan surgir de la expresión ectópica de la proteína Nav1.1 en células en las que normalmente no se expresa.

La estructura del nucleosoma y el panorama epigenético del promotor de SCN1A se analizan en neuronas inhibitoras gabaérgicas y excitadoras glutamérgicas de ratones y seres humanos. Esta información se usa para diseñar transactivadores de ZFP restringidos a neuronas inhibitoras gabaérgicas al dirigirse a secuencias a las que se puede acceder solamente alrededor del locus del SCN1A en este tipo de célula.

Las neuronas inhibitoras gabaérgicas de ratones transgénicos que expresan TdTomato con presencia de un promotor GAD67 y neuronas excitadoras glutamatéricas con presencia de GFP generadas mediante la cruce de ratones Emx1-IRES-Cre con ROSA26/stop/EGFP se aíslan mediante la clasificación celular activada por fluorescencia (FACS, por sus siglas en inglés). Se generan neuronas excitadoras y gabaérgicas humanas a partir de células madre pluripotentes inducidas (iPS, por sus siglas en inglés) y esto se confirma mediante inmunotinción y RT-PCR en busca de marcadores específicos para estos tipos de células, así como la actividad electrofisiológica. Las regiones genómicas accesibles alrededor del promotor del SCN1A en poblaciones neuronales de ratones y humanas se caracterizan mediante el uso del ensayo de cromatina accesible por transposasa (ATAC-Seq).

Se diseñan transactivadores de ZFP<sup>SCN1A</sup> que reconocen secuencias accesibles solamente en neuronas

gabaérgicas, en función de la accesibilidad diferencial de la cromatina de las regiones genómicas alrededor del promotor del SCN1A en neuronas inhibitoras y excitadoras. Se genera una serie de fusiones de transactivadores ZFP-VPR candidatos para dirigirse a diferentes regiones accesibles del SCN1A, donde se espera que la unión de los transactivadores regule al alza de manera potente la expresión de Nav1.1 en regiones inhibitoras, y revele cualquier expresión inducida no deseada de la expresión de Nav1.1 en neuronas excitadoras.

Se realizan estudios de expresión en neuronas derivadas de iPS humanas y neuronas primarias SCN1A<sup>+/-</sup> de ratones que modelan el síndrome de Dravet para determinar si los transactivadores ZFP<sup>SCN1A</sup> diseñados para reconocer secuencias de ADN a las que solamente se puede acceder en las neuronas inhibitoras proporcionan la especificidad necesaria cuando se expresan a partir de vectores de AAV con presencia de un promotor específico de interneuronas inhibitoras o de sinapsina-1 de ser humano de carácter panneuronal. Los niveles de expresión de Nav1.1 se miden mediante qRT-PCR, inmunoelectrotransferencia. y doble inmunofluorescencia con marcadores de tipos neuronales específicos para neuronas gabaérgicas inhibitoras (p. ej., GABA<sup>+</sup>, GAD65/67<sup>+</sup>, somatostatina y/o parvalbúmina) y glutamérgicas excitadoras (p. ej., Cux1<sup>+</sup>, FoxG1<sup>+</sup>, receptores GABA<sub>A</sub>, GABA<sup>-</sup>). La especificidad con respecto al tipo celular de los transactivadores ZFP<sup>SCN1A</sup> que se diseñan para dirigirse a diferentes secuencias diana en el promotor del SCN1A de ratones y seres humanos como la estructura de cromatina y la secuencia de ADN dentro de regiones sinténicas difiere de una especie a otra. Los controles en estos experimentos incluyen cultivos neuronales infectados con vectores de AAV similares que codifican GFP, ZFP sin dominios de transactivación o transactivadores sin dominios de unión a ADN de ZFP.

Se incorporan sitios de unión de microARN (miARN) dentro de la región no traducida 3' (3' UTR) de los transactivadores ZFP<sup>SCN1A</sup> que se restringen a tipos de células en los que tiene lugar una expresión no deseada (p. ej., las neuronas excitadoras glutamatérgicas). Esta estrategia se utilizó anteriormente para restringir la expresión de transgenes con suministro mediante AAV (Xie, et ál., «MicroRNA-regulated, systematically delivered rAAV9: a step closer to CNS-restricted transgene expression», *Mol. Ther.* 2011). Las diferencias en el perfil de expresión de miARN de las neuronas inhibitoras gabaérgicas y otros tipos de células se determinan mediante la secuenciación de ARN de tamaño pequeño.

*Ejemplo 5. Evaluación del potencial de la terapia génica con AAV-ZFP<sup>SCN1A</sup> para corregir los déficits de circulación de sodio en interneuronas gabaérgicas generadas mediante iPS derivadas de pacientes*

Una etapa fundamental hacia el desarrollo de uno o más transactivadores ZFP<sup>SCN1A</sup> para el síndrome

de Dravet consiste en demostrar que estos transactivadores artificiales tienen la función deseada en las neuronas humanas. Con este fin, se obtienen células iPS de pacientes con Dravet (n= 4-6) y pacientes sin Dravet (n= 4). Se representa un trasfondo genético sin Dravet dentro de estas células, obviando la necesidad de manipular artificialmente la expresión génica y, por lo tanto, han surgido células iPS como la línea celular del estado de la técnica para la investigación biomédica. Se utiliza la tecnología de edición genómica CRISPR-Cas9 para crear líneas celulares isogénicas mediante la reparación de la mutación genética en el SCN1A en la secuencia natural, o mediante la introducción de una mutación asociada a Dravet en un alelo normal dentro de una línea celular de control. De ese modo, las líneas isogénicas eliminan la variabilidad natural que surge a partir de la comparación de líneas celulares de diferentes sujetos humanos y, por lo tanto, son útiles para confirmar y aumentar los fenotipos específicos de la enfermedad. Se usa un protocolo de diferenciación de neuronas inhibitoras establecido y un producto en fase de desarrollo para diferenciar las líneas de células iPS en interneuronas inhibitoras gabaérgicas del prosencéfalo.

Las neuronas inhibitoras derivadas de pacientes con Dravet presentan circulaciones reducidas de sodio y deficiencias en la activación de los potenciales de acción tal como se ha determinado mediante mediciones electrofisiológicas mediante el pinzamiento zonal de células enteras. Se realizan mediciones similares para confirmar que las neuronas derivadas de Dravet descritas en la presente recapitulen estos fenotipos asociados a la enfermedad. Los defectos en la circulación de sodio tienen lugar en las neuronas inhibitoras, pero no en las excitadoras, de los pacientes con Dravet (Sun et al.) y, por lo tanto, solo se utilizan neuronas inhibitoras en la presente descripción. Los defectos en los canales de sodio inducidos por mutaciones en las neuronas inhibitoras derivadas de pacientes con Dravet pueden subsanarse mediante la expresión ectópica de SCN1A en estado natural (ref 20). Por lo tanto, los métodos descritos en la presente descripción son adecuados para evaluar la eficacia de los transactivadores ZFP<sup>SCN1A</sup> a la hora de restaurar la fisiología y el funcionamiento de los canales de sodio naturales en el contexto del síndrome de Dravet.

Se infectan cultivos de neuronas inhibitoras gabaérgicas con vectores de AAV que codifican transactivadores de ZFP<sup>SCN1A</sup> con presencia de promotores de neuronas universales o específicos de neuronas inhibitoras. Los cambios en los niveles de expresión de Nav1.1 se evalúan mediante inmunoelectrotransferencia. La restauración de las circulaciones de sodio funcionales en las neuronas inhibitoras se evalúa a través del pinzamiento zonal de células enteras de las células no transfectadas en comparación con las transfectadas. La unión de los transactivadores ZFP<sup>SCN1A</sup> a través del genoma en todas las neuronas inhibitoras derivadas de los pacientes se analiza mediante ChIP-seq y se correlaciona con cualquier cambio transcriptómico identificado detectado mediante RNA-seq. Los controles en estos experimentos son cultivos neuronales infectados con vectores de AAV similares que

codifican GFP, ZFP sin dominios transactivadores VPR, y dominios transactivadores VPR sin dominios de unión a ADN de ZFP.

*Ejemplo 6. Evaluación del potencial terapéutico de la intervención de AAV-ZFP<sup>SCN1A</sup> en diferentes edades y vías de suministro en ratones con SCN1A*

El amplio tropismo de los AAV es una propiedad fundamental para las aplicaciones de terapias génicas para genes de expresión amplia, pero puede convertirse en un problema significativo cuando un transgén de interés se expresa con especificidad con respecto al tipo de célula. Esto se ha solucionado en gran medida para los tejidos principales del organismo tales como el hígado, los músculos y el corazón a través del uso de promotores con especificidad de tejido tales como la proteína de unión a tiroxina (TBP, por sus siglas en inglés), la creatina cinasa y la troponina T, respectivamente. Se puede superponer un nivel de control adicional en promotores con especificidad de tejido para lograr un grado más alto de desdireccionamiento de los tejidos específicos mediante la incorporación de múltiples copias de sitios de unión para microARN de gran abundancia en esos tejidos, tales como miR-122 en el hígado y miR-1 en el músculo esquelético. El serotipo AAV-PHP.B recientemente descrito es excepcionalmente eficiente para la transferencia génica en el SNC después del suministro sistémico, donde transduce una amplia variedad de tipos de células. Además, su tropismo con respecto a los tejidos periféricos es casi en todos los casos igual de amplio que el de AAV9. La meta de una estrategia de terapia génica para el síndrome de Dravet es restaurar la expresión de Nav1.1 en las interneuronas inhibitoras gabaérgicas de manera exclusiva y al mismo tiempo prevenir efectos perjudiciales de la expresión ectópica en otras neuronas y otras partes. Se ha demostrado que los vectores de AAV y lentivirus que codifican la GFP con promotores pequeños (<2,8 kb) derivados de los genes neuropéptido Y (fnPY) y somatostatina (fSST) del pez globo (*Takifugu rubripes*) impulsan la expresión específica de las neuronas inhibitoras en el cerebro de los ratones a partir de la inyección intracraneal. Los vectores AAV-PHP.B que portan estos promotores que impulsan la expresión de la GFP se comparan con vectores de control en los que la expresión del transgén es impulsada por el promotor CAG potente y ubicuo y el promotor MeCP2 de ratones mínimo relativamente más débil. La especificidad de los vectores AAV-PHP.B-GFP con promotores fSST y fnYP con respecto a las neuronas inhibitoras gabaérgicas se estudia a partir del suministro al SNC mediante administración sistémica en ratones de 6 semanas de edad (vena de la cola) y el día 1 después del nacimiento (retroorbital), el suministro de LCR en recién nacidos y por último inyecciones unilaterales dirigidas al giro dentado (GD) (**Tabla 5**). La eficiencia de la transferencia génica en el SNC varía considerablemente según la vía de suministro y dado que se tratan ratones *Scn1a*<sup>+/-</sup> de diferentes edad

se realiza un análisis amplio para establecer el punto de referencia para la eficacia de la transducción neuronal y la especificidad de los promotores con respecto a las interneuronas inhibitoras gabaérgicas en todo el SNC para cada vía de suministro. Se ha demostrado anteriormente que los vectores de AAV que impulsan la expresión de GFP a partir de los promotores fNYP y fSST cortos son altamente específicos con respecto a las interneuronas inhibitoras en el hipocampo después de la inyección directa. Los vectores AAV-PHP.B de la presente descripción se validan de la misma manera que en estudios posteriores, donde se evalúa el impacto terapéutico de la restauración de Na<sub>v</sub>1.1 en las neuronas inhibitoras en la formación hipocampal de los ratones *Scn1a*<sup>+/-</sup>, que se ubica específicamente en el giro dentado y el revestimiento interno de la capa de células granulosas (los fundamentos se explicarán más adelante). Los experimentos se realizan en ratones 129SvJ/C57BL/6 generados en UMMS mediante la cruce de ratones 129SvJ con C57BL/6 obtenidos a través de Jackson Laboratories (Bar Harbor, ME). Los ratones se sacrifican un mes después de la inyección y se recolecta el cerebro y la médula espinal para el análisis histológico de la especificidad y eficiencia de la transducción mediante el uso de doble inmunofluorescencia con anticuerpos para marcadores con especificidad celular y GFP. La eficiencia y especificidad de la transferencia génica con respecto a las interneuronas inhibitoras gabaérgicas se evalúa en todo el cerebro y la médula espinal mediante doble tinción inmunofluorescente con anticuerpos para la descarboxilasa del ácido glutámico (GAD; marcador para las neuronas gabaérgicas) y GFP. Además, se evalúa la especificidad preferencial de los promotores y/o AAV-PHP.B con respecto a subconjuntos de interneuronas inhibitoras que expresan somatostatina (SST), parvalbúmina (PV), calretinina (CR), péptido intestinal vasoactivo (VIP, por sus siglas en inglés) o neuropéptido Y (NPY) mediante el uso de anticuerpos específicos para esas proteínas y GFP. Se recolecta el hígado, el corazón y músculo esquelético de los ratones tratados mediante administración sistémica e ICV para evaluar la expresión de GFP a nivel histológico y se utilizan inmunoelectrotransferencias para determinar la posibilidad de la expresión ectópica en los tejidos periféricos.

**Tabla 5. Grupos experimentales**

| Vía de suministro | Cantidad de ratones por cohorte |                    |                    |                    |
|-------------------|---------------------------------|--------------------|--------------------|--------------------|
|                   | Sistémica                       |                    | ICV                | IC                 |
| Edad              | 6 semanas*                      | PND1 #             | PND1 #             | 8 semanas*         |
| Dosis (vg)        | 2x10 <sup>12</sup>              | 4x10 <sup>11</sup> | 4x10 <sup>10</sup> | 1x10 <sup>10</sup> |
| AAV-fSST-GFP      | 6                               | 6-8                | 6-8                | 4                  |
| AAV-fNYP-GFP      | 6                               | 6-8                | 6-8                | 4                  |

|                |   |     |     |   |
|----------------|---|-----|-----|---|
| AAV-CAG-GFP    | 6 | 6-8 | 6-8 | 4 |
| AAV-MeCP2-GFP  | 6 | 6-8 | 6-8 | - |
| Vehículo (PBS) | 2 | -   | -   |   |

\* Los grupos están compuestos por la misma cantidad de ratones de ambos sexos.

# Una camada inyectada por vector

**Abreviaturas:** ICV: inyección intracerebroventricular; IC: inyección intracraneal; PND1: día 1 después del nacimiento

5

Se administraron inyecciones bilaterales a ratones *Scn1a*<sup>+/-</sup> de seis semanas de edad en el giro dentado de vectores AAV-PHP.B que codificaban diferentes proteínas transactivadoras ZFP<sup>Scn1a</sup>, una construcción con el dominio de activación ZFP<sup>Scn1a</sup>, pero sin el dominio de unión a ADN para controlar el impacto solamente del activador, o el mismo volumen de solución salina amortiguada con fosfato (PBS, por sus siglas en inglés) (n=3 machos + 3 hembras/grupo). Los vectores de AAV monocatenarios usados en estos experimentos también portan un casete de IRES-GFP anterógrado con respecto al ADNc de ZFP<sup>Scn1a</sup> para facilitar la identificación de las células transducidas. Se evalúan al menos dos transactivadores ZFP<sup>Scn1a</sup>, que pueden tener una activación más amplia en una variedad de neuronas, así como los dos transactivadores ZFP<sup>SCN1A</sup> restringidos a neuronas inhibitoras gabaérgicas más auspiciosos descritos anteriormente. Un mes después de la inyección, se recolecta el cerebro y el hipocampo de un hemisferio cerebral se disecciona para evaluar los niveles de expresión de las proteínas ZFP<sup>Scn1a</sup>, Na<sub>v</sub>1.1, Na<sub>v</sub>1.3, GAD65, GAD67 mediante inmunoelectrotransferencia mediante el uso de beta-actina o tubulina como controles de carga. El otro hemisferio cerebral se examina mediante estudios histológicos mediante el uso de cortes cerebrales en serie (10 µm) para analizar el % de interneuronas inhibitoras transducidas en el giro dentado y el estrato interno de la capa de células granulosas mediante doble tinción inmunofluorescente con anticuerpos para GAD y GFP, o GAD y una etiqueta epitópica incluida en todas las proteínas ZFP<sup>Scn1a</sup> (etiqueta de HA o myc). Asimismo, se determina el porcentaje de neuronas con presencia de GAD que expresan Na<sub>v</sub>1.1 y Na<sub>v</sub>1.3 para demostrar la restauración de los patrones normales de expresión de los canales de sodio. Además de la detección por inmunofluorescencia de la expresión de las proteínas Na<sub>v</sub>1.1 y Na<sub>v</sub>1.3, se evalúan los cambios en los niveles de ARNm en las interneuronas gabaérgicas mediante el uso de sondas de RNAscope para Na<sub>v</sub>1.1, Na<sub>v</sub>1.3, ZFP<sup>Scn1a</sup> y GAD. RNAscope es una técnica de hibridación *in situ* de sensibilidad alta para analizar los niveles de ARNm en las neuronas del cerebro. La combinación de estas dos estrategias para evaluar los cambios en los niveles de Na<sub>v</sub>1.1 como resultado de la expresión de ZFP<sup>Scn1a</sup> proporciona una comprensión integral de cómo se logran los cambios en las interneuronas mediante la estrategia de terapia génica de la presente descripción.

La eficacia terapéutica de la terapia génica con AAV-PHP.B-ZFP<sup>Scn1a</sup> se analiza en ratones *Scn1a*<sup>+/-</sup> de ambos sexos comenzando el día 1 después del nacimiento, o a las 6 semanas de edad a través de la vena de la cola. Los controles incluyen ratones tratados con un vector de AAV que codifica una proteína similar a una ZFP sin el dominio de unión a ADN de la ZFP, así como ratones *Scn1a*<sup>+/-</sup> de la misma edad no tratados y compañeros de camada en estado natural (n=15 machos y 15 hembras por grupo). Se sacrifica un subconjunto de ratones en cada grupo (n=3 machos y 3 hembras) a las 12 semanas de edad para evaluar la eficiencia de la transferencia génica con respecto a las interneuronas gabaérgicas mediante inmunoelectrotransferencia así como inmunofluorescencia con anticuerpos para GAD (y otros marcadores de tipos neuronales específicos, p. ej., GAD65, GAD67) y la ZFP, y la restauración de la expresión de Nav1.1 en esas células en todo el cerebro y la médula espinal. Además, se evalúa la expresión ectópica de la ZFP, junto con la expresión de Nav1.1 en los tejidos periféricos. El otro subconjunto de animales en cada grupo (n=24) se usa para estudiar el impacto en la supervivencia (hasta 1 año de edad), el desempeño motor y el comportamiento, que se evalúa cada dos meses entre los 2-12 meses de edad. La función motora y la coordinación se evalúan mediante el uso de las pruebas de rotarod con aceleración y cruce de la viga, dado que los ratones *Scn1a*<sup>+/-</sup> presentan una coordinación deficiente de los miembros anteriores y posteriores para el PND21. Además, se utilizan pruebas del comportamiento en las que los ratones *Scn1a*<sup>+/-</sup> muestran deficiencias en el desempeño, que incluyen: campo abierto, laberinto en cruz elevado, nidificación, enterramiento de canicas, y laberinto de Barnes para evaluar el aprendizaje espacial y la memoria, que parecen estar gravemente afectados en los ratones *Scn1a*<sup>+/-</sup>. Las convulsiones espontáneas características de los pacientes con el síndrome de Dravet también se manifiestan en los ratones *Scn1a*<sup>+/-</sup> y la frecuencia aumenta con la edad y la temperatura corporal. Además, se produce la muerte repentina y prematura de los ratones *Scn1a*<sup>+/-</sup> después de convulsiones tónico-clónicas. Por lo tanto, se utiliza una videovigilancia continua las 24 h a los 2, 6 y 12 meses de edad para evaluar la frecuencia y duración de las convulsiones. Los estudios de interacción social mediante el uso de lecturas de preferencia de cámara en respuesta a objetos y olores nuevos y se tiene en cuenta al ratón si se detecta un cambio significativo en los resultados primarios medidos en las pruebas descritas anteriormente. Se recolecta el cerebro, la médula espinal y los órganos periféricos para su evaluación en cuanto a los criterios de valoración humanitarios del experimento para realizar los análisis moleculares e histológicos descritos anteriormente.

#### *Ejemplo 7. Los sistemas dCas9 y ZFP aumentan la expresión del gen SCN1A en células humanas*

Para examinar la capacidad de ZFP1-ZFP3 de regular al alza la transcripción de SCN1A, los dominios

de unión a ADN de ZFP1-ZFP3 se fusionan con un dominio activador de la transcripción híbrido de VP64, p53 y RTA (VPR) para formar un transactivador quimérico. El dominio activador de la fusión VPR actúa para captar complejos reguladores de la transcripción y aumentar la capacidad de acceder a la cromatina y ayuda a lograr niveles altos de expresión del gen. Por lo tanto, el dominio de ZFP dirigirá el activador de VPR a la secuencia altamente conservada en la región promotora proximal para aumentar la expresión del gen SCN1A.

Además, para examinar la capacidad de los sistemas dCas9 de dirigirse al SCN1A para regular al alza la transcripción del SCN1A, se formaron complejos a partir de tres ARN guía dirigidos al SCN1A con proteína dCas9.

Las células HEK293T se transfectaron transitoriamente con una de las siguientes condiciones de experimentación: (1) una construcción VPR-ZFP1; (2) una construcción VPR-ZFP2; (3) una construcción VPR-ZFP3; (4) las tres construcciones VPR-ZFP1, VPR-ZFP2 y VPR-ZFP3; (5) una construcción dCas9-VPR y un ARN guía 1 de SCN1A; (6) una construcción dCas9-VPR y un ARN guía 2 de SCN1A; (7) una construcción dCas9-VPR y un ARN guía 3 de SCN1A; (8) una construcción dCas9-VPR y los tres ARN guía 1 de SCN1A, ARN guía 2 de SCN1A y ARN guía 3 de SCN1A; y (9) una construcción dCas9-VPR sin ningún ARN guía (control). La expresión del gen SCN1A se midió mediante qRT-PCR. La intensidad de la activación de SCN1A se normalizó con respecto al experimento de control (la construcción dCas9-VPR sin ningún ARN guía).

Todas las condiciones de experimentación evaluadas produjeron aumentos en la activación génica de SCN1A en relación con el experimento de control (FIG. 8). Estos datos demuestran que las proteínas de dedos de cinc descritas en este ejemplo y a lo largo de la presente descripción son capaces de dirigirse al SCN1A para influir en su expresión génica. Estos datos demuestran además que las secuencias de ARN guía de este ejemplo (SEQ ID NO: 83-94) son capaces de dirigir la dCas9 al SCN1A con el fin de influir en su expresión génica.

**Tabla 6.** Ácidos nucleicos guía dirigidos al SCN1A (secuencia separadora en negrita)

|                 | Secuencia de nucleótidos (ADN)                                                                                                          | Secuencia de nucleótidos (ARN)                                                                                                          |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Guía 1 de SCN1A | <b>GAGGTACCATAGAGTGAGGCGGT</b><br>TTTAGAGCTAGAAATAGCAAGTTAA<br>AATAAGGCTAGTCCGTTATCAACTT<br>GAAAAAGTGGCACCGAGTCGGTGC<br>(SEQ ID NO: 83) | <b>GAGGUACCAUAGAGUGAGGCGGUUU</b><br>UAGAGCUAGAAAUAGCAAGUUAAAAU<br>AAGGCUAGUCCGUUAUCAACUUGAAA<br>AAGUGGCACCGAGUCGGUGC<br>(SEQ ID NO: 84) |
| Guía 2 de SCN1A | <b>ACCGAGGCGAGGATGAAGCCGAG</b><br>GTTTTAGAGCTAGAAATAGCAAGTT                                                                             | <b>ACCGAGGCGAGGAUGAAGCCGAGGU</b><br>UUUAGAGCUAGAAAUAGCAAGUUAAA                                                                          |

|                    | Secuencia de nucleótidos (ADN)                                                                                                                     | Secuencia de nucleótidos (ARN)                                                                                                            |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
|                    | AAAATAAGGCTAGTCCGTTATCAAC<br>TTGAAAAAGTGGCACCGAGTCGGT<br>GC (SEQ ID NO: 87)                                                                        | AUAAGGCUAGUCCGUUAUCAACUUGA<br>AAAAGUGGCACCGAGUCGGUGC (SEQ<br>ID NO: 88)                                                                   |
| Guía 3 de<br>SCN1A | <b>ACCGAAGCCGAGAGGATACTGCA</b><br><b>GGTTTTAGAGCTAGAAATAGCAAG</b><br>TTAAAATAAGGCTAGTCCGTTATCA<br>ACTTGAAAAAGTGGCACCGAGTCG<br>GTGC (SEQ ID NO: 91) | <b>ACCGAAGCCGAGAGGAUACUGCAGG</b><br>UUUUAGAGCUAGAAAUAGCAAGUUA<br>AAUAAGGCUAGUCCGUUAUCAACUUG<br>AAAAAGUGGCACCGAGUCGGUGC<br>(SEQ ID NO: 92) |