

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 970 541**

51 Int. Cl.:

C12N 5/071 (2010.01)

C12N 9/22 (2006.01)

C12N 15/11 (2006.01)

G01N 33/68 (2006.01)

G01N 21/64 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.03.2020 PCT/US2020/023131**

87 Fecha y número de publicación internacional: **24.09.2020 WO20190932**

96 Fecha de presentación y número de la solicitud europea: **17.03.2020 E 20718950 (7)**

97 Fecha y número de publicación de la concesión europea: **13.12.2023 EP 3768826**

54 Título: **Plataforma de cribado CRISPR/CAS para identificar modificadores genéticos de la siembra o agregación de tau**

30 Prioridad:
18.03.2019 US 201962820086 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
29.05.2024

73 Titular/es:
REGENERON PHARMACEUTICALS, INC.
(100.0%)
777 Old Saw Mill River Road
Tarrytown, NY 10591, US

72 Inventor/es:
PRISSETTE, MARINE;
KOSS, MATTHEW;
FURY, WEN y
ZAMBROWICZ, BRIAN

74 Agente/Representante:
VALLEJO LÓPEZ, Juan Pedro

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 970 541 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Plataforma de cribado CRISPR/CAS para identificar modificadores genéticos de la siembra o agregación de tau

5 **Antecedentes**

La agregación o fibrilación anormal de proteínas es una característica definitoria de muchas enfermedades, incluyendo notablemente varias enfermedades neurodegenerativas, tales como la enfermedad de Alzheimer (EA), la enfermedad de Parkinson (EP), la demencia frontotemporal (DFT), la esclerosis lateral amiotrófica (ELA), la encefalopatía traumática crónica (ETC) y la enfermedad de Creutzfeldt-Jakob (ECJ), entre otras. En muchas de estas enfermedades, la fibrilación de ciertas proteínas en agregados insolubles no solo es un sello distintivo de la enfermedad, sino que también se ha implicado como factor causante de neurotoxicidad. Además, estas enfermedades se caracterizan por la propagación de la patología de agregados a través del sistema nervioso central siguiendo patrones estereotípicos, un proceso que se correlaciona con la progresión de la enfermedad. La identificación de genes y vías genéticas que modifican los procesos de agregación anormal de proteínas, o la propagación de agregados de célula a célula, son, por lo tanto, de gran valor para comprender mejor la etiología de las enfermedades neurodegenerativas, así como para idear estrategias para intervención terapéutica.

Nathaniel et al. (Elucidating Cellular Trafficking Pathways Controlling Prion-like Spread of Tau Aggregation Using CRISPR interference Screens, *Molecular Biology of the Cell* 27(25): 3947 (Resumen P887)) describen una plataforma basada en la interferencia CRISPR (CRISPRi) para identificar genes relacionados con la agregación de tau. CRISPRi usa una Cas9 inactivada fusionada a un represor transcripcional para atenuar genes diana.

Similarmente, Chen et al. (CRISPR-based genome-wide screen identifies protein homeostasis factors that control Tau aggregation and propagation in human cells, *Molecular Biology of the Cell* 27(25): 3947 (Resumen M264)) desvelan un cribado de todo el genoma basado en CRISPR que identificó genes que desempeñan una función en la agregación de tau.

Konermann et al. (*Nature* 517: 583-588, 2015) describen la activación CRISPR (CRISPRa), que usa una Cas9 inactivada y un ARN_g para dirigir un activador transcripcional a un gen diana y activar así su expresión. El activador transcripcional puede estar fusionado a Cas9 inactivada, o ser reclutado por un ARN_g modificado.

Kampmann (*ACS Chem. Biol.* 13(2): 406-416, 2018) revisa los cribados genéticos de CRISPRi y CRISPRa para células de mamífero. Una biblioteca de ARN_g puede transducirse en células diana usando una biblioteca lentivírica con baja multiplicidad de infección, lo que en su mayoría resulta en una integración estable de una sola construcción de expresión de ARN_g por célula.

El informe anual de investigación del año fiscal 2016 del proyecto de investigación de la Universidad de Tokio titulado "Identificación de reguladores genéticos de la agregación intracelular mediante cribado CRISPR de todo el genoma" describe un método de cribado FACS para agregados intracelulares de tau. La formación de agregados intracelulares detectables se consiguió aplicando fibras de tau recombinante marcadas con fluorescencia a líneas celulares que expresaban tau de forma constitutiva. Los agregados se detectaron perforando las membranas celulares para liberar tau monomérica, luego realizando FACS en las células perforadas.

Holmes et al. (*Proc. Natl. Acad. Sci. U.S.A.* 111(41): E4376-E4385, 2014) describen un ensayo de biosensor de citometría de flujo basada en FRET para la siembra de tau. Se desarrolló una línea celular biosensora que expresaba de forma estable un dominio de repetición de tau fusionado con proteína fluorescente cian (CFP) o proteína fluorescente amarilla (YFP). La exposición de la línea celular biosensora a semillas de tau provocó la agregación de tau, generando una señal de FRET.

50 **Sumario**

En el presente documento se proporcionan métodos de detección de modificadores genéticos de la agregación de tau.

Algunos de dichos métodos (CRISPRn) comprenden: (a) proporcionar una población de células que comprende una proteína Cas, un primer dominio de repetición de tau unido a un primer indicador, y un segundo dominio de repetición de tau unido a un segundo indicador; en donde las células son células de mamífero, en donde el primer indicador y el segundo indicador son proteínas fluorescentes, y en donde el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET); (b) introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos dirigidos a una pluralidad de genes; (c) cultivar la población de células para permitir la edición y expansión del genoma, en donde la pluralidad de ARN guía únicos forman complejos con la proteína Cas, y la proteína Cas corta la pluralidad de genes dando como resultado la inactivación de la función génica para producir una población de células modificada genéticamente; (d) poner en contacto la población de células modificada genéticamente con un agente de siembra de tau para producir una población de células sembrada, en donde la etapa (d) comprende cultivar la población de células modificada

genéticamente en presencia de medio acondicionado obtenido de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, en donde el medio acondicionado se obtuvo después de estar en células positivas a la agregación de tau confluentes durante 1 a 7 días; (e) cultivar la población sembrada de células para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población sembrada de células para producir una población de células positiva a la agregación; y (f) determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células modificada genéticamente en la etapa (c), en donde el enriquecimiento de un ARN guía en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células cultivada en la etapa (c) indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía potencia la agregación de tau.

En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau es un dominio de repetición de tau humano. En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende una mutación pro-agregación. Opcionalmente, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende una mutación de tau P301S.

En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende un dominio de cuatro repeticiones de tau. En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende SEQ ID NO: 11. En algunos de dichos métodos, el primer dominio de repetición de tau y el segundo dominio de repetición de tau son el mismo. En algunos de dichos métodos, el primer dominio de repetición de tau y el segundo dominio de repetición de tau son el mismo y cada uno comprende el dominio de cuatro repeticiones de tau que comprende una mutación de tau P301S.

En algunos de dichos métodos, el primer indicador es la proteína fluorescente cian (CFP) y el segundo indicador es la proteína fluorescente amarilla (YFP).

En algunos de dichos métodos, la proteína Cas es una proteína Cas9. Opcionalmente, la proteína Cas es Cas9 de *Streptococcus pyogenes*. Opcionalmente, la proteína Cas comprende SEQ ID NO: 21. Opcionalmente, la proteína Cas está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 22.

En algunos de dichos métodos, la proteína Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se expresan de forma estable en la población de células. En algunos de dichos métodos, los ácidos nucleicos que codifican la proteína Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se integran genómicamente en la población de células.

En algunos de dichos métodos, las células son células humanas. Opcionalmente, las células son células HEK293T.

En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce a una concentración seleccionada de forma que la mayoría de las células reciban solo uno de los ARN guía únicos. En algunos de dichos métodos, la pluralidad de ARN guía únicos se dirige a 100 o más genes, 1000 o más genes, o 10000 o más genes. En algunos de dichos métodos, la biblioteca es una biblioteca de todo el genoma.

En algunos de dichos métodos, en cada uno de la pluralidad de genes diana se seleccionan de media una pluralidad de secuencias diana. Opcionalmente, se seleccionan al menos tres secuencias diana de media en cada uno de los genes seleccionados. Opcionalmente, se seleccionan de aproximadamente tres a aproximadamente seis secuencias diana (por ejemplo, aproximadamente tres, aproximadamente cuatro o aproximadamente seis) de media en cada uno de la pluralidad de genes diana.

En algunos de dichos métodos, cada ARN guía se dirige a un exón constitutivo. Opcionalmente, cada ARN guía se dirige a un exón constitutivo en 5'. En algunos de dichos métodos, cada ARN guía se dirige a un primer exón, un segundo exón o un tercer exón.

En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción vírica. Opcionalmente, cada uno de la pluralidad de ARN guía únicos está en un vector vírico separado. Opcionalmente, la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción lentivírica. En algunos de dichos métodos, la población de células se infecta con una multiplicidad de infección inferior a aproximadamente 0,3.

En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce en la población de células junto con un marcador de selección, y la etapa (b) comprende además la selección de células que comprenden el marcador de selección. Opcionalmente, el marcador de selección confiere resistencia a un fármaco. Opcionalmente, el marcador de selección confiere resistencia a la puromicina o a la zeocina. Opcionalmente, el marcador de selección selecciona de neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y

blastocidina S desaminasa. Opcionalmente, el marcador de selección se selecciona de neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa, blastocidina S desaminasa y proteína de resistencia a la bleomicina.

5 En algunos de dichos métodos, la población de células en las que se introduce la pluralidad de ARN guía únicos en la etapa (b) comprende más de aproximadamente 300 células por ARN guía único.

En algunos de dichos métodos, la etapa (c) es de aproximadamente 3 a aproximadamente 9 días. Opcionalmente, la etapa (c) es de aproximadamente 6 días.

10 En algunos de dichos métodos, el medio acondicionado se obtuvo después de estar en células confluentes positivas a la agregación de tau durante aproximadamente 4 días. Opcionalmente, la etapa (d) comprende cultivar la población de células modificada genéticamente en aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio recién preparado. En algunos de dichos métodos, la población de células modificada genéticamente no se cocultiva con las células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado.

15 En algunos de dichos métodos, la etapa (e) es de aproximadamente 2 a aproximadamente 6 días. Opcionalmente, la etapa (e) es de aproximadamente 4 días. En algunos de dichos métodos, la población de células positiva a la agregación en la etapa (e) se identifica por citometría de flujo.

20 En algunos de dichos métodos, la abundancia se determina mediante secuenciación de nueva generación. En algunos de dichos métodos, un ARN guía se considera enriquecido si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c).

25 En algunos de dichos métodos, la etapa (f) comprende determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c) en un primer momento en la etapa (c) y/o un segundo momento en la etapa (c). Opcionalmente, el primer momento en la etapa (c) se encuentra en un primer pase del cultivo de la población de células, y el segundo momento se encuentra en la mitad del cultivo de la población de células para permitir la edición y expansión del genoma. Opcionalmente, el primer momento en la etapa (c) es después de aproximadamente tres días de cultivo, y el segundo momento en la etapa (c) es después de aproximadamente seis días de cultivo.

30 En algunos de dichos métodos, un gen se considera un modificador genético de la agregación de tau, en donde la inactivación del gen potencia la agregación de tau, si: (1) la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c) tanto en el primer momento en la etapa (c) como en el segundo momento en la etapa (c); y/o (2) la abundancia de al menos dos ARN guía únicos dirigidos al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación de la etapa (e) en relación con la población de células cultivada de la etapa (c) en el primer momento de la etapa (c) o en el segundo momento de la etapa (c).

35 En algunos de dichos métodos, se siguen las siguientes etapas en la etapa (f) para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación del gen potencia la agregación de tau: (1) identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación producida en la etapa (e); (2) calcular la probabilidad aleatoria de que los ARN guía identificados en la etapa (f)(1) estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, en donde x es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b), en donde m es la variedad de ARN guía únicos identificados en la etapa (f)(1), en donde n es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b) que se dirigen al gen, y en donde n' es la variedad de ARN guía únicos identificados en la etapa (f)(1) que se dirigen al gen; (3) calcular las puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (f)(1), en donde la puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación producida en la etapa (e) dividida por la abundancia relativa del ARN guía en la población de células cultivada en la etapa (c), y en donde la abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos; y (4) seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral.

40 Algunos de dichos métodos (CRISPRa) comprenden: (a) proporcionar una población de células que comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, una proteína adaptadora quimérica que comprende una proteína adaptadora fusionada a uno o más dominios de activación transcripcional, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador; en donde las células son células de mamífero, en donde el primer indicador y el segundo indicador son proteínas fluorescentes, y en donde el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET); (b)

introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos dirigidos a una pluralidad de genes; (c) cultivar la población de células para permitir la activación transcripcional y la expansión, en donde la pluralidad de ARN guía únicos forman complejos con la proteína Cas quimérica y la proteína adaptadora quimérica, y los complejos activan la transcripción de la pluralidad de genes dando lugar a un aumento de la expresión génica para producir una población de células modificada genéticamente; (d) poner en contacto la población de células modificada genéticamente con un agente de siembra de tau para producir una población de células sembrada, en donde la etapa (d) comprende cultivar la población de células modificada genéticamente en presencia de medio acondicionado obtenido de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau está presente de forma estable en un estado agregado, en donde el medio acondicionado se obtuvo después de estar en células positivas a la agregación de tau confluentes durante 1 a 7 días; (e) cultivar la población de células sembrada para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación; y (f) determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células modificada genéticamente en la etapa (c), en donde el enriquecimiento de un ARN guía en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células cultivada en la etapa (c) indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía potencia la agregación de tau.

En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau es un dominio de repetición de tau humano. En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende una mutación pro-agregación. Opcionalmente, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende una mutación de tau P301S.

En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende un dominio de cuatro repeticiones de tau. En algunos de dichos métodos, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprende SEQ ID NO: 11. En algunos de dichos métodos, el primer dominio de repetición de tau y el segundo dominio de repetición de tau son el mismo. En algunos de dichos métodos, el primer dominio de repetición de tau y el segundo dominio de repetición de tau son el mismo y cada uno comprende el dominio de cuatro repeticiones de tau que comprende una mutación de tau P301S.

En algunos de dichos métodos, el primer indicador es la proteína fluorescente cian (CFP) y el segundo indicador es la proteína fluorescente amarilla (YFP).

En algunos de dichos métodos, la proteína Cas es una proteína Cas9. Opcionalmente, la proteína Cas es Cas9 de *Streptococcus pyogenes*. En algunos de dichos métodos, la proteína Cas quimérica comprende la proteína Cas inactiva por nucleasa fusionada a un dominio de activación transcripcional VP64, opcionalmente en donde la proteína Cas quimérica comprende de extremo N a extremo C: la proteína Cas inactiva por nucleasa; una señal de localización nuclear; y el dominio activador transcripcional VP64. En algunos de dichos métodos, la proteína adaptadora es una proteína de la cubierta de MS2, y en donde uno o más dominios de activación transcripcional en la proteína adaptadora quimérica comprenden un dominio de activación transcripcional p65 y un dominio de activación transcripcional HSF1, opcionalmente en donde la proteína adaptadora quimérica comprende de extremo N a extremo C: la proteína de la cubierta de MS2; una señal de localización nuclear; el dominio de activación transcripcional p65; y el dominio de activación transcripcional HSF1. En algunos de dichos métodos, la proteína Cas quimérica comprende SEQ ID NO: 36, opcionalmente en donde la proteína Cas quimérica está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 38. En algunos de dichos métodos, la proteína adaptadora quimérica comprende SEQ ID NO: 37, opcionalmente en donde la proteína adaptadora quimérica está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 39.

En algunos de dichos métodos, la proteína Cas quimérica, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se expresan de forma estable en la población de células. En algunos de dichos métodos, los ácidos nucleicos que codifican la proteína Cas quimérica, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se integran genómicamente en la población de células.

En algunos de dichos métodos, las células son células humanas. Opcionalmente, las células son células HEK293T.

En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce a una concentración seleccionada de forma que la mayoría de las células reciban solo uno de los ARN guía únicos. En algunos de dichos métodos, la pluralidad de ARN guía únicos se dirige a 100 o más genes, 1000 o más genes, o 10000 o más genes. En algunos de dichos métodos, la biblioteca es una biblioteca de todo el genoma.

En algunos de dichos métodos, en cada uno de la pluralidad de genes diana se seleccionan de media una pluralidad de secuencias diana. Opcionalmente, se seleccionan al menos tres secuencias diana de media en cada uno de los

genes seleccionados. Opcionalmente, se seleccionan de aproximadamente tres a aproximadamente seis secuencias diana (por ejemplo, aproximadamente tres, aproximadamente cuatro o aproximadamente seis) de media en cada uno de la pluralidad de genes diana. Opcionalmente, se seleccionan aproximadamente tres secuencias diana de media en cada uno de la pluralidad de genes diana.

5 En algunos de dichos métodos, cada ARN guía se dirige a una secuencia diana de ARN guía dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción. En algunos de dichos métodos, cada ARN guía comprende uno o más elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica. Opcionalmente, cada ARN guía comprende dos elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica. Opcionalmente, un primer elemento de unión al adaptador se encuentra dentro de un primer bucle de cada uno de los uno o más ARN guía, y un segundo elemento de unión al adaptador se encuentra dentro de un segundo bucle de cada uno de los uno o más ARN guía. Opcionalmente, el elemento de unión al adaptador comprende la secuencia establecida en SEQ ID NO: 33. En algunos de dichos métodos, en donde cada uno de uno o más ARN guía es un único ARN guía que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de ARN CRISPR transactivador (ARNtracr), y el primer bucle es el tetrabucle correspondiente a los residuos 13-16 de SEQ ID NO: 17, y el segundo bucle es el tallo-bucle 2 correspondiente a los residuos 53-56 de SEQ ID NO: 17.

20 En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción vírica. Opcionalmente, cada uno de la pluralidad de ARN guía únicos está en un vector vírico separado. Opcionalmente, la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción lentivírica. En algunos de dichos métodos, la población de células se infecta con una multiplicidad de infección inferior a aproximadamente 0,3.

25 En algunos de dichos métodos, la pluralidad de ARN guía únicos se introduce en la población de células junto con un marcador de selección, y la etapa (b) comprende además la selección de células que comprenden el marcador de selección. Opcionalmente, el marcador de selección confiere resistencia a un fármaco. Opcionalmente, el marcador de selección confiere resistencia a la puromicina o a la zeocina. Opcionalmente, el marcador de selección se selecciona de neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y blasticidina S desaminasa. Opcionalmente, el marcador de selección se selecciona de neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa, blasticidina S desaminasa y proteína de resistencia a la bleomicina.

35 En algunos de dichos métodos, la población de células en las que se introduce la pluralidad de ARN guía únicos en la etapa (b) comprende más de aproximadamente 300 células por ARN guía único.

En algunos de dichos métodos, la etapa (c) es de aproximadamente 3 a aproximadamente 9 días. Opcionalmente, la etapa (c) es de aproximadamente 6 días.

40 En algunos de dichos métodos, el medio acondicionado se obtuvo después de estar en células confluentes positivas a la agregación de tau durante aproximadamente 4 días. Opcionalmente, la etapa (d) comprende cultivar la población de células modificada genéticamente en aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio recién preparado. En algunos de dichos métodos, la población de células modificada genéticamente no se cocultiva con las células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado.

50 En algunos de dichos métodos, la etapa (e) es de aproximadamente 2 a aproximadamente 6 días. Opcionalmente, la etapa (e) es de aproximadamente 4 días. En algunos de dichos métodos, la población de células positiva a la agregación en la etapa (e) se identifica por citometría de flujo.

En algunos de dichos métodos, la abundancia se determina mediante secuenciación de nueva generación. En algunos de dichos métodos, un ARN guía se considera enriquecido si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c).

55 En algunos de dichos métodos, la etapa (f) comprende determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c) en un primer momento en la etapa (c) y/o un segundo momento en la etapa (c). Opcionalmente, el primer momento en la etapa (c) se encuentra en un primer pase del cultivo de la población de células, y el segundo momento se encuentra en la mitad del cultivo de la población de células para permitir la edición y expansión del genoma. Opcionalmente, el primer momento en la etapa (c) es después de aproximadamente tres días de cultivo, y el segundo momento en la etapa (c) es después de aproximadamente seis días de cultivo.

65 En algunos de dichos métodos, un gen se considera un modificador genético de la agregación de tau, en donde la activación transcripcional del gen potencia la agregación de tau, si: (1) la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población

de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c) tanto en el primer momento en la etapa (c) como en el segundo momento en la etapa (c); y/o (2) la abundancia de al menos dos ARN guía únicos dirigidos al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación de la etapa (e) en relación con la población de células cultivada de la etapa (c) en el primer momento de la etapa (c) o en el segundo momento de la etapa (c).

En algunos de dichos métodos, se siguen las siguientes etapas en la etapa (f) para identificar un gen como modificador genético de la agregación de tau, en donde la activación transcripcional del gen potencia la agregación de tau: (1) identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación producida en la etapa (e); (2) calcular la probabilidad aleatoria de que los ARN guía identificados en la etapa (f)(1) estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, en donde x es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b), en donde m es la variedad de ARN guía únicos identificados en la etapa (f)(1), en donde n es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b) que se dirigen al gen, y en donde n' es la variedad de ARN guía únicos identificados en la etapa (f)(1) que se dirigen al gen; (3) calcular puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (f)(1), en donde la puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación producida en la etapa (e) dividida por la abundancia relativa del ARN guía en la población de células cultivada en la etapa (c), y donde la abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos; y (4) seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral.

Breve descripción de las figuras

La **Figura 1** (no a escala) muestra un esquema de la isoforma tau 2N4R. Las líneas de biosensores de tau incluyen solo tau4RD-YFP y tau4RD-CFP como transgenes, no la 2N4R completa.

La **Figura 2** muestra un esquema de cómo se monitoriza la formación de agregados mediante transferencia de energía por resonancia de fluorescencia (FRET) en líneas celulares de biosensores tau. La proteína tau^{4RD}-CFP se excita con luz violeta y emite luz azul. La proteína de fusión tau^{4RD}-YFP se excita con luz azul y emite luz amarilla. Si no hay agregación, la excitación con luz violeta no producirá FRET. Si hay agregación de tau, la excitación con luz violeta dará lugar a FRET y emisión de luz amarilla.

La **Figura 3A** muestra la expresión relativa de ARNm de Cas9 en clones de células biosensoras tau^{4RD}-CFP/tau^{4RD}-YFP (TCY) transducidos con construcciones lentivíricas de expresión de Cas9 en relación con el clon Cas9H1, que es un clon TCY de control de bajo rendimiento de expresión de Cas9 que se aisló previamente.

La **Figura 3B** muestra la eficiencia de corte en el locus *PERK* y el locus *SNCA* en los clones Cas9 TCY tres y siete días después de la transducción con ARN_g dirigidos a *PERK* y *SNCA* respectivamente.

La **Figura 4** muestra un esquema de la estrategia para la inactivación de genes diana en la célula biosensora Cas9 TCY usando una biblioteca de ARN_g de CRISPR/Cas9 de todo el genoma.

La **Figura 5** es un esquema que muestra la derivación de subclones de tau^{4RD}-YFP Ag[+] que contienen agregados de tau que se propagan de forma estable cuando las células tau^{4RD}-YFP se siembran con fibrillas de tau^{4RD}. También se muestra una imagen de microscopía de fluorescencia que muestra el subclon con agregados de tau.

La **Figura 6** es un esquema que muestra que el medio acondicionado de subclones de tau^{4RD}-YFP Ag[+] obtenido después de tres días en células confluentes puede proporcionar una fuente de actividad de agregación de tau, mientras que el medio de subclones de tau^{4RD}-YFP Ag[-] no lo hace. El medio acondicionado se aplicó a las células receptoras en forma de 75 % de medio acondicionado y 25 % de medio recién preparado. Se muestran imágenes de análisis de clasificación celular activada por fluorescencia (FACS) para cada uno. El eje x muestra CFP (excitación láser de 405 nm) y el eje y muestra FRET (excitación a partir de la emisión de CFP). El cuadrante superior derecho es FRET[+], el cuadrante inferior derecho es CFP[+] y el cuadrante inferior izquierdo es doble negativo.

La **Figura 7** es un esquema que muestra la estrategia para un cribado con nucleasas CRISPR (CRISPRn) de todo el genoma para identificar genes modificadores que promueven la agregación de tau.

La **Figura 8** es un esquema que muestra los conceptos de abundancia y enriquecimiento para el análisis de secuenciación de nueva generación (NGS) usando el cribado CRISPRn de todo el genoma.

La **Figura 9** muestra un esquema para el cribado secundario de los genes diana 1-14 identificados en el cribado de todo el genoma para genes modificadores que promueven la agregación de tau.

La **Figura 10** es un gráfico que muestra la inducción de FRET por el medio acondicionado de agregados de tau en células biosensoras Cas9 TCY transducidas con construcciones de expresión lentivírica para ARN_g dirigidos a los genes diana 1-14. El cribado secundario confirmó que los genes diana 2 y 8 modulan la susceptibilidad celular a la siembra/agregación de tau.

La **Figura 11** muestra imágenes de análisis FACS de células biosensoras Cas9 TCY transducidas con construcciones de expresión lentivírica para el ARN_g1 del gen diana 2, el ARN_g5 del gen diana 8, un ARN_g no diana y ningún ARN_g. Las células se cultivaron en medio acondicionado o en medio recién preparado. El eje x muestra CFP (excitación láser de 405 nm) y el eje y muestra FRET (excitación a partir de emisión CFP). El cuadrante superior derecho es FRET[+], el cuadrante inferior derecho es CFP[+] y el cuadrante inferior izquierdo es doble negativo. La inactivación de los genes diana 2 u 8 aumenta la formación de agregados de tau en respuesta

al medio acondicionado de agregados de tau, pero no al medio recién preparado.

La **Figura 12** muestra un esquema del cribado secundario en células biosensoras Cas9 TCY transducidas con construcciones de expresión lentivírica para ARN_g dirigidos a los genes diana 2 y 8, incluyendo el análisis de expresión de ARNm, el análisis de expresión de proteínas y el análisis FRET. Se usaron dos ARN_g contra el gen diana 2 (g1 y g3), un ARN_g contra el gen diana 8 (g5) y un ARN_g no diana (g3) como control no diana.

La **Figura 13** muestra la expresión relativa del gen diana 2 y el gen diana 8 en las células biosensoras Cas9 TCY evaluadas mediante qRT-PCR en el día 6 tras la transducción con las construcciones lentivíricas de expresión de ARN_g.

La **Figura 14** muestra la expresión de la proteína 2 (codificada por el gen diana 2) y la proteína 8 (codificada por el gen diana 8) en las células biosensoras Cas9 TCY evaluadas por transferencia Western en el día 13 tras la transducción con las construcciones lentivíricas de expresión de ARN_g.

La **Figura 15** muestra la agregación de tau medida por el porcentaje de células FRET[+] en células biosensoras Cas9 TCY en el día 10 tras la transducción con las construcciones lentivíricas de expresión de ARN_g. No se usó lipofectamina.

La **Figura 16** muestra la expresión del gen diana 2 y el gen diana 8 en los clones de células Cas9 TCY atenuados evaluados mediante transferencia Western.

La **Figura 17** muestra la expresión de agregación de tau en los clones de linfocitos TCY Cas9 atenuados del gen diana 2 y del gen diana 8 como se evalúa mediante FRET.

La **Figura 18** muestra la expresión del gen diana 2 y del gen diana 8 en los clones de células Cas9 TCY atenuados como se evalúa mediante transferencia Western, y la fosforilación de tau en las posiciones S262 y S356 en dichos clones como se evalúa mediante transferencia Western.

La **Figura 19** muestra que el lisado de células enteras del clon18 de tau-YFP Ag[+] puede inducir la agregación de tau y la señal FRET en células biosensoras de tau. Se probaron diferentes cantidades de lisado de células enteras y diferentes condiciones de sonicación para producir el lisado.

La **Figura 20** muestra que el lisado de células enteras del clon18 de tau-YFP Ag[+] puede inducir la agregación de tau y la señal FRET en células biosensoras de tau. Se probaron diferentes cantidades de lisado de células enteras y se probaron diferentes cantidades de lipofectamina.

La **Figura 21** muestra que el lisado de células enteras del clon 18 de tau-YFP Ag[+] puede inducir la agregación de tau y la señal FRET en las células biosensoras de tau, pero el lisado de células enteras de los clones Ag[-] no. Se probaron diferentes cantidades de lisado de células enteras y se probaron diferentes cantidades de lipofectamina.

La **Figura 22** muestra un esquema de la estrategia de un cribado con nucleasas CRISPR (CRISPRn) de todo el genoma para identificar genes modificadores que impidan la agregación de tau.

La **Figura 23** es un gráfico que muestra la identificación de genes con ARN_g enriquecidos de forma única en muestras FRET[-].

La **Figura 24** es un gráfico que muestra la identificación de genes con ARN_g agotados de forma única en muestras FRET[-].

La **Figura 25** muestra un esquema que muestra la estrategia de cribado secundario para confirmar los genes modificadores identificados que impiden la agregación de tau.

La **Figura 26** muestra un esquema que muestra la estrategia para un cribado de activación CRISPR (CRISPRa) de todo el genoma para identificar genes modificadores que impidan la agregación de tau.

La **Figura 27** muestra un esquema que muestra la estrategia para un cribado con nucleasas CRISPR (CRISPRn) de todo el genoma para identificar genes modificadores que promuevan la desagregación de tau.

La **Figura 28** muestra la apertura usada para clasificar las poblaciones de células Ag[+], motitas[+] y Ag[-].

La **Figura 29** muestra un esquema de una estrategia de bloqueo de timidina usada en el cribado de nucleasas CRISPR (CRISPRn) de todo el genoma para identificar genes modificadores que promueven la desagregación de tau.

Definiciones

Los términos "proteína", "polipéptido" y "péptido", usados indistintamente en el presente documento, incluyen formas poliméricas de aminoácidos de cualquier longitud, incluidos aminoácidos codificados y no codificados y aminoácidos modificados o derivados química o bioquímicamente. Los términos también incluyen polímeros que han sido modificados, tales como los polipéptidos que tienen esqueletos peptídicos modificados. El término "dominio" se refiere a cualquier parte de una proteína o polipéptido que tenga una función o estructura particular.

Se dice que las proteínas tienen un "extremo N" y un "extremo C". El término "extremo N" se refiere al inicio de una proteína o polipéptido, terminado por un aminoácido con un grupo amina libre (-NH₂). El término "extremo C" se refiere al final de una cadena de aminoácidos (proteína o polipéptido), terminado por un grupo carboxilo libre (-COOH).

Los términos "ácido nucleico" y "polinucleótido", usados indistintamente en el presente documento, incluyen formas poliméricas de nucleótidos de cualquier longitud, incluidos ribonucleótidos, desoxirribonucleótidos o análogos o versiones modificadas de los mismos. Incluyen ADN o ARN monocatenario, bicatenario y multicatenario, ADN genómico, ADNc, híbridos de ADN-ARN y polímeros que comprenden bases purínicas, bases pirimidínicas u otras bases nucleotídicas naturales, modificadas químicamente, modificadas bioquímicamente, no naturales o derivatizadas.

- Se dice que los ácidos nucleicos tienen "extremos 5'" y "extremos 3'" debido a que los mononucleótidos reaccionan para formar oligonucleótidos de manera que el fosfato 5' de un anillo de pentosa mononucleotídico se une al oxígeno 3' de su vecino en una dirección mediante un enlace fosfodiéster. Un extremo de un oligonucleótido se denomina "extremo 5'" si su fosfato 5' no está unido al oxígeno 3' de un anillo de pentosa mononucleotídico. Un extremo de un oligonucleótido se denomina "extremo 3'" si su oxígeno 3' no está unido a un fosfato 5' de otro anillo de pentosa mononucleotídico. También puede decirse que una secuencia de ácido nucleico, aunque sea interna a un oligonucleótido mayor, tiene extremos 5' y 3'. En una molécula de ADN lineal o circular, los elementos discretos se denominan "aguas arriba" o 5' de los elementos "aguas abajo" o 3'.
- El término "genómicamente integrado" se refiere a un ácido nucleico que se ha introducido en una célula de tal forma que la secuencia de nucleótidos se integra en el genoma de la célula. Puede usarse cualquier protocolo para la incorporación estable de un ácido nucleico en el genoma de una célula.
- El término "vector diana" se refiere a un ácido nucleico recombinante que puede introducirse mediante recombinación homóloga, ligadura mediada por unión de extremos no homólogos o cualquier otro medio de recombinación en una posición diana del genoma de una célula.
- El término "vector vírico" se refiere a un ácido nucleico recombinante que incluye al menos un elemento de origen vírico e incluye elementos suficientes para o permisivos para la encapsidación en una partícula de vector vírico. El vector y/o la partícula pueden utilizarse para transferir ADN, ARN u otros ácidos nucleicos a células *ex vivo* o *in vivo*. Se conocen numerosas formas de vectores víricos.
- El término "natural" incluye entidades que tienen una estructura y/o actividad como se encuentra en un estado o contexto normal (en contraste con mutante, enfermo, alterado, etc.). Los genes y polipéptidos naturales existen a menudo en múltiples formas diferentes (por ejemplo, alelos).
- El término "secuencia endógena" se refiere a una secuencia de ácido nucleico que se produce de forma natural en una célula u organismo. Por ejemplo, una secuencia *MAPT* endógena de una célula u organismo se refiere a una secuencia *MAPT* nativa que se produce de forma natural en el locus *MAPT* de la célula u organismo.
- Las moléculas o secuencias "exógenas" incluyen moléculas o secuencias que no están normalmente presentes en una célula en esa forma. La presencia normal incluye la presencia con respecto a la etapa de desarrollo particular y las condiciones ambientales de la célula. Una molécula o secuencia exógena, por ejemplo, puede incluir una versión mutada de una secuencia endógena correspondiente dentro de la célula, tal como una versión humanizada de la secuencia endógena, o puede incluir una secuencia correspondiente a una secuencia endógena dentro de la célula, pero en una forma diferente (es decir, no dentro de un cromosoma). Por el contrario, las moléculas o secuencias endógenas incluyen moléculas o secuencias que están normalmente presentes en esa forma en una célula particular en una etapa de desarrollo particular en condiciones ambientales particulares.
- El término "heterólogo", cuando se usa en el contexto de un ácido nucleico o una proteína, indica que el ácido nucleico o la proteína comprenden al menos dos segmentos que no se encuentran juntos de forma natural en la misma molécula. Por ejemplo, el término "heterólogo", cuando se usa con referencia a segmentos de un ácido nucleico o segmentos de una proteína, indica que el ácido nucleico o la proteína comprenden dos o más subsecuencias que no se encuentran en la misma relación entre sí (por ejemplo, unidas) en la naturaleza. Como un ejemplo, una región "heteróloga" de un vector de ácido nucleico es un segmento de ácido nucleico dentro o unido a otra molécula de ácido nucleico que no se encuentra en asociación con la otra molécula en la naturaleza. Por ejemplo, una región heteróloga de un vector de ácido nucleico podría incluir una secuencia codificante flanqueada por secuencias que no se encuentran asociadas a la secuencia codificante en la naturaleza. Similarmente, una región "heteróloga" de una proteína es un segmento de aminoácidos dentro o unido a otra molécula peptídica que no se encuentra en asociación con la otra molécula peptídica en la naturaleza (por ejemplo, una proteína de fusión, o una proteína con una marca). Similarmente, un ácido nucleico o una proteína puede comprender una marca heteróloga o una secuencia de secreción o localización heteróloga.
- El término "locus" se refiere a una localización específica de un gen (o secuencia significativa), secuencia de ADN, secuencia codificante de polipéptido o posición en un cromosoma del genoma de un organismo. Por ejemplo, un "locus *MAPT*" puede referirse a la localización específica de un gen *MAPT*, secuencia de ADN *MAPT*, secuencia codificante de proteína-tau asociada a microtúbulos, o posición *MAPT* en un cromosoma del genoma de un organismo que se ha identificado como dónde reside dicha secuencia. Un "locus *MAPT*" puede comprender un elemento regulador de un gen *MAPT*, incluyendo, por ejemplo, un potenciador, un promotor, una región no traducida (UTR) 5' y/o 3', o una combinación de los mismos.
- El término "gen" se refiere a una secuencia de ADN en un cromosoma que codifica un producto (por ejemplo, un producto de ARN y/o un producto polipeptídico) e incluye la región codificante interrumpida con intrones no codificantes y secuencia localizada adyacente a la región codificante en los extremos 5' y 3' de tal manera que el gen corresponde al ARNm de longitud completa (incluyendo las secuencias no traducidas 5' y 3'). El término "gen" también incluye otras

secuencias no codificantes que incluyen secuencias reguladoras (por ejemplo, promotores, potenciadores y sitios de unión a factores de transcripción), señales de poliadenilación, sitios de entrada de ribosomas internos, silenciadores, secuencia aislante y regiones de unión a la matriz. Estas secuencias pueden estar cerca de la región codificante del gen (por ejemplo, en 10 kb) o en sitios distantes, e influyen en el nivel o la tasa de transcripción y traducción del gen.

El término "alelo" se refiere a una forma variante de un gen. Algunos genes tienen una variedad de formas diferentes, que se localizan en la misma posición, o locus genético, en un cromosoma. Un organismo diploide tiene dos alelos en cada locus genético. Cada par de alelos representa el genotipo de un locus genético específico. Los genotipos se describen como homocigóticos si hay dos alelos idénticos en un locus concreto y como heterocigóticos si los dos alelos difieren.

Un "promotor" es una región reguladora de ADN que suele incluir una caja TATA capaz de dirigir a la ARN polimerasa II para que inicie la síntesis de ARN en el lugar de inicio de la transcripción adecuado para una secuencia polinucleotídica particular. Un promotor puede comprender además otras regiones que influyen en la tasa de iniciación de la transcripción. Las secuencias promotoras desveladas en el presente documento modulan la transcripción de un polinucleótido unido operativamente. Un promotor puede ser activo en uno o más de los tipos de células desvelados en el presente documento (por ejemplo, una célula humana, una célula pluripotente, un embrión en estadio unicelular, una célula diferenciada o una combinación de los mismos). Un promotor puede ser, por ejemplo, un promotor constitutivamente activo, un promotor condicional, un promotor inducible, un promotor temporalmente restringido (por ejemplo, un promotor regulado por el desarrollo), o un promotor espacialmente restringido (por ejemplo, un promotor específico de célula o específico de tejido). Pueden encontrarse ejemplos de promotores, por ejemplo, en el documento de patente WO 2013/176772.

"Unión operativa" o estar "unido operativamente" incluye la yuxtaposición de dos o más componentes (por ejemplo, un promotor y otro elemento de secuencia) de manera que ambos componentes funcionen normalmente y permitan la posibilidad de que al menos uno de los componentes pueda mediar en una función que se ejerce sobre al menos uno de los otros componentes. Por ejemplo, un promotor puede estar unido operativamente a una secuencia codificante si el promotor controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores transcripcionales. La unión operativa puede incluir que dichas secuencias sean contiguas entre sí o que actúen en *trans* (por ejemplo, una secuencia reguladora puede actuar a distancia para controlar la transcripción de la secuencia codificante).

El término "variante" se refiere a una secuencia de nucleótidos que difiere de la secuencia más predominante en una población (por ejemplo, en un nucleótido) o a una secuencia de proteínas diferente de la secuencia más prevalente en una población (por ejemplo, en un aminoácido).

El término "fragmento", cuando se refiere a una proteína, significa una proteína que es más corta o tiene menos aminoácidos que la proteína de longitud completa. El término "fragmento", cuando se refiere a un ácido nucleico, significa un ácido nucleico que es más corto o tiene menos nucleótidos que el ácido nucleico de longitud completa. Un fragmento puede ser, por ejemplo, un fragmento aminoterminal (es decir, la eliminación de una parte del extremo carboxiterminal de la proteína), un fragmento carboxiterminal (es decir, la eliminación de una parte del extremo aminoterminal de la proteína) o un fragmento interno.

"Identidad de secuencia" o "identidad", en el contexto de dos polinucleótidos o secuencias polipeptídicas, se refiere a los residuos de las dos secuencias que son iguales cuando se alinean para obtener la máxima correspondencia en una ventana de comparación especificada. Cuando se usa el porcentaje de identidad de secuencia en referencia a las proteínas, las posiciones de residuos que no son idénticas a menudo difieren por sustituciones conservadas de aminoácidos, donde los residuos de aminoácidos se sustituyen por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobia) y, por lo tanto, no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservativas, el porcentaje de identidad de secuencia puede ajustarse al alza para corregir la naturaleza conservativa de la sustitución. Se dice que las secuencias que difieren por dichas sustituciones conservativas tienen "similitud de secuencia" o "similitud". Los medios para realizar este ajuste son bien conocidos. Normalmente, esto implica puntuar una sustitución conservativa como una discordancia parcial en lugar de total, aumentando así el porcentaje de identidad de secuencia. Así, por ejemplo, donde a un aminoácido idéntico se le da una puntuación de 1 y a una sustitución no conservativa se le da una puntuación de cero, a una sustitución conservativa se le da una puntuación entre cero y 1. La puntuación de las sustituciones conservativas se calcula, por ejemplo, como se implementa en el programa PC/GENE (Intelligenetics, Mountain View, California).

"Porcentaje de identidad de secuencia" incluye el valor determinado comparando dos secuencias óptimamente alineadas (mayor número de residuos perfectamente coincidentes) a lo largo de una ventana de comparación, en donde la porción de la secuencia polinucleotídica en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones ni deleciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las que aparece la base de ácido nucleico o el residuo de aminoácido idéntico en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número

total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia. A menos que se especifique lo contrario (por ejemplo, que la secuencia más corta incluya una secuencia heteróloga unida), la ventana de comparación es la longitud completa de la más corta de las dos secuencias que se comparan.

5 A menos que se indique lo contrario, los valores de identidad/similitud de secuencias incluyen el valor obtenido usando GAP Versión 10 usando los siguientes parámetros: % de identidad y % de similitud para una secuencia de nucleótidos usando peso por hueco de 50 y peso por longitud de 3, y la matriz de puntuación nwsgapdna.cmp; % de identidad y
10 % de similitud para una secuencia de aminoácidos usando peso por hueco de 8 y peso por longitud de 2, y la matriz de puntuación BLOSUM62; o cualquier programa equivalente. "Programa equivalente" incluye cualquier programa de comparación de secuencias que, para dos secuencias cualesquiera en cuestión, genere un alineamiento que tenga idénticas coincidencias de nucleótidos o residuos de aminoácidos y un idéntico porcentaje de identidad de secuencia cuando se compara con el correspondiente alineamiento generado por GAP Versión 10.

15 El término "sustitución conservativa de aminoácidos" se refiere a la sustitución de un aminoácido que normalmente está presente en la secuencia por un aminoácido diferente de tamaño, carga o polaridad similares. Los ejemplos de sustituciones conservativas incluyen la sustitución de un residuo no polar (hidrófobo), tal como isoleucina, valina o leucina por otro residuo no polar. Asimismo, ejemplos de sustituciones conservativas incluyen la sustitución de un residuo polar (hidrófilo) por otro, tal como entre arginina y lisina, entre glutamina y asparagina, o entre glicina y serina.
20 Además, la sustitución de un residuo básico, tal como lisina, arginina o histidina por otro, o la sustitución de un residuo ácido, tal como ácido aspártico o ácido glutámico por otro residuo ácido son ejemplos adicionales de sustituciones conservativas. Los ejemplos de sustituciones no conservativas incluyen la sustitución de un residuo de aminoácido no polar (hidrófobo), tal como isoleucina, valina, leucina, alanina o metionina por un residuo polar (hidrófilo), tal como cisteína, glutamina, ácido glutámico o lisina y/o un residuo polar por un residuo no polar. A continuación, se resumen las clasificaciones típicas de aminoácidos.
25

Tabla 1. Clasificaciones de aminoácidos.

Alanina	Ala	A	No polar	Neutro	1,8
Arginina	Arg	R	Polar	Positivo	-4,5
Asparagina	Asn	N	Polar	Neutro	-3,5
Ácido aspártico	Asp	D	Polar	Negativo	-3,5
Cisteína	Cys	C	No polar	Neutro	2,5
Ácido glutámico	Glu	E	Polar	Negativo	-3,5
Glutamina	Gln	Q	Polar	Neutro	-3,5
Glicina	Gly	G	No polar	Neutro	-0,4
Histidina	Su	H	Polar	Positivo	-3,2
Isoleucina	Isla	I	No polar	Neutro	4,5
Leucina	Leu	L	No polar	Neutro	3,8
Lisina	Lys	K	Polar	Positivo	-3,9
Metionina	Met	M	No polar	Neutro	1,9
Fenilalanina	Phe	F	No polar	Neutro	2,8
Prolina	Pro	P	No polar	Neutro	-1,6
Serina	Ser	S	Polar	Neutro	-0,8
Treonina	Thr	T	Polar	Neutro	-0,7
Triptófano	Tip	W	No polar	Neutro	-0,9
Tirosina	Tyr	Y	Polar	Neutro	-1,3
Valina	Val	V	No polar	Neutro	4,2

30 Una secuencia "homóloga" (por ejemplo, secuencia de ácido nucleico) incluye una secuencia que es idéntica o sustancialmente similar a una secuencia de referencia conocida, de manera que es, por ejemplo, al menos 50 %, al menos 55 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a la secuencia de referencia conocida. Las secuencias homólogas pueden incluir, por ejemplo, secuencias ortólogas y secuencias parálogas. Los genes homólogos, por ejemplo, descienden normalmente de una secuencia de ADN
35 ancestral común, ya sea a través de un evento de especiación (genes ortólogos) o un evento de duplicación genética (genes parálogos). Los genes "ortólogos" incluyen genes de especies diferentes que evolucionaron a partir de un gen ancestral común por especiación. Los genes ortólogos conservan normalmente la misma función a lo largo de la evolución. Los genes "parálogos" incluyen genes relacionados por duplicación dentro de un genoma. Los parálogos pueden desarrollar nuevas funciones en el curso de la evolución.

40 El término "in vitro" incluye entornos artificiales y a procesos o reacciones que ocurren en un entorno artificial (por ejemplo, un tubo de ensayo o una célula o línea celular aislada). El término "in vivo" incluye entornos naturales (por

ejemplo, una célula u organismo o cuerpo) y a procesos o reacciones que ocurren dentro de un entorno natural. El término "ex vivo" incluye células que han sido extraídas del cuerpo de un individuo y a procesos o reacciones que ocurren dentro de dichas células.

5 El término "gen indicador" se refiere a un ácido nucleico que tiene una secuencia que codifica un producto génico (normalmente una enzima) que se ensaya de forma fácil y cuantificable cuando una construcción que comprende la secuencia del gen indicador unida operativamente a un elemento promotor y/o potenciador heterólogo se introduce en células que contienen (o que se puede hacer que contengan) los factores necesarios para la activación de los elementos promotores y/o potenciadores. Los ejemplos de genes indicadores incluyen, pero no se limitan a, genes que codifican beta-galactosidasa (lacZ), genes de la cloranfenicol acetiltransferasa bacteriana (cat), genes de la luciferasa de luciérnaga, genes que codifican la beta-glucuronidasa (GUS) y genes que codifican las proteínas fluorescentes. Una "proteína indicadora" es una proteína codificada por un gen indicador.

15 El término "proteína indicadora fluorescente", como se usa en el presente documento, se refiere a una proteína indicadora que detectable basándose en fluorescencia, en donde la fluorescencia puede ser directamente de la proteína indicadora, de la actividad de la proteína indicadora sobre un sustrato fluorogénico o de una proteína con afinidad para unirse a un compuesto marcado con fluorescencia. Los ejemplos de proteínas fluorescentes incluyen proteínas fluorescentes verdes (por ejemplo, GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP y ZsGreen1), proteínas fluorescentes amarillas (por ejemplo, YFP, eYFP, Citrine, Venus, YPet, PhiYFP, y ZsYellow1), proteínas fluorescentes azules (por ejemplo, BFP, eBFP, eBFP2, Azurite, mKalamal, GFPuv, Sapphire y T-sapphire), proteínas fluorescentes cian (por ejemplo, CFP, eCFP, Cerulean, CyPet, AmCyan1 y Midoriishi-Cyan), proteínas fluorescentes rojas (por ejemplo, RFP, mKate, mKate2, mPlum, monómero de DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, monómero de DsRed, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry y Jred), proteínas fluorescentes naranjas (por ejemplo, mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine y tdTomato), y cualquier otra proteína fluorescente adecuada cuya presencia en las células pueda detectarse por métodos de citometría de flujo.

La reparación en respuesta a las roturas de doble cadena (DSB) se produce principalmente a través de dos vías de reparación del ADN conservadas: la recombinación homóloga (HR) y la unión de extremos no homólogos (NHEJ). Véase Kasparek & Humphrey (2011) *Seminars in Cell & Dev. Biol.* 22:886-897. Similarmente, la reparación de un ácido nucleico diana mediada por un ácido nucleico donante exógeno puede incluir cualquier proceso de intercambio de información genética entre los dos polinucleótidos.

35 El término "recombinación" incluye cualquier proceso de intercambio de información genética entre dos polinucleótidos y puede producirse por cualquier mecanismo. La recombinación puede producirse mediante reparación dirigida por homología (HDR) o recombinación homóloga (HR). La HDR o HR incluye una forma de reparación del ácido nucleico que puede requerir homología de secuencia de nucleótidos, usa una molécula "donante", tal como un molde para la reparación de una molécula "diana" (es decir, la que experimentó la rotura de doble cadena) y conduce a la transferencia de información genética del donante a la diana. Sin desear quedar ligado a teoría particular alguna, dicha transferencia puede implicar la corrección de los desajustes del ADN heterodúplex que se forma entre la diana rota y el donante, y/o la hibridación de cadenas dependiente de la síntesis, en la que el donante se usa para resintetizar la información genética que pasará a formar parte de la diana, y/o procesos relacionados. En algunos casos, el polinucleótido donante, una parte del polinucleótido donante, una copia del polinucleótido donante o una parte de una copia del polinucleótido donante se integra en el ADN diana. Véanse Wang et al. (2013) *Cell* 153:910-918; Mandalos et al. (2012) *PLOS ONE* 7:e45768:1-9; y Wang et al. (2013) *Nat Biotechnol.* 31:530-532.

La NHEJ incluye la reparación de roturas de doble cadena en un ácido nucleico mediante la ligadura directa de los extremos de la rotura entre sí o con una secuencia exógena sin necesidad de un molde homólogo. La ligadura de secuencias no contiguas mediante NHEJ puede dar lugar a deleciones, inserciones o translocaciones cerca del lugar de la rotura de doble cadena. Por ejemplo, la NHEJ también puede dar lugar a la integración dirigida de un ácido nucleico donante exógeno mediante la ligadura directa de los extremos de la rotura con los extremos del ácido nucleico donante exógeno (es decir, captura basada en la NHEJ). Dicha integración dirigida mediada por NHEJ puede ser preferible para la inserción de un ácido nucleico donante exógeno cuando las vías de reparación dirigidas por homología (HDR) no son fácilmente usables (por ejemplo, en células que no se dividen, células primarias y células que realizan mal la reparación del ADN basada en homología). Además, a diferencia de la reparación dirigida por homología, no es necesario el conocimiento de grandes regiones de identidad de secuencia que flanquean el sitio de corte, que puede ser beneficioso cuando se intenta la inserción dirigida en organismos que tienen genomas para los que hay un conocimiento limitado de la secuencia genómica. La integración puede realizarse mediante la ligadura de extremos romos entre el ácido nucleico donante exógeno y la secuencia genómica cortada, o mediante la ligadura de extremos cohesivos (es decir, que tienen salientes en 5' o 3') usando un ácido nucleico donante exógeno que está flanqueado por salientes compatibles con los generados por un agente de nucleasa en la secuencia genómica cortada. Véanse, por ejemplo, los documentos de patente US 2011/020722, WO 2014/033644, WO 2014/089290 y Maresca et al. (2013) *Genome Res.* 23(3):539-546. Si se ligan extremos romos, puede ser necesaria la resección de la diana y/o del donante para generar regiones de microhomología necesarias para la unión de fragmentos, que puede crear alteraciones no deseadas en la secuencia diana.

Las composiciones o métodos que "comprenden" o "incluyen" uno o más elementos citados pueden incluir otros elementos no citados específicamente. Por ejemplo, una composición que "comprende" o "incluye" una proteína puede contener la proteína sola o en combinación con otros componentes. La expresión transitoria "que consiste esencialmente en" significa que el alcance de una reivindicación debe interpretarse de manera que abarque los elementos especificados que se citan en la reivindicación y aquellos que no afectan materialmente a la(s) característica(s) básica(s) y novedosa(s) de la invención reivindicada. Por lo tanto, no se pretende que el término "que consiste esencialmente en", cuando se usa en una reivindicación de esta invención, deba interpretarse como equivalente a "que comprende".

"Opcional" u "opcionalmente" significa que el evento o circunstancia descrito posteriormente puede ocurrir o no y que la descripción incluye casos en los que el evento o circunstancia ocurre y casos en los que no.

La designación de un intervalo de valores incluye todos los números enteros dentro del intervalo o que lo definen, y todos los subintervalos definidos por números enteros dentro del intervalo.

A menos que se deduzca lo contrario del contexto, el término "aproximadamente" engloba valores dentro de un margen estándar de error de medición (por ejemplo, EEM) de un valor establecido.

El término "y/o" se refiere y abarca todas y cada una de las combinaciones posibles de uno o más de los elementos enumerados asociados, así como la falta de combinaciones cuando se interpreta de forma alternativa ("o").

El término "o" se refiere a cualquier miembro de una lista particular y también incluye cualquier combinación de miembros de esa lista.

Las formas en singular de los artículos "un", "una", "el" y "la" incluyen referencias al plural, a menos que el contexto dicte claramente lo contrario. Por ejemplo, los términos "una proteína" o "al menos una proteína" pueden incluir una pluralidad de proteínas, incluidas sus mezclas.

Estadísticamente significativo significa $p \leq 0,05$.

Descripción detallada

I. Perspectiva general

Se proporcionan células biosensoras de tau preparadas para proteína Cas y métodos de fabricación y uso de dichas células para cribar modificadores genéticos de la siembra o agregación de tau. Se proporcionan células biosensoras de tau preparadas para el mediador de activación sinérgica (SAM) CRISPR/Cas y métodos de fabricación y uso de dichas células para el cribado de modificadores genéticos de la siembra o agregación de tau. Se proporcionan células biosensoras de tau preparadas para proteína Cas y métodos de fabricación y uso de dichas células para detectar modificadores genéticos de la desagregación de tau. Se proporcionan células biosensoras de tau preparadas para el mediador de activación sinérgica (SAM) CRISPR/Cas y métodos de fabricación y uso de dichas células para cribar modificadores genéticos de la desagregación de tau. También se proporcionan reactivos y métodos para sensibilizar dichas células a la actividad de siembra de tau o a la agregación de tau. También se proporcionan reactivos y métodos para inducir la agregación de tau.

Con el fin de identificar genes y vías que modifiquen los procesos de agregación anormal de la proteína tau, se desarrolló una plataforma para realizar cribados con bibliotecas de ARN_g de nucleasa CRISPR (por ejemplo, CRISPR/Cas9) (CRISPR_n) para identificar genes que regulan el potencial de las células para ser "sembradas" por agregados de proteínas asociadas a la enfermedad de tau (por ejemplo, genes que, cuando se inactivan, hacen que las células sean más susceptibles a la formación de agregados de tau cuando se exponen a una fuente de proteína tau fibrilada). Para identificar aún más los genes y las vías que modifican los procesos de agregación anormal de la proteína tau, se desarrolló una plataforma para realizar cribados con bibliotecas de ARN_g de activación CRISPR (CRISPR_a) para identificar genes que regulan el potencial de las células para ser "sembradas" por agregados de proteína asociados a la enfermedad tau (por ejemplo, genes que, cuando se activan transcripcionalmente, hacen que las células sean más susceptibles a la formación de agregados de tau cuando se exponen a una fuente de proteína fibrilada tau). Similarmente, se desarrolló una plataforma para realizar cribados con bibliotecas de ARN_g de nucleasa CRISPR (por ejemplo, CRISPR/Cas9) (CRISPR_n) para identificar genes que, cuando se inactivan, evitan la agregación de tau o promueven la desagregación de tau. Asimismo, se desarrolló una plataforma para realizar cribados con bibliotecas de ARN_g de activación de CRISPR (CRISPR_a) para identificar genes que, cuando se activan transcripcionalmente, evitan la agregación de tau o promueven la desagregación de tau. Una "semilla" se refiere a una o más proteínas que nuclea la agregación de otras proteínas con un dominio de agregación similar. La actividad de siembra de una muestra se refiere a la capacidad de una muestra de nuclear (es decir, inducir) la agregación de una proteína con un dominio de agregación similar. La identificación de dichos genes puede dilucidar los mecanismos de propagación de agregados de tau de célula a célula y las vías genéticas que rigen la susceptibilidad de las neuronas a formar agregados de tau en el contexto de las enfermedades neurodegenerativas.

Los cribados emplean una línea celular de mamífero biosensora de tau (por ejemplo, una línea celular humana o HEK293T) que consiste en células que expresan de forma estable el dominio de repetición de tau (por ejemplo, el dominio de cuatro repeticiones de tau, tau_{4RD}) con una mutación patógena (por ejemplo, la mutación patógena P301S), unida a indicadores de proteína fluorescente únicos que pueden actuar juntos como un biosensor intracelular que produce una señal detectable cuando se agregan. En un ejemplo no limitante, las líneas celulares contienen dos transgenes que expresan de forma estable variantes de proteínas asociadas a la enfermedad fusionadas a la proteína fluorescente CFP (por ejemplo, eCFP) o a la proteína fluorescente YFP (por ejemplo, eYFP):tau^{4RD}-CFP/tau^{4RD}-YFP (TCY), en donde el dominio de repetición de tau (4RD) comprende la mutación patógena P301S. En estas líneas biosensoras, la agregación de proteínas tau-CFP/tau-YFP produce una señal de transferencia de energía por resonancia de fluorescencia (FRET), el resultado de una transferencia de energía fluorescente de la CFP donante a la YFP aceptora. El término CFP (proteína fluorescente cian), cuando se usa en el presente documento, incluye eCFP (proteína fluorescente cian mejorada), y el término YFP (proteína fluorescente amarilla), cuando se usa en el presente documento, incluye eYFP (proteína fluorescente amarilla mejorada). Las células FRET-positivas, que contienen agregados de tau, pueden clasificarse y aislarse por citometría de flujo. En el nivel basal, las células no estimuladas expresan los indicadores en un estado estable y soluble con una señal FRET mínima. Tras la estimulación (por ejemplo, transfección liposómica de partículas semilla), las proteínas indicadoras forman agregados, produciendo una señal FRET. Las células que contienen agregados pueden aislarse mediante FACS. Las líneas celulares que contienen agregados que se propagan de forma estable, Ag[+], pueden aislarse mediante dilución clónica en serie de líneas celulares Ag[-].

Se realizaron varias modificaciones en esta línea celular biosensora de tau para hacerla útil para el cribado genético usando bibliotecas de CRISPRn. En primer lugar, estas células biosensoras de tau se modificaron mediante la introducción de un transgén que expresa Cas (por ejemplo, Cas9 o SpCas9) para su uso en los cribados de CRISPRn. En segundo lugar, se desarrollaron reactivos y un método para sensibilizar las células a la actividad de siembra de tau y agregación de tau. Se desarrolló una línea celular en la que los agregados de tau persisten de forma estable en todas las células con crecimiento y múltiples pases en el tiempo. Estas células se usaron para producir medio acondicionado obteniendo el medio que ha estado en las células confluentes durante un periodo de tiempo. A continuación, este medio acondicionado puede aplicarse sobre biosensoras tau sin tratamiento previo-células tau en una proporción tal que pueda inducirse la agregación de tau en un pequeño porcentaje de estas células receptoras, sensibilizándolas así a la actividad de siembra de tau y a la agregación de tau. El medio acondicionado sin cocultivo no se había usado antes en este contexto como agente de siembra. Sin embargo, el medio acondicionado es particularmente útil para los cribados de todo el genoma a gran escala, debido a que las fibrillas tau producidas *in vitro* son un recurso limitado. Además, el medio acondicionado es más relevante fisiológicamente debido a que es producido y secretado por células en vez de *in vitro*.

Estas líneas celulares se usaron para desarrollar un método de cribado en el que las células biosensoras de tau que expresaban Cas sin agregados (Ag[-]) se transducían con una biblioteca de ARN guía de CRISPR para introducir mutaciones de inactivación en cada gen diana. Tras cultivar las células para permitir la edición y expansión del genoma, las células se cultivaron en medio acondicionado para sensibilizarlas a la actividad de siembra, y se identificaron las células en las que se producía agregación de tau. Se identificaron ARN guía que estaban enriquecidos en la subpoblación positiva a la agregación en relación con momentos anteriores durante la edición y expansión del genoma para identificar los genes que pueden regular la susceptibilidad de las células a la siembra de tau cuando se exponen a una fuente externa de actividad de siembra de tau.

Asimismo, se realizaron varias modificaciones en esta línea celular biosensora de tau para hacerla útil para el cribado genético usando bibliotecas de CRISPRa (por ejemplo, para su uso con un sistema CRISPR/Cas mediador de activación sinérgica (SAM)). En un sistema SAM a modo de ejemplo, varios dominios de activación interactúan para causar una activación transcripcional mayor que la que podría inducir cualquier factor por sí solo. Por ejemplo, un sistema SAM a modo de ejemplo comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional (por ejemplo, VP64) y una proteína adaptadora quimérica que comprende una proteína adaptadora (por ejemplo, la proteína de la cubierta de MS2 (MCP)) fusionada a uno o más dominios de activación transcripcional (por ejemplo, fusionada a p65 y HSF1). La MCP se une de forma natural a los bucles del tallo de MS2. En un sistema SAM a modo de ejemplo, la MCP interactúa con los bucles del tallo de MS2 manipulados en el ARN_g asociado a CRISPR y, de este modo, transporta los factores de transcripción unidos a la ubicación genómica adecuada.

En primer lugar, estas células biosensoras de tau se modificaron introduciendo uno o más transgenes que expresan la proteína Cas quimérica que comprende la proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional (por ejemplo, VP64) y la proteína adaptadora quimérica que comprende la proteína adaptadora (por ejemplo, proteína de la cubierta de MS2 (MCP)) fusionada a uno o más dominios de activación transcripcional (por ejemplo, fusionada a p65 y HSF1). Aunque en el presente documento se describen los sistemas SAM, también pueden usarse otros sistemas CRISPRa, tales como una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, en donde dichos sistemas no incluyen también una proteína adaptadora quimérica. En dichos casos, las células biosensoras de tau se modificarían introduciendo un transgén que exprese la proteína Cas quimérica.

En segundo lugar, se desarrollaron reactivos y un método para sensibilizar las células a la actividad de siembra de tau y agregación de tau. Se desarrolló una línea celular en la que los agregados de tau persisten de forma estable en todas las células con crecimiento y múltiples pases en el tiempo. Estas células se usaron para producir medio acondicionado obteniendo el medio que ha estado en las células confluentes durante un periodo de tiempo. A continuación, este medio acondicionado puede aplicarse sobre células tau biosensoras de tau sin tratamiento previo en una proporción tal que pueda inducirse la agregación de tau en un pequeño porcentaje de estas células receptoras, sensibilizándolas así a la actividad de siembra de tau y a la agregación de tau. El medio acondicionado sin cocultivo no se había usado antes en este contexto como agente de siembra. Sin embargo, el medio acondicionado es particularmente útil para los cribados de todo el genoma a gran escala, debido a que las fibrillas tau producidas *in vitro* son un recurso limitado. Además, el medio acondicionado es más relevante fisiológicamente debido a que es producido y secretado por células en vez de *in vitro*.

Estas líneas celulares se usaron para desarrollar un método de cribado en el que las células biosensoras de tau que expresaban SAM sin agregados (Ag[-]) se transducían con una biblioteca de ARN guía de CRISPRa para activar transcripcionalmente cada gen diana. Tras cultivar las células para permitir la edición y expansión del genoma, las células se cultivaron en medio acondicionado para sensibilizarlas a la actividad de siembra, y se identificaron las células en las que se producía agregación de tau. Se identificaron ARN guía que estaban enriquecidos en la subpoblación positiva a la agregación en relación con momentos anteriores durante la edición y expansión del genoma para identificar los genes que pueden regular la susceptibilidad de las células a la siembra de tau cuando se exponen a una fuente externa de actividad de siembra de tau.

II. Líneas celulares biosensoras Cas/tau y biosensoras SAM/tau y métodos de generación

A. Células biosensoras Cas/tau y células Biosensoras SAM/tau

En el presente documento se desvelan células que no solo expresan un primer dominio de repetición de tau (por ejemplo, que comprende el dominio de unión a microtúbulos de tau (MBD)) unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador, sino que también expresan una proteína Cas, tal como Cas9. También se desvelan en el presente documento células que no solo expresan un primer dominio de repetición de tau (por ejemplo, que comprende el dominio de unión a microtúbulos tau (MBD)) unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador, sino que también expresan una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional y una proteína adaptadora quimérica que comprende una proteína adaptadora fusionada a uno o más dominios de activación transcripcional. El primer dominio de repetición de tau unido al primer indicador puede expresarse de forma estable, y el segundo dominio de repetición de tau unido al segundo indicador puede expresarse de forma estable. Por ejemplo, el ADN que codifica el primer dominio de repetición de tau unido al primer indicador puede integrarse genómicamente, y el ADN que codifica el segundo dominio de repetición de tau unido al segundo indicador puede integrarse genómicamente. Similarmente, la proteína Cas puede expresarse de forma estable en las células biosensoras Cas/tau. Por ejemplo, el ADN que codifica la proteína Cas puede integrarse genómicamente. Similarmente, la proteína Cas quimérica y/o la proteína adaptadora quimérica pueden expresarse de forma estable en las células biosensoras SAM/tau. Por ejemplo, el ADN que codifica la proteína Cas quimérica puede integrarse genómicamente y/o el ADN que codifica la proteína adaptadora quimérica puede integrarse genómicamente. Las células pueden ser negativas a la agregación de tau o positivas a la agregación de tau.

1. Dominios de tau y de repetición de tau unidos a indicadores

La proteína tau asociada a los microtúbulos es una proteína que promueve el ensamblaje y la estabilidad de los microtúbulos y se expresa predominantemente en las neuronas. Tau desempeña una función en la estabilización de los microtúbulos neuronales y, por tanto, en la promoción del crecimiento axónico. En la enfermedad de Alzheimer (EA) y una familia de enfermedades neurodegenerativas relacionadas llamadas tauopatías, la proteína tau está anormalmente hiperfosforilada y agregada en haces de filamentos (filamentos helicoidales pareados), que se manifiestan, tal como ovillos neurofibrilares. Las tauopatías son un grupo de afecciones neurodegenerativas heterogéneas caracterizadas por la deposición de tau anormal en el cerebro.

El dominio de repetición de tau puede proceder de una proteína tau de cualquier animal o mamífero, tal como el ser humano, el ratón o la rata. En un ejemplo específico, el dominio de repetición de tau es de una proteína tau humana. A una proteína tau humana a modo de ejemplo se le asigna el número de acceso UniProt P10636. Las proteínas tau son productos de empalmes alternativos de un único gen que en los humanos se designa *MAPT* (proteína tau asociada a microtúbulos). El dominio de repetición de tau lleva los motivos de secuencia responsables de la agregación (es decir, es el dominio propenso a la agregación de tau). Dependiendo del empalme, el dominio de repetición de la proteína tau tiene tres o cuatro regiones de repetición que constituyen el núcleo de la proteína propenso a la agregación, que a menudo se denomina dominio de repetición (RD). Específicamente, el dominio de repetición de tau representa el núcleo de la región de unión a microtúbulos y alberga los motivos de hexapéptidos en R2 y R3 que son responsables de la agregación de tau. En el cerebro humano, existen seis isoformas de tau que oscilan de 352 a 441 aminoácidos de longitud. Estas isoformas varían en el extremo carboxilo según la presencia de tres o cuatro dominios de repetición (R1-R4), además de la presencia o ausencia de uno o dos dominios de inserción en el extremo amino.

Se cree que los dominios de repetición, situados en la mitad del extremo carboxilo de tau, son importantes para la unión a microtúbulos, así como para la agregación patológica de tau en filamentos helicoidales pareados (PHF), que son los componentes centrales de los ovillos neurofibrilares que se encuentran en las tauopatías. Las secuencias a modo de ejemplo de los cuatro dominios de repetición (R1-R4) se proporcionan en SEQ ID NOS: 1-4, respectivamente. Las secuencias codificantes a modo de ejemplo de los cuatro dominios de repetición (R1-R4) se proporcionan en SEQ ID NO: 5-8. En la SEQ ID NO: 9 se proporciona una secuencia a modo de ejemplo del dominio de cuatro repeticiones de tau. En la SEQ ID NO: 10 se proporciona una secuencia a modo de ejemplo del dominio de cuatro repeticiones de tau con la mutación P301S. En la SEQ ID NO: 11 se proporciona una secuencia a modo de ejemplo del dominio de cuatro repeticiones de tau con la mutación P301S. En la SEQ ID NO: 12 se proporciona una secuencia a modo de ejemplo del dominio de cuatro repeticiones de tau con la mutación P301S.

El dominio de repetición de tau usado en las células biosensoras Cas/tau o en las células biosensoras SAM/tau puede comprender el dominio de unión a microtúbulos de tau (MBD). El dominio de repetición de tau usado en las células biosensoras Cas/tau o las células biosensoras SAM/tau puede comprender uno o más o todos de los cuatro dominios de repetición (R1-R4). Por ejemplo, el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en uno o más o todos de SEQ ID NOS: 1, 2, 3 y 4, o secuencias al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idénticas a SEQ ID NOS: 1, 2, 3 y 4. En un ejemplo específico, el dominio de repetición de tau es el dominio de cuatro repeticiones de tau (R1-R4) que se encuentra en varias isoformas de tau. El dominio de cuatro repeticiones de tau puede usarse en lugar de tau de longitud completa debido a que forma de forma fiable fibrillas en células cultivadas. Por ejemplo, el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 9 o SEQ ID NO: 11 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 9 o SEQ ID NO: 11. En un ejemplo específico, el ácido nucleico que codifica el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 12 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 12, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 11. En otro ejemplo específico, el ácido nucleico que codifica el segundo dominio de repetición de tau unido al segundo indicador puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 10 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 10, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 9. El primer y segundo dominios de repetición de tau en las células desveladas en el presente documento pueden ser iguales, similares o diferentes.

Uno o ambos del primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en las células. Por ejemplo, los ácidos nucleicos que codifican uno o ambos del primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células y unirse operativamente a promotores activos en la célula.

Los dominios de repetición de tau usados en las células desveladas en el presente documento también pueden comprender una mutación patógena de tau, tal como una mutación pro-agregación. Dicha mutación puede ser, por ejemplo, una mutación que está asociada con (por ejemplo, segrega con) o causa una tauopatía. Como ejemplo, la mutación puede ser una mutación sensibilizadora de la agregación que sensibiliza a tau a la siembra, pero no da lugar a que tau se agregue fácilmente por sí misma. Por ejemplo, la mutación puede ser la mutación P301S asociada a la enfermedad. Por mutación P301S se entiende la mutación P301S de tau humana o una mutación correspondiente en otra proteína tau cuando está óptimamente alineada con la proteína tau humana. La mutación P301S en tau muestra una alta sensibilidad a la siembra, pero no se agrega fácilmente por sí sola. Así, aunque en el nivel basal las proteínas indicadoras tau que comprenden la mutación P301S existen en una forma estable y soluble dentro de la célula, la exposición a semillas exógenas de tau conduce a la agregación de la proteína indicadora tau. Otras mutaciones de tau incluyen, por ejemplo, K280del, P301L, V337M, P301L/V337M y K280del/I227P/I308P.

El primer dominio de repetición de tau puede unirse al primer indicador y el segundo dominio de repetición de tau puede unirse al segundo indicador por cualquier medio. Por ejemplo, el indicador puede fusionarse al dominio de repetición de tau (por ejemplo, como parte de una proteína de fusión).

El primer indicador y el segundo indicador son un par de indicadores únicos que pueden actuar juntos como un biosensor intracelular que produce una señal detectable cuando se agregan la primera y segunda proteínas. Específicamente, los indicadores son proteínas fluorescentes. El primer y segundo indicadores son un par de transferencia de energía por resonancia de fluorescencia (FRET), de tal manera que la FRET puede usarse para medir la agregación de proteínas. Los ejemplos de pares FRET (fluoróforos donante y aceptor) son bien conocidos. Véase, por ejemplo, Bajar et al. (2016) *Sensors (Basel)* 16(9): 1488. Las técnicas típicas de microscopía de fluorescencia se basan en la absorción por un fluoróforo de luz a una longitud de onda (excitación), seguida de la posterior emisión de fluorescencia secundaria a una longitud de onda más larga. El mecanismo de transferencia de energía de resonancia de fluorescencia implica un fluoróforo donante en un estado electrónico excitado, que puede transferir su energía de excitación a un cromóforo aceptor cercano de forma no radiativa a través de interacciones dipolo-dipolo de largo alcance. Por ejemplo, el donante de energía FRET puede ser el primer indicador, y el aceptor de energía FRET puede ser el segundo indicador. Alternativamente, el donante de energía FRET puede ser el segundo indicador, y el aceptor

de energía FRET puede ser el primer indicador. En un ejemplo específico, el primer y segundo indicadores son CFP e YFP. Se proporcionan proteínas y secuencias de codificación a modo de ejemplo para CFP, por ejemplo, en SEQ ID NO: 13 y 14, respectivamente. Se proporcionan proteínas y secuencias codificantes a modo de ejemplo para YFP, por ejemplo, en SEQ ID NO: 15 y 16, respectivamente. Como ejemplo específico, CFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 13 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 13. Como otro ejemplo específico, la YFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 15 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 15.

En un ejemplo no limitante, las células biosensoras desveladas en el presente documento contienen dos transgenes que expresan de forma estable variantes de la proteína tau asociadas a la enfermedad y fusionadas a la proteína fluorescente CFP o la proteína fluorescente YFP, respectivamente (τ^{4RD} -CFP/ τ^{4RD} -YFP (TCY)), en donde el dominio de cuatro repeticiones de tau (4RD) comprende la mutación patógena P301S. En estas líneas biosensoras, la agregación de las proteínas tau-CFP/tau-YFP produce una señal FRET, el resultado de la transferencia de energía fluorescente de la CFP donante a la YFP receptora. Las células FRET-positivas, que contienen agregados de tau, pueden clasificarse y aislarse por citometría de flujo. En el nivel basal, las células no estimuladas expresan los indicadores en un estado estable y soluble con una señal FRET mínima. Tras la estimulación (por ejemplo, transfección liposómica de partículas semilla), las proteínas indicadoras forman agregados, produciendo una señal FRET.

Las células biosensoras Cas/tau desveladas en el presente documento pueden ser células positiva a la agregación (Ag[+]) en las que el dominio de repetición de tau se presenta de forma estable en un estado agregado, lo que significa que los agregados del dominio de repetición de tau persisten de forma estable en todas las células con crecimiento y múltiples pases a lo largo del tiempo. Alternativamente, las células biosensoras Cas/tau desveladas en el presente documento pueden ser de agregación negativa (Ag[-]).

2. Proteínas Cas y proteínas Cas quiméricas

Las células biosensoras Cas/tau desveladas en el presente documento también comprenden ácidos nucleicos (ADN o ARN) que codifican proteínas Cas. Opcionalmente, la proteína Cas se expresa de forma estable. Opcionalmente, las células comprenden una secuencia codificante Cas genómicamente integrada. Asimismo, las células biosensoras SAM/tau desveladas en el presente documento también comprenden ácidos nucleicos (ADN o ARN) que codifican proteínas Cas quiméricas que comprenden una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional (por ejemplo, VP64). Opcionalmente, la proteína Cas quimérica se expresa de forma estable. Opcionalmente, las células comprenden una secuencia codificante de Cas quimérica genómicamente integrada.

Las proteínas Cas forman parte de los sistemas de repeticiones palindrómicas cortas regularmente intercaladas agrupadas (CRISPR)/asociadas a CRISPR (Cas). Los sistemas CRISPR/Cas incluyen transcritos y otros elementos implicados en la expresión de, o que dirigen la actividad de, genes Cas. Un sistema CRISPR/Cas puede ser, por ejemplo, un sistema de tipo I, de tipo II, de tipo III o de tipo V (por ejemplo, subtipo V-A o subtipo V-B). Los métodos y composiciones desvelados en el presente documento pueden emplear sistemas CRISPR/Cas utilizando complejos CRISPR (que comprenden un ARN guía (ARNg) complejo con una proteína Cas) para la unión o corte de ácidos nucleicos dirigida al sitio.

Los sistemas CRISPR/Cas usados en las composiciones y métodos desvelados en el presente documento pueden ser de origen no natural. Un sistema "de origen no natural" incluye cualquier cosa que indique la participación de la mano del hombre, tal como que uno o más componentes del sistema que están alterados o mutados con respecto a su estado natural, que están al menos sustancialmente libres de al menos otro componente con el que están asociados de forma natural en la naturaleza, o que están asociados con al menos otro componente con el que no están asociados de forma natural. Por ejemplo, algunos sistemas CRISPR/Cas emplean complejos CRISPR no naturales que comprenden un ARNg y una proteína Cas que no se encuentran juntos de forma natural, emplean una proteína Cas que no se encuentra de forma natural o emplean un ARNg que no se encuentra de forma natural.

Las proteínas Cas comprenden generalmente al menos un dominio de reconocimiento o unión a ARN que puede interactuar con ARN guía. Las proteínas Cas también pueden comprender dominios de la nucleasa (por ejemplo, dominios DNasa o dominios RNasa), dominios de unión a ADN, dominios helicasa, dominios de interacción proteína-proteína, dominios de dimerización y otros dominios. Algunos de dichos dominios (por ejemplo, los dominios DNasa) pueden ser de una proteína Cas nativa. Otros de dichos dominios pueden añadirse para hacer una proteína Cas modificada. Un dominio de la nucleasa posee actividad catalítica para el corte de ácido nucleico, que incluye la ruptura de los enlaces covalentes de una molécula de ácido nucleico. El corte puede producir extremos romos o protuberantes, y puede ser monocatenario o bicatenario. Por ejemplo, una proteína Cas9 natural creará normalmente un producto de corte romo. Alternativamente, una proteína Cpf1 natural (por ejemplo, FnCpf1) puede dar lugar a un producto de corte con un saliente en 5' de 5 nucleótidos, produciéndose el corte después del 18.º par de bases de la secuencia PAM en la cadena no diana y después de la 23.ª base en la cadena diana. Una proteína Cas puede tener una actividad de corte completa para crear una rotura de doble cadena en un locus genómico diana (por ejemplo, una rotura de doble cadena con extremos romos), o puede ser una melladora que cree una rotura de cadena simple en un locus genómico

diana.

Los ejemplos de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 o Csx12), Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 y Cu1966, y homólogos o versiones modificadas de las mismas.

Una proteína Cas a modo de ejemplo es una proteína Cas9 o una proteína derivada de una proteína Cas9. Las proteínas Cas9 son de un sistema CRISPR/Cas de tipo II y normalmente comparten cuatro motivos clave con una arquitectura conservada. Los motivos 1, 2 y 4 son motivos similares a RuvC, y el motivo 3 es un motivo HNH. Las proteínas Cas9 a modo de ejemplo son de *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus sp.*, *Staphylococcus aureus*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycooides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas sp.*, *Crocospaera watsonii*, *Cyanothece sp.*, *Microcystis aeruginosa*, *Synechococcus sp.*, *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccii*, *Candidatus Desulfurudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finexgoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochroamatium vinosum*, *Marinobacter sp.*, *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc sp.*, *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira sp.*, *Lynngbya sp.*, *Microcoleus chthonoplastes*, *Oscillatoria sp.*, *Petrotoga mobilis*, *Thermosiphon africanus*, *Acaryochloris marina*, *Neisseria meningitidis* o *Campylobacter jejuni*. Ejemplos adicionales de miembros de la familia Cas9 se describen en el documento de patente WO 2014/131833. Cas9 de *S. pyogenes* (SpCas9) (número de acceso SwissProt asignado Q99ZW2) es una proteína Cas9 a modo de ejemplo. La proteína SpCas9 a modo de ejemplo y la secuencia codificante se exponen en SEQ ID NOS: 21 y 22, respectivamente. Cas9 de *S. aureus* (SaCas9) (número de acceso UniProt asignado J7RUA5) es otra proteína Cas9 a modo de ejemplo. Cas9 de *Campylobacter jejuni* (CjCas9) (número de acceso UniProt asignado Q0P897) es otra proteína Cas9 a modo de ejemplo. Véase, por ejemplo, Kim et al. (2017) *Nat. Comm.* 8:14500. SaCas9 es más pequeña que SpCas9, y CjCas9 es más pequeña que SaCas9 y SpCas9. Cas9 de *Neisseria meningitidis* (Nme2Cas9) es otra proteína Cas9 a modo de ejemplo. Véase, por ejemplo, Edraki et al. (2019) *Mol. Cell* 73(4):714-726. Las proteínas Cas9 de *Streptococcus thermophilus* (por ejemplo, Cas9 de *Streptococcus thermophilus* LMD-9 codificada por el locus CRISPR1 (St1Cas9) o Cas9 de *Streptococcus thermophilus* del locus CRISPR3 (St3Cas9)) son otras proteínas Cas9 a modo de ejemplo. Cas9 de *Francisella novicida* (FnCas9) o la variante RHA de Cas9 de *Francisella novicida* que reconoce una PAM alternativa (sustituciones E1369R/E1449H/R1556A) son otras proteínas Cas9 a modo de ejemplo. Estas y otras proteínas Cas9 a modo de ejemplo se revisan, por ejemplo, en Cebrián-Serrano y Davies (2017) *Mamm. Genome* 28(7):247-261.

Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 21 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 21. Como otro ejemplo específico, una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa y uno o más dominios de activación transcripcional puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 36 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 36.

Otro ejemplo de una proteína Cas es la proteína Cpfl (CRISPR de *Prevotella* y *Francisella* 1). Cpfl es una proteína de gran tamaño (aproximadamente 1300 aminoácidos) que contiene un dominio de la nucleasa similar a RuvC homólogo al dominio correspondiente de Cas9 junto con un homólogo al grupo rico en arginina característico de Cas9. Sin embargo, Cpfl carece del dominio de la nucleasa HNH que está presente en las proteínas Cas9, y el dominio similar a RuvC es contiguo en la secuencia Cpfl, a diferencia de Cas9, donde contiene inserciones largas que incluyen el dominio HNH. Véase, por ejemplo, Zetsche et al. (2015) *Cell* 163(3):759-771. Las proteínas Cpfl a modo de ejemplo proceden de *Francisella tularensis* 1, *Francisella tularensis subsp. novicida*, *Prevotella albensis*, *Lachnospiraceae bacteria MC2017 1*, *Butyrivibrio proteoclasticus*, *Peregrinibacteria bacteria GW2011 GWA2 33 10*, *Parcubacteria bacteria GW2011 GWC2 44 17*, *Smithella sp. SCADC*, *Acidaminococcus sp. BV3L6*, *Lachnospiraceae bacteria MA2020*, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi 237*, *Leptospira inadai*, *Lachnospiraceae bacteria ND2006*, *Porphyromonas crevioricanis 3*, *Prevotella disiens* y *Porphyromonas macacae*. Cpfl de *Francisella novicida* U112 (FnCpfl1; número de acceso UniProt asignado A0Q7Q2) es una proteína Cpfl a modo de ejemplo.

Las proteínas Cas pueden ser proteínas naturales (es decir, las que se dan en la naturaleza), proteínas Cas modificadas (es decir, variantes de proteínas Cas) o fragmentos de proteínas Cas naturales o modificadas. Las proteínas Cas también pueden ser variantes o fragmentos activos con respecto a la actividad catalítica de las proteínas Cas naturales o modificadas. Las variantes o fragmentos activos con respecto a la actividad catalítica pueden comprender al menos 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más de identidad de secuencia con la proteína Cas natural o modificada o una porción de la misma, en donde las variantes activas

retienen la capacidad de cortar en un sitio de corte deseado y, por lo tanto, retienen la actividad inductora de rotura de doble cadena. Los ensayos para la actividad inductora de la rotura de la doble cadena son conocidos y generalmente miden la actividad global y la especificidad de la proteína Cas en sustratos de ADN que contienen el sitio de corte.

5 Un ejemplo de una proteína Cas modificada es la proteína SpCas9-HF1 modificada, que es una variante de alta fidelidad de Cas9 de *Streptococcus pyogenes* que alberga alteraciones (N497A/R661A/Q695A/Q926A) diseñadas para reducir los contactos no específicos con el ADN. Véase, por ejemplo, Kleinstiver et al. (2016) Nature 529(7587):490-495. Otro ejemplo de una proteína Cas modificada es la variante eSpCas9 modificada (K848A/K1003A/R1060A) diseñada para reducir los efectos inespecíficos. Véase, por ejemplo, Slaymaker et al. (2016) Science 351(6268):84-88. Otras variantes de SpCas9 incluyen K855A y K810A/K1003A/R1060A. Estas y otras proteínas Cas modificadas se revisan, por ejemplo, en Cebrian-Serrano y Davies (2017) Mamm. Genome 28(7):247-261. Otro ejemplo de una proteína Cas9 modificada es xCas9, que es una variante de SpCas9 que puede reconocer una gama ampliada de secuencias PAM. Véase, por ejemplo, Hu et al. (2018) Nature 556:57-63.

15 Las proteínas Cas pueden modificarse para aumentar o disminuir una o más de afinidad de unión al ácido nucleico, especificidad de unión al ácido nucleico y actividad enzimática. Las proteínas Cas también pueden modificarse para cambiar cualquier otra actividad o propiedad de la proteína, tal como la estabilidad. Por ejemplo, una proteína Cas puede truncarse para eliminar dominios que no son esenciales para la función de la proteína o para optimizar (por ejemplo, potenciar o reducir) la actividad o una propiedad de la proteína Cas. Como otro ejemplo, uno o más dominios de la nucleasa de la proteína Cas pueden modificarse, delecionarse o inactivarse (por ejemplo, para su uso en las células biosensoras SAM/tau que comprenden una proteína Cas inactiva por nucleasa).

25 Las proteínas Cas pueden comprender al menos un dominio de la nucleasa, tal como un dominio DNasa. Por ejemplo, una proteína CpfI natural comprende generalmente un dominio similar a RuvC que corta ambas cadenas de ADN diana, quizás en una configuración dimérica. Las proteínas Cas también pueden comprender al menos dos dominios de la nucleasa, tales como dominios DNasa. Por ejemplo, una proteína Cas9 natural comprende generalmente un dominio de la nucleasa similar a RuvC y un dominio de la nucleasa similar a HNH. Cada uno de los dominios RuvC y HNH puede cortar una cadena diferente de ADN bicatenario para hacer una rotura de doble cadena en el ADN. Véase, por ejemplo, Jinek et al. (2012) Science 337:816-821.

30 Uno, varios o todos los dominios de la nucleasa pueden delecionarse o mutarse para que dejen de ser funcionales o tengan una actividad nucleasa reducida. Por ejemplo, si se deleciona o muta uno de los dominios de la nucleasa en una proteína Cas9, la proteína Cas9 resultante puede denominarse melladora y puede generar una rotura de cadena simple en un ADN diana de doble cadena, pero no una rotura de doble cadena (es decir, puede cortar la cadena complementaria o la cadena no complementaria, pero no ambas). Si se delecionan o mutan ambos dominios de la nucleasa, la proteína Cas resultante (por ejemplo, Cas9) tendrá una capacidad reducida para cortar ambas cadenas de un ADN bicatenario (por ejemplo, una proteína Cas sin nucleasa o con nucleasa inactiva, o una proteína Cas catalíticamente muerta (dCas)). Un ejemplo de una mutación que convierte Cas9 en una melladora es una mutación D10A (aspartato a alanina en la posición 10 de Cas9) en el dominio RuvC de Cas9 de *S. pyogenes*. Similarmente, H939A (histidina a alanina en la posición 839 del aminoácido), H840A (histidina a alanina en la posición 840 del aminoácido) o N863A (asparagina a alanina en la posición 863 del aminoácido) en el dominio HNH de Cas9 de *S. pyogenes* puede convertir la Cas9 en una melladora. Otros ejemplos de mutaciones que convierten Cas9 en una melladora incluyen las mutaciones correspondientes a Cas9 de *S. thermophilus*. Véase, por ejemplo, Sapranaukas et al. (2011) Nucleic Acids Res. 39(21):9275-9282 y el documento de patente WO 2013/141680. Dichas mutaciones pueden generarse usando métodos, tales como la mutagénesis dirigida al sitio, la mutagénesis mediada por PCR o la síntesis total de genes. Pueden encontrarse ejemplos de otras mutaciones que crean melladoras, por ejemplo, en los documentos de patente WO 2013/176772 y WO 2013/142578. Si se delecionan o mutan todos los dominios de la nucleasa en una proteína Cas (por ejemplo, se delecionan o mutan ambos dominios de la nucleasa en una proteína Cas9), la proteína Cas resultante (por ejemplo, Cas9) tendrá una capacidad reducida para cortar ambas cadenas de un ADN de doble cadena (por ejemplo, una proteína Cas nula en nucleasa o inactiva en nucleasa). Un ejemplo específico es un doble mutante D10A/H840A de Cas9 de *S. pyogenes* o un doble mutante correspondiente en un Cas9 de otra especie cuando se alinea óptimamente con Cas9 de *S. pyogenes*. Otro ejemplo específico es un doble mutante D10A/N863A de Cas9 de *S. pyogenes* o un doble mutante correspondiente en un Cas9 de otra especie cuando se alinea óptimamente con Cas9 de *S. pyogenes*.

55 Los ejemplos de mutaciones inactivadoras en los dominios catalíticos de xCas9 son los mismos que los descritos anteriormente para SpCas9. También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de las proteínas Cas9 de *Staphylococcus aureus*. Por ejemplo, la enzima Cas9 de *Staphylococcus aureus* (SaCas9) puede comprender una sustitución en la posición N580 (por ejemplo, sustitución N580A) y una sustitución en la posición D10 (por ejemplo, sustitución D10A) para generar una proteína Cas inactiva por nucleasa. Véase, por ejemplo, el documento de patente WO 2016/106236. También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de Nme2Cas9 (por ejemplo, combinación de D16A y H588A). También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de St1Cas9 (por ejemplo, la combinación de D9A, D598A, H599A y N622A). También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de St3Cas9 (por ejemplo, la combinación de D10A y N870A). También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de CjCas9 (por ejemplo, la combinación de D8A y H559A). También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de FnCas9 y RHA FnCas9 (por ejemplo, N995A).

También se conocen ejemplos de mutaciones inactivadoras en los dominios catalíticos de las proteínas Cpfl. Con referencia a las proteínas Cpfl de *Francisella novicida* U112 (FnCpf1), *Acidaminococcus* sp. BV3L6 (AsCpf1), *Lachnospiraceae bacterium* ND2006 (LbCpf1) y *Moraxella bovoculi* 237 (MbCpf1 Cpfl), dichas mutaciones pueden incluir mutaciones en las posiciones 908, 993 o 1263 de AsCpf1 o posiciones correspondientes en ortólogos de Cpfl, o posiciones 832, 925, 947 o 1180 de LbCpf1 o posiciones correspondientes en ortólogos de Cpfl. Dichas mutaciones pueden incluir, por ejemplo, una o más de las mutaciones D908A, E993A y D1263A de AsCpf1, o mutaciones correspondientes en ortólogos de Cpfl, o D832A, E925A, D947A y D1180A de LbCpf1, o mutaciones correspondientes en ortólogos de Cpfl. Véase, por ejemplo, el documento de patente US 2016/0208243.

Las proteínas Cas también pueden unirse operativamente a polipéptidos heterólogos como proteínas de fusión. Por ejemplo, una proteína Cas puede fusionarse a un dominio de corte, un dominio de modificación epigenética, un dominio de activación transcripcional o un dominio represor transcripcional. Véase el documento de patente WO 2014/089290. Por ejemplo, las proteínas Cas pueden unirse de forma operable o fusionarse a un dominio de activación transcripcional para su uso en las células biosensoras SAM/tau. Ejemplos de dominios de activación transcripcional incluyen un dominio de activación de VP16 del virus del herpes simple, VP64 (que es un derivado tetramérico de VP16), un dominio de activación de NFκB p65, dominios de activación de p53 1 y 2, un dominio de activación de CREB (proteína de unión a elemento de respuesta a AMPc), un dominio de activación de E2A y un dominio de activación de NFAT (factor nuclear de linfocitos T activados). Otros ejemplos incluyen dominios de activación de Oct1, Oct-2A, SP1, AP-2, CTF1, P300, CBP, PCAF, SRC1, PVALF, ERF-2, OsGAI, HALF-1, C1, AP1, ARF-5, ARF-6, ARF-7, ARF-8, CPRF1, CPRF4, MYC-RP/GP, TRAB1PC4, y HSF1. Véanse, por ejemplo, los documentos de patente US 2016/0237456, EP3045537 y WO 2011/146121. En algunos casos, puede usarse un sistema de activación transcripcional que comprenda una proteína de fusión dCas9-VP64 emparejada con MS2-p65-HSF1. Los ARN guía en dichos sistemas pueden diseñarse con secuencias aptámeras añadidas al tetrabucle y al tallo-bucle 2 del ARN guía diseñados para unirse a las proteínas de la cubierta del bacteriófago MS2 dimerizado. Véase, por ejemplo, Konermann et al. (2015) *Nature* 517(7536):583-588. Ejemplos de dominios represores transcripcionales incluyen dominios represores tempranos de AMPc inducible (ICER), dominios represores de la caja A asociada a Kruppel (KRAB-A), dominios represores ricos en glicina YY1, represores similares a Sp1, represores E(spl), represor IκB y MeCP2. Otros ejemplos incluyen dominios represores transcripcionales de A/B, KOX, gen temprano inducible por TGF-beta (TIEG), v-erbA, SID, SID4X, MBD2, MBD3, DNMT1, DNMG3A, DNMT3B, Rb, ROM2, Véanse, por ejemplo, los documentos de patente EP3045537 y WO 2011/146121. Las proteínas Cas también pueden fusionarse a un polipéptido heterólogo que proporcione una estabilidad mayor o menor. El dominio fusionado o el polipéptido heterólogo pueden localizarse en el extremo N, el extremo C o internamente dentro de la proteína Cas.

Las proteínas Cas también pueden unirse operativamente a polipéptidos heterólogos como proteínas de fusión. Como un ejemplo, una proteína Cas puede fusionarse a uno o más polipéptidos heterólogos que permiten la localización subcelular. Dichos polipéptidos heterólogos pueden incluir, por ejemplo, una o más señales de localización nuclear (NLS), tales como la NLS monopartita de SV40 y/o una NLS bipartita de alfa-importina para la localización en el núcleo, una señal de localización mitocondrial para la localización en las mitocondrias, una señal de retención en el RE y similares. Véase, por ejemplo, Lange et al. (2007) *J. Biol. Chem.* 282:5101-5105. Dichas señales de localización subcelular pueden localizarse en el extremo N, el extremo C, o en cualquier lugar dentro de la proteína Cas. Una NLS puede comprender un tramo de aminoácidos básicos, y puede ser una secuencia monopartita o una secuencia bipartita. Opcionalmente, una proteína Cas puede comprender dos o más NLS, incluyendo una NLS (por ejemplo, una NLS de alfa-importina o una NLS monopartita) en el extremo N y una NLS (por ejemplo, una NLS de SV40 o una NLS bipartita) en el extremo C. Una proteína Cas también puede comprender dos o más NLS en el extremo N y/o dos o más NLS en el extremo C.

Las proteínas Cas también pueden unirse operativamente a un dominio de penetración celular o a un dominio de transducción de proteínas. Por ejemplo, el dominio de penetración celular puede derivar de la proteína TAT del VIH-1, el motivo de penetración celular TLM del virus de la hepatitis B humana, MPG, Pep-1, VP22, un péptido de penetración celular del virus del herpes simple o una secuencia peptídica de poliarginina. Véanse, por ejemplo, los documentos de patente WO 2014/089290 y WO 2013/176772. El dominio de penetración celular puede estar situado en el extremo N, en el extremo C o en cualquier lugar dentro de la proteína Cas.

Las proteínas Cas también pueden unirse operativamente a un polipéptido heterólogo para facilitar el seguimiento o la purificación, tal como una proteína fluorescente, una marca de purificación o una marca epitópica. Los ejemplos de proteínas fluorescentes incluyen proteínas fluorescentes verdes (por ejemplo, GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ejemplo, YFP, eYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ejemplo, eBFP, eBFP2, Azurite, mKalamal, GFPuv, Sapphire, T-sapphire), proteínas fluorescentes cian (por ejemplo, eCFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (por ejemplo, mKate, mKate2, mPlum, monómero de DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, monómero de DsRed, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), proteínas fluorescentes naranjas (por ejemplo, mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato), y cualquier otra proteína fluorescente adecuada. Los ejemplos de marcas incluyen glutatión-S-transferasa (GST), proteína de unión a quitina (CBP), proteína de unión a maltosa, tioredoxina (TRX), poli(NANP), etiqueta de purificación por afinidad en tándem (TAP), myc, AcV5,

AU1, AU5, E, ECS, E2, FLAG, hemaglutinina (HA), nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, histidina (His), proteína transportadora de carboxilo de biotina (BCCP) y calmodulina.

Las proteínas Cas pueden suministrarse en cualquier forma. Por ejemplo, una proteína Cas puede proporcionarse en forma de una proteína. Por ejemplo, una proteína Cas puede proporcionarse como una proteína Cas complejada con un ARN_g. Alternativamente, una proteína Cas puede proporcionarse en forma de un ácido nucleico que codifica la proteína Cas, tal como un ARN (por ejemplo, ARN mensajero (ARN_m)) o ADN. Opcionalmente, el ácido nucleico que codifica la proteína Cas puede estar optimizado en codones para su traducción eficiente en proteína en una célula u organismo particular. Por ejemplo, el ácido nucleico que codifica la proteína Cas puede modificarse para sustituir codones que tengan una mayor frecuencia de uso en una célula bacteriana, una célula de levadura, una célula humana, una célula no humana, una célula de mamífero, una célula de roedor, una célula de ratón, una célula de rata o cualquier otra célula huésped de interés, en comparación con la secuencia polinucleotídica natural. Por ejemplo, el ácido nucleico que codifica la proteína Cas puede estar optimizado en codones para su expresión en una célula humana. Cuando un ácido nucleico que codifica la proteína Cas se introduce en la célula, la proteína Cas puede expresarse transitoria, condicional o constitutivamente en la célula.

Las proteínas Cas suministradas como ARN_m pueden modificarse para mejorar la estabilidad y/o las propiedades de inmunogenicidad. Las modificaciones pueden realizarse en uno o más nucleósidos dentro del ARN_m. Ejemplos de modificaciones químicas de las nucleobases del ARN_m incluyen la pseudouridina, la 1-metil-pseudouridina y la 5-metil-citidina. Por ejemplo, puede usarse ARN_m de Cas con caperuza y poliadenilado que contenga N1-metil-pseudouridina. Similarmente, los ARN_m de Cas pueden modificarse por agotamiento de uridina usando codones sinónimos.

Los ácidos nucleicos que codifican proteínas Cas pueden integrarse de forma estable en el genoma de una célula y unirse operativamente a un promotor activo en la célula. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 22 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 22, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 21. En otro ejemplo específico, el ácido nucleico que codifica una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa y uno o más dominios de activación transcripcional puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 38 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 38, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 36. Alternativamente, los ácidos nucleicos que codifican proteínas Cas pueden unirse operativamente a un promotor en una construcción de expresión. Las construcciones de expresión incluyen cualquier construcción de ácido nucleico capaz de dirigir la expresión de un gen u otra secuencia de ácido nucleico de interés (por ejemplo, un gen Cas) y que puede transferir dicha secuencia de ácido nucleico de interés a una célula diana. Los promotores que pueden usarse en una construcción de expresión incluyen promotores activos, por ejemplo, en una o más de una célula eucariota, una célula humana, una célula no humana, una célula de mamífero, una célula de mamífero no humano, una célula de roedor, una célula de ratón, una célula de rata, una célula pluripotente, una célula madre embrionaria (ES), una célula madre adulta, una célula progenitora de desarrollo restringido, una célula madre pluripotente inducida (iPS), o un embrión en estadio unicelular. Dichos promotores pueden ser, por ejemplo, promotores condicionales, promotores inducibles, promotores constitutivos o promotores específicos de tejido.

3. Proteínas adaptadoras quiméricas

Las células biosensoras SAM/tau desveladas en el presente documento pueden comprender no solo ácidos nucleicos (ADN o ARN) que codifican una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional (por ejemplo, VP64), sino opcionalmente también ácidos nucleicos (ADN o ARN) que codifican una proteína adaptadora quimérica que comprende una proteína adaptadora (por ejemplo, la proteína de la cubierta de MS2 (MCP)) fusionada a uno o más dominios de activación transcripcional (por ejemplo, fusionada a p65 y HSF1). Opcionalmente, la proteína Cas quimérica y/o la proteína adaptadora quimérica se expresan de forma estable. Opcionalmente, las células comprenden una secuencia codificante de proteína Cas quimérica genómicamente integrada y/o una secuencia codificante de proteína adaptadora quimérica genómicamente integrada.

Dichas proteínas adaptadoras quiméricas comprenden: (a) un adaptador (es decir, dominio adaptador o proteína adaptadora) que se une específicamente a un elemento de unión al adaptador dentro de un ARN guía; y (b) uno o más dominios heterólogos de activación transcripcional. Por ejemplo, dichas proteínas de fusión pueden comprender 1, 2, 3, 4, 5, o más dominios de activación transcripcional (por ejemplo, dos o más dominios heterólogos de activación transcripcional o tres o más dominios heterólogos de activación transcripcional). En un ejemplo, dichas proteínas adaptadoras quiméricas pueden comprender: (a) un adaptador (es decir, un dominio adaptador o proteína adaptadora) que se une específicamente a un elemento de unión al adaptador en un ARN guía; y (b) dos o más dominios de activación transcripcional. Por ejemplo, la proteína adaptadora quimérica puede comprender: (a) un adaptador de proteína de la cubierta de MS2 que se une específicamente a uno o más aptámeros MS2 en un ARN guía (por ejemplo, dos aptámeros MS2 en ubicaciones separadas en un ARN guía); y (b) uno o más (por ejemplo, dos o más dominios

de activación transcripcional). Por ejemplo, los dos dominios de activación transcripcional pueden ser dominios de activación transcripcional p65 y HSF1 o fragmentos funcionales o variantes de los mismos. Sin embargo, también se proporcionan proteínas adaptadoras quiméricas en las que los dominios de activación transcripcional comprenden otros dominios de activación transcripcional o fragmentos funcionales o variantes de los mismos.

5 El uno o más dominios de activación transcripcional pueden fusionarse directamente al adaptador. Alternativamente, el uno o más dominios de activación transcripcional pueden unirse al adaptador mediante un conector o una combinación de conectores o mediante uno o más dominios adicionales. Asimismo, si están presentes dos o más dominios de activación transcripcional, pueden fusionarse directamente entre sí o pueden unirse entre sí mediante un conector o una combinación de conectores o mediante uno o más dominios adicionales. Los conectores que pueden usarse en estas proteínas de fusión pueden incluir cualquier secuencia que no interfiera con la función de las proteínas de fusión. Los conectores a modo de ejemplo son cortos (por ejemplo, 2-20 aminoácidos) y normalmente son flexibles (por ejemplo, comprenden aminoácidos con un alto grado de libertad, tales como glicina, alanina y serina).

15 El uno o más dominios de activación transcripcional y el adaptador pueden estar en cualquier orden dentro de la proteína adaptadora quimérica. Como una opción, el uno o más dominios de activación transcripcional puede estar carboxiterminal al adaptador y el adaptador puede estar aminoterminal al uno o más dominios de activación transcripcional. Por ejemplo, el uno o más dominios de activación transcripcional pueden estar en el extremo C de la proteína adaptadora quimérica, y el adaptador puede estar en el extremo N de la proteína adaptadora quimérica. Sin embargo, el uno o más dominios de activación transcripcional pueden estar carboxiterminal al adaptador sin estar en el extremo C de la proteína adaptadora quimérica (por ejemplo, si una señal de localización nuclear está en el extremo C de la proteína adaptadora quimérica). Asimismo, el adaptador puede estar aminoterminal a uno o más dominios de activación transcripcional sin estar en el extremo N de la proteína adaptadora quimérica (por ejemplo, si una señal de localización nuclear está en el extremo N de la proteína adaptadora quimérica). Como otra opción, el uno o más dominios de activación transcripcional puede estar aminoterminal al adaptador y el adaptador puede estar carboxiterminal al uno o más dominios de activación transcripcional. Por ejemplo, uno o más dominios de activación transcripcional pueden estar en el extremo N de la proteína adaptadora quimérica, y el adaptador puede estar en el extremo C de la proteína adaptadora quimérica. Como otra opción más, si la proteína adaptadora quimérica comprende dos o más dominios de activación transcripcional, los dos o más dominios de activación transcripcional pueden flanquear el adaptador.

Las proteínas adaptadoras quiméricas también pueden unirse operativamente o fusionarse a polipéptidos heterólogos adicionales. El polipéptido heterólogo fusionado o unido puede localizarse en el extremo N, en el extremo C o en cualquier parte interna de la proteína adaptadora quimérica. Por ejemplo, una proteína adaptadora quimérica puede comprender además una señal de localización nuclear. Un ejemplo específico de dicha proteína comprende una proteína de la cubierta de MS2 (adaptadora) unida (directamente o mediante una NLS) a un dominio de activación transcripcional p65 carboxiterminal a la proteína de la cubierta de MS2 (MCP), y un dominio de activación transcripcional HSF1 carboxiterminal al dominio de activación transcripcional p65. Dicha proteína puede comprender de extremo N a extremo C: una MCP; una señal de localización nuclear; un dominio de activación transcripcional p65; y un dominio de activación transcripcional HSF1. Por ejemplo, una proteína adaptadora quimérica puede comprender, consistir esencialmente en, o consistir en una secuencia de aminoácidos al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a la secuencia de proteína adaptadora quimérica MCP-p65-HSF1 establecida en SEQ ID NO: 37. Similarmente, un ácido nucleico que codifica una proteína adaptadora quimérica puede comprender, consistir esencialmente en, o consistir en una secuencia al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a la secuencia codificante de la proteína adaptadora quimérica MCP-p65-HSF1 establecida en SEQ ID NO: 39.

Los adaptadores (es decir, los dominios adaptadores o las proteínas adaptadoras) son dominios de unión a ácidos nucleicos (por ejemplo, dominios de unión a ADN y/o dominios de unión a ARN) que reconocen y se unen específicamente a secuencias distintas (por ejemplo, se unen a secuencias distintas de ADN y/o ARN, tales como aptámeros, de forma específica de secuencia). Los aptámeros incluyen ácidos nucleicos que, gracias a su capacidad para adoptar una conformación tridimensional específica, pueden unirse a una molécula diana con gran afinidad y especificidad. Dichos adaptadores pueden unirse, por ejemplo, a una secuencia de ARN y una estructura secundaria específicas. Estas secuencias (es decir, los elementos de unión al adaptador) se pueden manipular en un ARN guía. Por ejemplo, un aptámero MS2 puede manipularse en un ARN guía para unirse específicamente a una proteína de la cubierta de MS2 (MCP). Por ejemplo, el adaptador puede comprender, consistir esencialmente en, o consistir en una secuencia de aminoácidos al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a la secuencia de MCP establecida en SEQ ID NO: 40. Similarmente, un ácido nucleico que codifica el adaptador puede comprender, consistir esencialmente en, o consistir en una secuencia de aminoácidos al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a la secuencia codificante de MCP establecida en SEQ ID NO: 41. Ejemplos específicos de adaptadores y dianas incluyen, por ejemplo, combinaciones de proteínas de unión a ARN/aptámeros que existen dentro de la diversidad de proteínas de la cubierta de bacteriófagos. Véanse, por ejemplo, los documentos de patente US 2019-0284572 y WO 2019/183123.

65 Las proteínas adaptadoras quiméricas desveladas en el presente documento comprenden uno o más dominios de activación transcripcional. Dichos dominios de activación transcripcional pueden ser dominios de activación

transcripcional naturales, pueden ser fragmentos funcionales o variantes funcionales de dominios de activación transcripcional naturales, o pueden ser dominios de activación transcripcional manipulados o sintéticos. Los dominios de activación transcripcional que pueden usarse incluyen los descritos, por ejemplo, en los documentos de patente US 2019-0284572 y WO 2019/183123.

5

4. Tipos de células

Las células biosensoras Cas/tau desveladas en el presente documento pueden ser cualquier tipo de célula de mamífero y pueden ser *in vitro*, *ex vivo* o *in vivo*. Una línea celular o población de células biosensoras Cas/tau puede ser una línea celular o población de células monoclonal. Similarmente, las células biosensoras SAM/tau desveladas en el presente documento pueden ser cualquier tipo de célula de mamífero y pueden ser *in vitro*, *ex vivo* o *in vivo*. Una línea celular o población de células biosensoras de SAM/tau puede ser una línea celular o población de células monoclonal. La célula puede proceder de cualquier fuente. Por ejemplo, dichas células pueden ser células humanas o células de mamífero no humanos, tales como células de roedor, células de ratón o células de rata. Los mamíferos incluyen, por ejemplo, humanos, primates no humanos, monos, simios superiores, gatos, perros, caballos, toros, ciervos, bisontes, ovejas, roedores (por ejemplo, ratones, ratas, hámsteres, cobayas), ganado (por ejemplo, especies bovinas, tales como vacas y bueyes; especies ovinas, tales como ovejas y cabras; y especies porcinas, tales como cerdos y jabalíes). También se incluyen los animales domésticos y los animales agrícolas. El término "animal no humano" excluye a los seres humanos. En un ejemplo específico, las células biosensoras Cas/tau son células humanas (por ejemplo, células HEK293T). Similarmente, en un ejemplo específico, las células biosensoras SAM/tau son células humanas (por ejemplo, células HEK293T).

La célula puede ser, por ejemplo, una célula totipotente no humana o una célula pluripotente (por ejemplo, una célula madre embrionaria (ES), tal como una célula ES de roedor, una célula ES de ratón o una célula ES de rata). Las células totipotentes incluyen células indiferenciadas que pueden dar lugar a cualquier tipo de célula, y las células pluripotentes incluyen células indiferenciadas que poseen la capacidad de convertirse en más de un tipo celular diferenciado. Dichas células pluripotentes y/o totipotentes no humanas pueden ser, por ejemplo, células ES o células similares a ES, tales como las células madre pluripotentes inducidas (iPS). Las células ES incluyen células no humanas totipotentes o pluripotentes derivadas de embriones que son capaces de contribuir a cualquier tejido del embrión en desarrollo tras su introducción en un embrión. Solo para referencia, las células ES pueden derivar de la masa celular interna de un blastocisto y son capaces de diferenciarse en células de cualquiera de las tres capas germinales de los vertebrados (endodermo, ectodermo y mesodermo).

La célula también puede ser una célula somática primaria o una célula que no sea una célula somática primaria. Las células somáticas pueden incluir cualquier célula que no sea un gameto, célula germinal, gametocito o célula madre indiferenciada. La célula también puede ser una célula primaria. Las células primarias incluyen células o cultivos de células que han sido aisladas directamente de un organismo, órgano o tejido. Las células primarias incluyen células que no son ni transformadas ni inmortales. Incluyen cualquier célula obtenida de un organismo, órgano o tejido que no haya pasado previamente por un cultivo de tejidos o que haya pasado previamente por un cultivo de tejidos pero que sea incapaz de ser pasada indefinidamente por un cultivo de tejidos. Dichas células pueden aislarse mediante técnicas convencionales e incluyen, por ejemplo, células somáticas, células hematopoyéticas, células endoteliales, células epiteliales, fibroblastos, células mesenquimatosas, queratinocitos, melanocitos, monocitos, células mononucleares, adipocitos, preadipocitos, neuronas, neuroglíocitos, hepatocitos, mioblastos esqueléticos y células de músculo liso. Por ejemplo, las células primarias pueden derivar de tejidos conjuntivos, tejidos musculares, tejidos del sistema nervioso o tejidos epiteliales.

Dichas células también incluyen las que normalmente no proliferarían indefinidamente pero que, debido a una mutación o alteración, han eludido la senescencia celular normal y, en su lugar, pueden seguir sometándose a división. Dichas mutaciones o alteraciones pueden producirse de forma natural o ser inducidas intencionadamente. Los ejemplos de células inmortalizadas incluyen las células de ovario de hámster chino (CHO), las células de riñón embrionario humano (por ejemplo, las células HEK293T) y las células de fibroblastos embrionarios de ratón (por ejemplo, las células 3T3). Son bien conocidos numerosos tipos de células inmortalizadas. Las células inmortalizadas o primarias incluyen células que se usan normalmente para el cultivo o para la expresión de genes o proteínas recombinantes.

La célula también puede ser una célula diferenciada, tal como una célula neuronal (por ejemplo, una célula neuronal humana).

B. Métodos de generación de células biosensoras Cas/Tau y células biosensoras SAM/Tau

Las células biosensoras Cas/tau desveladas en el presente documento pueden generarse por cualquier medio conocido. El primer dominio de repetición de tau unido al primer indicador, el segundo dominio de repetición de tau unido al segundo indicador y la proteína Cas pueden introducirse en la célula en cualquier forma (por ejemplo, ADN, ARN o proteína) por cualquier medio conocido. Similarmente, las células biosensoras SAM/tau desveladas en el presente documento pueden generarse por cualquier medio conocido. El primer dominio de repetición de tau unido al primer indicador, el segundo dominio de repetición de tau unido al segundo indicador, la proteína Cas quimérica y la

proteína adaptadora quimérica pueden introducirse en la célula en cualquier forma (por ejemplo, ADN, ARN o proteína) por cualquier medio conocido. "Introducir" incluye presentar a la célula el ácido nucleico o la proteína de tal manera que la secuencia acceda al interior de la célula. Los métodos proporcionados en el presente documento no dependen de un método particular para introducir un ácido nucleico o una proteína en la célula, solo de que el ácido nucleico o la proteína accedan al interior de al menos una célula. Los métodos para introducir ácidos nucleicos y proteínas en diversos tipos de células son conocidos e incluyen, por ejemplo, métodos de transfección estable, métodos de transfección transitoria y métodos mediados por virus. Opcionalmente, se pueden usar vectores dirigidos.

Los protocolos de transfección, así, tal como los protocolos para introducir ácidos nucleicos o proteínas en las células, pueden variar. Los métodos de transfección no limitantes incluyen métodos de transfección basados en productos químicos que usan liposomas; nanopartículas; fosfato de calcio (Graham et al. (1973) *Virology* 52 (2): 456-67, Bacchetti et al. (1977) *Proc. Natl. Acad. Sci. USA* 74 (4): 1590-4 y Kriegler, M (1991). *Transfer and Expression: A Laboratory Manual*. Nueva York: W. H. Freeman and Company. pp. 96-97); dendrímeros; o polímeros catiónicos, tales como DEAE-dextrano o polietilénimina. Los métodos no químicos incluyen electroporación, sonoporación y transfección óptica. La transfección basada en partículas incluye el uso de una pistola de genes o la transfección asistida por imanes (Bertram (2006) *Current Pharmaceutical Biotechnology* 7, 277-28). También pueden usarse métodos víricos para la transfección.

La introducción de ácidos nucleicos o proteínas en una célula también puede estar mediada por electroporación, por inyección intracitoplasmática, por infección vírica, por adenovirus, por virus adenoasociado, por lentivirus, por retrovirus, por transfección, por transfección mediada por lípidos o por nucleofección. La nucleofección es una tecnología de electroporación mejorada que permite administrar sustratos de ácido nucleico no solo al citoplasma, sino también a través de la membrana nuclear hasta el núcleo. Además, el uso de la nucleofección en los métodos desvelados en el presente documento requiere normalmente muchas menos células que la electroporación normal (por ejemplo, solo aproximadamente 2 millones en comparación con los 7 millones de la electroporación normal). En un ejemplo, la nucleofección se realiza usando el sistema LONZA® NUCLEOFECTOR™.

La introducción de ácidos nucleicos o proteínas en una célula también puede realizarse mediante microinyección. La microinyección de un ARNm se realiza preferentemente en el citoplasma (por ejemplo, para administrar el ARNm directamente a la maquinaria de traducción), mientras que la microinyección de una proteína o de un ADN que codifica una proteína es preferentemente en el núcleo. Alternativamente, la microinyección puede llevarse a cabo mediante inyección tanto en el núcleo como en el citoplasma: primero puede introducirse una aguja en el núcleo e inyectarse una primera cantidad, y mientras se retira la aguja de la célula puede inyectarse una segunda cantidad en el citoplasma. Los métodos para llevar a cabo la microinyección son bien conocidos. Véase, por ejemplo, Nagy et al. (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003, *Manipulating the Mouse Embryo*. Cold Spring Harbor, Nueva York: Cold Spring Harbor Laboratory Press); Meyer et al. (2010) *Proc. Natl. Acad. Sci. USA* 107:15022-15026 y Meyer et al. (2012) *Proc. Natl. Acad. Sci. USA* 109:9354-9359.

Otros métodos para introducir ácidos nucleicos o proteínas en una célula pueden incluir, por ejemplo, la administración de vectores, la administración mediada por partículas, la administración mediada por exosomas, la administración mediada por nanopartículas lipídicas, la administración mediada por péptidos de penetración celular o la administración mediada por dispositivos implantables. En el presente documento se describen en otra parte métodos de administración de ácidos nucleicos o proteínas a un sujeto para modificar células *in vivo*.

En un ejemplo, el primer dominio de repetición de tau unido al primer indicador, el segundo dominio de repetición de tau unido al segundo indicador y la proteína Cas pueden introducirse mediante transducción vírica, tal como la transducción lentivírica.

El cribado de células que comprenden el primer dominio de repetición de tau unido al primer indicador, el segundo dominio de repetición de tau unido al segundo indicador y la proteína Cas puede realizarse por cualquier medio conocido.

Como un ejemplo, pueden usarse genes indicadores para detectar células que tengan la proteína Cas, el primer dominio de repetición de tau unido al primer indicador, o el segundo dominio de repetición de tau unido al segundo indicador. Los genes indicadores fluorescentes a modo de ejemplo incluyen los que codifican la proteína fluorescente verde (GFP), la proteína fluorescente verde mejorada (eGFP), la proteína fluorescente cian (CFP), la proteína fluorescente amarilla (YFP), la proteína fluorescente amarilla mejorada (eYFP), la proteína fluorescente azul (BFP), la proteína fluorescente azul mejorada (eBFP), DsRed, ZsGreen, MmGFP, mPlum, mCherry, tdTomato, mStrawberry, J-Red, mOrange, mKO, mCitrine, Venus, YPet, Emerald, CyPet, Cerulean y T-Sapphire. Dado que el primer indicador y el segundo indicador son proteínas fluorescentes (por ejemplo, CFP e YFP), las células que contienen estos indicadores pueden seleccionarse por citometría de flujo para seleccionar células doblemente positivas. A continuación, las células doblemente positivas pueden combinarse para generar una línea policlonal, o pueden generarse líneas monoclonales a partir de células doblemente positivas individuales.

Como otro ejemplo, pueden usarse marcadores de selección para cribar células que tienen la proteína Cas, el primer dominio de repetición de tau unido al primer indicador, o el segundo dominio de repetición de tau unido al segundo indicador. Algunos marcadores de selección a modo de ejemplo son la neomicina fosfotransferasa (neo^r), la

higromicina B fosfotransferasa (*hyg^r*), la puromicina-N-acetiltransferasa (*puro^r*), la blastidina S desaminasa (*bsr^r*), la xantina/guanina fosforribosil transferasa (*gpt*), o la timidina cinasa del virus del herpes simple (HSV-k). Otro marcador de selección a modo de ejemplo es la proteína de resistencia a la bleomicina, codificada por el gen *Sh ble* (gen de la bleomicina de *Streptoalloteichus hindustanus*), que confiere resistencia a la zeocina (fleomicina D1).

5 Las células positivas a la agregación (Ag[+]) en las que el dominio de repetición de tau se presenta de forma estable en un estado agregado, lo que significa que los agregados del dominio de repetición de tau persisten de forma estable en todas las células con crecimiento y múltiples pasajes a lo largo del tiempo, pueden generarse usando medio acondicionado obtenido de células cultivadas positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, tal como se describe en el presente documento. Esta es la "siembra mínima" desvelada en los ejemplos del presente documento. Medio acondicionado se refiere al medio gastado obtenido de células cultivadas. Contiene metabolitos, factores de crecimiento y proteínas de matriz extracelular secretadas al medio por las células cultivadas. El uso de medio acondicionado no implica el cocultivo con células Ag[+] (es decir, las células Ag[-] sin tratamiento previo que no se cocultivan con células Ag[+]). Con un ejemplo, el medio acondicionado puede generarse obteniendo medio que haya estado en células Ag[+] confluentes. El medio puede haber estado en las células Ag[+] confluentes durante 1 a 7 días, por ejemplo, aproximadamente 2 a aproximadamente 6, aproximadamente 3 a aproximadamente 5 o aproximadamente 4 días. A continuación, el medio acondicionado puede aplicarse a las células biosensoras Cas/tau sin tratamiento previo (Ag[-]) en combinación con medio recién preparado. Similarmente, el medio acondicionado puede aplicarse a continuación a las células biosensoras SAM/tau sin tratamiento previo (Ag[-]) en combinación con medio recién preparado. La proporción entre medio acondicionado y medio recién preparado puede ser, por ejemplo, aproximadamente 10:1, aproximadamente 9:1, aproximadamente 8:1, aproximadamente 7:1, aproximadamente 6:1, aproximadamente 5:1, aproximadamente 4:1, aproximadamente 3:1, aproximadamente 2:1, aproximadamente 1:1, aproximadamente 1:2, aproximadamente 1:3, aproximadamente 1:4, aproximadamente 1:5, aproximadamente 1:6, aproximadamente 1:7, aproximadamente 1:8, aproximadamente 1:9, o aproximadamente 1:10. Por ejemplo, la proporción de medio acondicionado de medio recién preparado puede ser de aproximadamente 5:1 a aproximadamente 1:1, de aproximadamente 4:1 a aproximadamente 2:1, o de aproximadamente 3:1. Por ejemplo, puede comprender cultivar la población de células modificada genéticamente en aproximadamente 90 % de medio acondicionado y aproximadamente 10 % de medio recién preparado, aproximadamente 85 % de medio acondicionado y aproximadamente 15 % de medio recién preparado, aproximadamente 80 % de medio acondicionado y aproximadamente 20 % de medio recién preparado, aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio recién preparado, aproximadamente 70 % de medio acondicionado y aproximadamente 30 % de medio recién preparado, aproximadamente 65 % de medio acondicionado y aproximadamente 35 % de medio recién preparado, aproximadamente 60 % de medio acondicionado y aproximadamente 40 % de medio recién preparado, aproximadamente 55 % de medio acondicionado y aproximadamente 45 % de medio recién preparado, aproximadamente 50 % de medio acondicionado y aproximadamente 50 % de medio recién preparado, aproximadamente 45 % de medio acondicionado y aproximadamente 55 % de medio recién preparado, aproximadamente 40 % de medio acondicionado y aproximadamente 60 % de medio recién preparado, aproximadamente 35 % de medio acondicionado y aproximadamente 65 % de medio recién preparado, aproximadamente 30 % de medio acondicionado y aproximadamente 70 % de medio recién preparado, aproximadamente 25 % de medio acondicionado y aproximadamente 75 % de medio recién preparado, aproximadamente 20 % de medio acondicionado y aproximadamente 80 % de medio recién preparado, aproximadamente 15 % de medio acondicionado y aproximadamente 85 % de medio recién preparado, o aproximadamente 10 % de medio acondicionado y aproximadamente 90 % de medio recién preparado. En un ejemplo, puede comprender cultivar la población de células modificada genéticamente en un medio que comprende al menos aproximadamente 50 % de medio acondicionado y no más de aproximadamente 50 % de medio recién preparado. En un ejemplo específico, puede comprender cultivar la población de células modificada genéticamente en aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio recién preparado. Opcionalmente, el medio acondicionado se aplica a las células Ag[-] sin tratamiento previo sin lipofectamina o sin liposomas (por ejemplo, liposomas catiónicos) o sin fosfolípidos. Opcionalmente, la población de células modificada genéticamente no se cocultiva con las células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado.

El medio acondicionado sin cocultivo no se había usado antes en este contexto como agente de siembra. Sin embargo, el medio acondicionado es especialmente útil para los cribados de todo el genoma a gran escala, ya que las fibrillas tau producidas *in vitro* son un recurso limitado. Además, el medio acondicionado es más relevante fisiológicamente debido a que es producido y secretado por células en vez de *in vitro*. El uso de medio acondicionado, tal como se describe en el presente documento, proporciona un refuerzo de la actividad de siembra de tau (por ejemplo, ~0,1 % como se mide por inducción FRET como se desvela en otra parte del presente documento) para sensibilizar las células a la agregación de tau.

III. Bibliotecas inactivadas de ARN guía

Los métodos de cribado CRISPRn descritos en el presente documento hacen uso de bibliotecas inactivadas de ARN guía (ARNg) CRISPR, tales como bibliotecas inactivadas de ARNg de todo el genoma. Las nucleasas Cas, tales como Cas9, pueden programarse para inducir roturas de doble cadena en loci genómicos específicos mediante ARNg

diseñados para dirigirse a secuencias diana específicas. Dado que la especificidad de direccionamiento de las proteínas Cas se confiere mediante ARNg cortos, es posible el cribado funcional a escala genómica agrupada. Dichas bibliotecas tienen varias ventajas sobre bibliotecas, tales como bibliotecas de ARNh_c, que reducen la expresión de proteínas al dirigirse al ARNm. En cambio, las bibliotecas de ARNg consiguen la inactivación mediante mutaciones de desplazamiento de marco introducidas en las regiones codificantes genómicas de los genes.

Los métodos de cribado CRISPRa desvelados en el presente documento hacen uso de bibliotecas de activación transcripcional de ARN guía (ARNg) CRISPR, tales como las bibliotecas de activación transcripcional de ARNg de todo el genoma. Los sistemas SAM pueden programarse para activar la transcripción de genes en loci genómicos específicos mediante ARNg diseñados para dirigirse a secuencias diana específicas. Dado que la especificidad de direccionamiento de las proteínas Cas se confiere mediante ARNg cortos, es posible el cribado funcional a escala genómica agrupada.

Los ARNg de una biblioteca pueden dirigirse a cualquier número de genes. Por ejemplo, los ARNg pueden dirigirse a aproximadamente 50 genes o más, aproximadamente 100 genes o más, aproximadamente 200 genes o más, aproximadamente 300 genes o más, aproximadamente 400 genes o más, aproximadamente 500 genes o más, aproximadamente 1000 genes o más, aproximadamente 2000 genes o más, aproximadamente 3000 genes o más, aproximadamente 4000 genes o más, aproximadamente 5000 genes o más, aproximadamente 10000 genes o más, o aproximadamente 20000 genes o más. En algunas bibliotecas, los ARNg pueden seleccionarse para dirigirse a genes de una vía de señalización particular. Algunas bibliotecas son bibliotecas de todo el genoma.

Las bibliotecas de todo el genoma incluyen uno o más ARNg (por ejemplo, ARNgu) dirigidos a cada gen del genoma. El genoma objetivo puede ser de cualquier tipo de genoma. Por ejemplo, el genoma puede ser un genoma eucariota, un genoma de mamífero, un genoma de mamífero no humano, un genoma de roedor, un genoma de ratón, un genoma de rata o un genoma humano. En un ejemplo, el genoma diana es un genoma humano.

Los ARNg pueden dirigirse a cualquier número de secuencias dentro de cada gen diana individual. En algunas bibliotecas, una pluralidad de secuencias diana son diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, aproximadamente 2 a aproximadamente 10, aproximadamente 2 a aproximadamente 9, aproximadamente 2 a aproximadamente 8, aproximadamente 2 a aproximadamente 7, aproximadamente 2 a aproximadamente 6, aproximadamente 2 a aproximadamente 5, aproximadamente 2 a aproximadamente 4 o aproximadamente 2 a aproximadamente 3 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5 o al menos aproximadamente 6 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Como ejemplo específico, aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana. Como otro ejemplo específico, aproximadamente 3 a aproximadamente 6 o aproximadamente 4 a aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana.

Por ejemplo, las bibliotecas pueden dirigirse a genes con una cobertura promedio de aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9 o aproximadamente 10 ARNg por gen. En un ejemplo específico, una biblioteca puede dirigirse a genes con una cobertura promedio de aproximadamente 3-4 ARNg por gen o aproximadamente 6 ARNg por gen.

Los ARNg pueden dirigirse a cualquier localización deseada en los genes diana. Los ARNg CRISPRn pueden diseñarse para dirigirse a regiones codificantes de genes, de modo que el corte por la proteína Cas correspondiente provoque mutaciones por inserción/delección (indel) con desplazamiento de marco que den lugar a un alelo con pérdida de función. Más específicamente, las mutaciones con desplazamiento de marco pueden conseguirse mediante roturas selectivas de la doble cadena de ADN y su posterior reparación mutagénica a través de la vía de unión de extremos no homólogos (NHEJ), que produce indels en el lugar de la rotura. El indel que se introduce en el DSB es aleatorio, y algunos indels dan lugar a mutaciones con desplazamiento de marco que provocan la terminación prematura del gen.

En algunas bibliotecas CRISPRn, cada ARNg se dirige a un exón constitutivo si es posible. En algunas bibliotecas CRISPRn, cada ARNg se dirige a un exón constitutivo en 5' si es posible. En algunos métodos, cada ARNg se dirige a un primer exón, un segundo exón o un tercer exón (desde el extremo 5' del gen) si es posible.

Como un ejemplo, los ARNg de la biblioteca CRISPRn pueden dirigirse a exones constitutivos. Los exones constitutivos son exones que se conservan de forma consistente tras el corte y empalme. Los exones que se expresan en todos los tejidos pueden considerarse exones constitutivos para los ARNg diana. Los ARNg de la biblioteca pueden dirigirse a exones constitutivos cerca del extremo 5' de cada gen. Opcionalmente, el primer y último exones de cada gen pueden excluirse como posibles dianas. Opcionalmente, cualquier exón que contenga un sitio de corte y empalme alternativo puede ser excluido como posibles dianas. Opcionalmente, los dos primeros exones que cumplan los criterios anteriores se seleccionan como posibles dianas. Opcionalmente, los exones 2 y 3 se seleccionan como posibles dianas (por ejemplo, si no se identifican exones constitutivos). Además, los ARNg de la biblioteca pueden seleccionarse y diseñarse para minimizar los efectos inespecíficos.

En un ejemplo específico, la biblioteca o bibliotecas de ARNg CRISPRn de todo el genoma comprenden ARNgu dirigidos a exones constitutivos 5' de > 18.000 genes del genoma humano con una cobertura promedio de ~6 ARNgu por gen, y cada sitio diana se seleccionó para minimizar la modificación inespecífica.

Los ARNg CRISPRa pueden diseñarse para dirigirse a secuencias adyacentes al sitio de inicio de la transcripción de un gen. Por ejemplo, la secuencia diana puede estar dentro de 1000, 900, 800, 700, 600, 500, 400, 300, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5 o 1 par de bases del sitio de inicio de la transcripción. Por ejemplo, cada ARNg de la biblioteca CRISPRa puede dirigirse a una secuencia dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción. Opcionalmente, la secuencia diana está dentro de la región de 200 pares de bases aguas arriba del sitio de inicio de la transcripción y 1 par de bases aguas abajo del sitio de inicio de la transcripción (-200 a +1).

Los ARNg de la biblioteca de todo el genoma pueden tener cualquier forma. Por ejemplo, la biblioteca de ARNg puede encapsidarse en un vector vírico, tal como vectores retrovíricos, vectores lentivíricos o vectores adenovíricos. En un ejemplo específico, la biblioteca de ARNg se encapsida en vectores lentivíricos. Los vectores pueden comprender además genes indicadores o marcadores de selección para facilitar la selección de las células que reciben los vectores. Los ejemplos de dichos genes indicadores y marcadores de selección se describen en el presente documento en otra parte. Como un ejemplo, el marcador de selección puede ser uno que imparta resistencia a un fármaco, tal como neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y blasticidina S desaminasa. Otro marcador de selección a modo de ejemplo es la proteína de resistencia a la bleomicina, codificada por el gen *Sh ble* (gen de la bleomicina de *Streptoalloteichus hindustanus*), que confiere resistencia a la zeocina (fleomicina D1). Por ejemplo, las células pueden seleccionarse con un fármaco (por ejemplo, puromicina) de modo que solo las células transducidas con una construcción de ARN guía se conserven para ser usadas para llevar a cabo el cribado.

A. ARN guía

Un "ARN guía" o "ARNg" es una molécula de ARN que se une a una proteína Cas (por ejemplo, la proteína Cas9) y dirige la proteína Cas a una localización específica dentro de un ADN diana. Los ARN guía pueden comprender dos segmentos: un "segmento que se dirige al ADN" y un "segmento que se une a la proteína". "Segmento" incluye una sección o región de una molécula, tal como un tramo contiguo de nucleótidos en un ARN. Algunos ARNg, tales como los de Cas9, pueden comprender dos moléculas de ARN separadas: un "ARN-activador" (por ejemplo, ARNtracr) y un "ARN-destinatario" (por ejemplo, ARN CRISPR o ARNcr). Otros ARNg son una única molécula de ARN (polinucleótido de ARN único), que también puede denominarse "ARNg de molécula única", "ARN guía único" o "ARNgu". Véanse, por ejemplo, los documentos de patente WO 2013/176772, WO 2014/065596, WO 2014/089290, WO 2014/093622, WO 2014/099750, WO 2013/142578 y WO 2014/131833. Para Cas9, por ejemplo, un ARN de guía único puede comprender un ARNcr fusionado a un ARNtracr (por ejemplo, a través de un conector). Para Cpf1, por ejemplo, solo se necesita un ARNcr para lograr la unión a una secuencia diana. Los términos "ARN guía" y "ARNg" incluyen tanto los ARNg de molécula doble (es decir, modulares) como los ARNg de molécula única.

Un ARNg de dos moléculas a modo de ejemplo comprende una molécula similar a un ARNcr ("ARN CRISPR" o "ARN-destinatario" o "ARNcr" o "ARNcr repetido") y una molécula correspondiente similar a un ARNtracr ("ARN CRISPR trans-activo" o "ARN-activador" o "ARNtracr"). Un ARNcr comprende tanto el segmento de ADN diana (monocatenario) del ARNg, tal como un tramo de nucleótidos que forma una mitad del dúplex de ARNbc del segmento de unión a proteínas del ARNg. Un ejemplo de cola de ARNcr, situada aguas abajo (3') del segmento dirigido al ADN comprende, consiste esencialmente en, o consiste en GUUUUAGAGCUAUGCU (SEQ ID NO: 23). Cualquiera de los segmentos dirigidos al ADN descritos en el presente documento puede unirse al extremo 5' de SEQ ID NO: 23 para formar un ARNcr.

Un ARNtracr correspondiente (ARN-activador) comprende un tramo de nucleótidos que forma la otra mitad del dúplex de ARNbc del segmento de unión a proteínas del ARNg. Un tramo de nucleótidos de un ARNcr es complementario y se hibrida con un tramo de nucleótidos de un ARNtracr para formar el dúplex de ARNbc del dominio de unión a proteínas del ARNg. Como tal, se puede decir que cada ARNcr tiene un ARNtracr correspondiente. Un ejemplo de secuencia de ARNtracr comprende, consiste esencialmente en, o consiste en AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUU (SEQ ID NO: 24). Otros ejemplos de secuencias de tracrRNA comprenden, consisten esencialmente en, o consisten en cualquiera de

AAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG

CACCGAGUCGGUGCUUUU (SEQ ID NO: 28),

o

GUUGGAACCAUUCAAAACAGCAUAGCAAGUAAAAUAAGGCCUAGUCCGUUAUCA

ACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 29).

5 En los sistemas en los que se necesitan tanto un ARNcr como un ARNtracr, el ARNcr y el ARNtra correspondiente se hibridan para formar un ARNg. En los sistemas en los que solo se necesita un ARNcr, el ARNcr puede ser el ARNg. El ARNcr proporciona además el segmento de ADN monocatenario diana que se hibrida con la cadena complementaria de un ADN diana. Si se usa para la modificación dentro de una célula, la secuencia exacta de una molécula de ARNcr o ARNtracr dada puede diseñarse para que sea específica de la especie en la que se usarán las moléculas de ARN. Véanse, por ejemplo, Mali et al. (2013) Science 339:823-826; Jinek et al. (2012) Science 337:816-821; Hwang et al. (2013) Nat. Biotechnol. 31:227-229; Jiang et al. (2013) Nat. Biotechnol. 31:233-239; y Cong et al. (2013) Science 339:819-823.

10 El segmento dirigido al ADN (ARNcr) de un ARNg dado comprende una secuencia de nucleótidos que es complementaria a una secuencia de la cadena complementaria del ADN diana, como se describe con más detalle a continuación. El segmento dirigido al ADN de un ARNg interactúa con el ADN diana de una manera específica de la secuencia mediante hibridación (es decir, emparejamiento de bases). Como tal, la secuencia de nucleótidos del segmento dirigido al ADN puede variar y determina la ubicación dentro del ADN diana con la que interactuarán el ARNg y el ADN diana. El segmento de ADN diana de un ARNg objeto puede modificarse para hibridarse con cualquier secuencia deseada dentro de un ADN diana. Los ARNcr naturales difieren en función del sistema CRISPR/Cas y del organismo, pero a menudo contienen un segmento de direccionamiento de entre 21 y 72 nucleótidos de longitud, flanqueado por dos repeticiones directas (DR) de una longitud de entre 21 y 46 nucleótidos (véase, por ejemplo, el documento de patente WO 2014/131833). En el caso de *S. pyogenes*, las DR tienen una longitud de 36 nucleótidos y el segmento diana tiene 30 nucleótidos de longitud. La DR situada en 3' es complementaria y se hibrida con el ARNtracr correspondiente, que a su vez se une a la proteína Cas.

15 El segmento dirigido al ADN puede tener, por ejemplo, una longitud de al menos aproximadamente 12, 15, 17, 18, 19, 20, 25, 30, 35 o 40 nucleótidos. Dichos segmentos dirigidos al ADN pueden tener, por ejemplo, una longitud de aproximadamente 12 a aproximadamente 100, de aproximadamente 12 a aproximadamente 80, de aproximadamente 12 a aproximadamente 50, de aproximadamente 12 a aproximadamente 40, de aproximadamente 12 a aproximadamente 30, de aproximadamente 12 a aproximadamente 25, o de aproximadamente 12 a aproximadamente 20 nucleótidos. Por ejemplo, el segmento dirigido al ADN diana puede ser de aproximadamente 15 a aproximadamente 25 nucleótidos (por ejemplo, de aproximadamente 17 a aproximadamente 20 nucleótidos, o de aproximadamente 17, 18, 19 o 20 nucleótidos). Véase, por ejemplo, el documento de patente US 2016/0024523. Para Cas9 de *S. pyogenes*, un segmento dirigido al ADN típico tiene entre 16 y 20 nucleótidos de longitud o entre 17 y 20 nucleótidos de longitud. En el caso de Cas9 de *S. aureus*, el segmento dirigido al ADN típico tiene entre 21 y 23 nucleótidos de longitud. En el caso de Cpf1, un segmento dirigido al ADN típico tiene al menos 16 nucleótidos de longitud o al menos 18 nucleótidos de longitud.

20 Los ARNtracr pueden tener cualquier forma (por ejemplo, ARNtracr de longitud completa o ARNtracr parciales activos) y longitudes variables. Pueden incluir transcritos primarios o formas procesadas. Por ejemplo, los ARNtracr (como parte de un ARN guía único o como una molécula separada como parte de un ARNg de dos moléculas) pueden comprender, consistir esencialmente en, o consistir en la totalidad o una parte de una secuencia de ARNtracr natural (por ejemplo, aproximadamente o más de aproximadamente 20, 26, 32, 45, 48, 54, 63, 67, 85, o más nucleótidos de una secuencia de ARNtracr natural). Ejemplos de secuencias de tracrRNA naturales de *S. pyogenes* incluyen versiones de 171 nucleótidos, 89 nucleótidos, 75 nucleótidos y 65 nucleótidos. Véanse, por ejemplo, Deltcheva et al. (2011) Nature 471:602-607; documento de patente WO 2014/093661. Ejemplos de ARNtracr dentro de ARN guía único (ARNgu) incluyen los segmentos de ARNtracr que se encuentran dentro de las versiones +48, +54, +67 y +85 de los ARNgu, donde "+n" indica que hasta el nucleótido +n de ARNtracr natural está incluido en el ARNgu. Véase el documento de patente US 8.697.359.

25 El porcentaje de complementariedad entre el segmento dirigido al ADN del ARN guía y la cadena complementaria del ADN diana puede ser de al menos el 60 % (por ejemplo, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 97 %, al menos el 98 %, al menos el 99 % o el 100 %). El porcentaje de complementariedad entre el segmento dirigido al ADN y la cadena complementaria del ADN diana puede ser de al menos el 60 % en aproximadamente 20 nucleótidos contiguos. Como un ejemplo, el porcentaje de complementariedad entre el segmento dirigido al ADN y la cadena complementaria del ADN diana puede ser del 100 % en los 14 nucleótidos contiguos en el extremo 5' de la cadena complementaria del ADN diana y tan bajo como el 0 % en el resto. En tal caso, puede considerarse que el segmento dirigido al ADN tiene 14 nucleótidos de longitud. Como otro ejemplo, el porcentaje de complementariedad entre el segmento dirigido al ADN y la cadena complementaria del ADN diana puede ser del 100 % en los siete nucleótidos contiguos en el extremo 5' de la cadena complementaria del ADN diana y tan bajo como el 0 % en el resto. En tal caso, puede considerarse que el segmento dirigido al ADN tiene 7 nucleótidos de longitud. En algunos ARN guía, al menos 17 nucleótidos del segmento dirigido al ADN son complementarios a la cadena complementaria del ADN diana. Por ejemplo, el segmento dirigido al ADN puede tener 20 nucleótidos de longitud y puede incluir 1, 2 o 3 errores de emparejamiento con la cadena complementaria del ADN diana. En un ejemplo, los errores de emparejamiento no son adyacentes a la región

5 de la cadena complementaria correspondiente a la secuencia del motivo adyacente al protoespaciador (PAM) (es decir, el complemento inverso de la secuencia PAM) (por ejemplo, los errores de emparejamiento se encuentran en el extremo 5' del segmento dirigido al ADN del ARN guía, o los errores de emparejamiento se encuentran al menos a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o 19 pares de bases de distancia de la región de la cadena complementaria correspondiente a la secuencia PAM).

10 El segmento de unión a proteínas de un ARNg puede comprender dos tramos de nucleótidos que son complementarios entre sí. Los nucleótidos complementarios del segmento de unión a proteínas se hibridan para formar un dúplex de ARN bicatenario (ARNbc). El segmento de unión a proteínas de un ARNg sujeto interactúa con una proteína Cas, y el ARNg dirige la proteína Cas unida a una secuencia de nucleótidos específica dentro del ADN diana a través del segmento dirigido al ADN.

15 Los ARN guía únicos pueden incluir un segmento dirigido a ADN y una secuencia de armazón (es decir, la secuencia de unión a proteínas o a Cas del ARN guía). Por ejemplo, dichos ARN guía pueden tener un segmento 5' dirigido al ADN unido a una secuencia de armazón en 3'. Las secuencias de armazón a modo de ejemplo comprenden, consisten esencialmente en, o consisten en:

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA
AAAAGUGGCACCGAGUCGGUGCU (versión 1; SEQ ID NO: 17);

GUUGGAACCAUUCAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCA
ACUUGAAAAAGUGGCACCGAGUCGGUGC (versión 2; SEQ ID NO: 18);

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA
AAAAGUGGCACCGAGUCGGUGCU (versión 3; SEQ ID NO: 19);

20 y

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU (versión 4; SEQ ID NO: 20).

25 Las secuencias de armazón a modo de ejemplo comprenden, consisten esencialmente en, o consisten en:

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA
AAAAGUGGCACCGAGUCGGUGCUUUUUUU (versión 5; SEQ ID NO: 30);

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA
AAAAGUGGCACCGAGUCGGUGCUUUUU (versión 6; SEQ ID NO: 31);

30 o

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU (versión 7; SEQ ID NO: 32).

35 Los ARN guía dirigidos a cualquiera de las secuencias diana de ARN guía desveladas en el presente documento pueden incluir, por ejemplo, un segmento dirigido al ADN en el extremo 5' del ARN guía fusionado a cualquiera de las secuencias de armazón de ARN guía a modo de ejemplo en el extremo 3' del ARN guía. Es decir, cualquiera de los segmentos dirigidos al ADN desvelados en el presente documento puede unirse al extremo 5' de cualquiera de las secuencias de armazón anteriores para formar un ARN guía único (ARN guía quimérico).

40 Los ARN guía pueden incluir modificaciones o secuencias que proporcionen características adicionales deseables (por ejemplo, estabilidad modificada o regulada; direccionamiento subcelular; seguimiento con una marca fluorescente; un sitio de unión para una proteína o complejo proteico; y similares). Ejemplos de dichas modificaciones

incluyen, por ejemplo, una caperuza en 5' (por ejemplo, una caperuza de 7-metilguanilato (m7G)); una cola 3' poliadenilada (por ejemplo, una cola 3' poli(A)); una secuencia ribointerruptora (por ejemplo, para permitir la estabilidad regulada y/o la accesibilidad regulada por proteínas y/o complejos de proteína); una secuencia de control de la estabilidad; una secuencia que forma un dúplex de ARNbc (por ejemplo, una horquilla); una modificación o secuencia que dirija el ARN a una localización subcelular (por ejemplo, núcleo, mitocondrias, cloroplastos, etc.); una modificación o secuencia que proporcione seguimiento (por ejemplo, conjugación directa con una molécula fluorescente, conjugación con un resto que facilita la detección fluorescente, una secuencia que permite la detección fluorescente, etc.); una modificación o secuencia que proporciona un sitio de unión para proteínas (por ejemplo, proteínas que actúan sobre el ADN, incluidos activadores transcripcionales, represores transcripcionales, ADN metiltransferasas, ADN desmetilasas, histona acetiltransferasas, histona desacetilasas, etc.); y combinaciones de las mismas. Otros ejemplos de modificaciones incluyen estructuras dúplex de tallo-bucle manipuladas, regiones de protuberancia manipuladas, horquillas en 3' de la estructura dúplex de tallo-bucle manipuladas, o cualquier combinación de las mismas. Véase, por ejemplo, el documento de patente US 2015/0376586. Una protuberancia puede ser una región de nucleótidos no apareada dentro del dúplex constituido por la región similar al ARNcr y la región mínima similar al ARNtrcr. Una protuberancia puede comprender, en un lado del dúplex, un 5'-XXXY-3' no apareado donde X es cualquier purina e Y puede ser un nucleótido que puede formar un par de tibeo con un nucleótido en la cadena opuesta, y una región de nucleótidos no apareados en el otro lado del dúplex.

En algunos casos, puede usarse un sistema de activación transcripcional que comprenda una proteína de fusión dCas9-VP64 emparejada con MS2-p65-HSF1. Los ARN guía en dichos sistemas pueden diseñarse con secuencias aptámeras añadidas al tetrabucle y al tallo-bucle 2 del ARN_g diseñados para unirse a las proteínas de la cubierta del bacteriófago MS2 dimerizado. Véase, por ejemplo, Konermann et al. (2015) Nature 517(7536):583-588.

Los ácidos nucleicos no modificados pueden ser propensos a la degradación. Los ácidos nucleicos exógenos también pueden inducir una respuesta inmunitaria innata. Las modificaciones pueden ayudar a introducir estabilidad y reducir la inmunogenicidad. Los ARN guía pueden comprender nucleósidos modificados y nucleótidos modificados incluyendo, por ejemplo, uno o más de los siguientes: (1) alteración o sustitución de uno o ambos de los oxígenos del fosfato no de enlace y/o de uno o más de los oxígenos del fosfato de enlace en el enlace fosfodiéster del esqueleto; (2) alteración o sustitución de un constituyente del azúcar ribosa, tal como la alteración o sustitución del hidroxilo 2' en el azúcar ribosa; (3) sustitución del resto de fosfato por conectores defosfo; (4) modificación o sustitución de una nucleobase natural; (5) sustitución o modificación del esqueleto de ribosa-fosfato; (6) modificación del extremo 3' o del extremo 5' del oligonucleótido (por ejemplo, eliminación, modificación o sustitución de un grupo fosfato terminal o conjugación de un resto); y (7) modificación del azúcar. Otras posibles modificaciones del ARN guía incluyen modificaciones de o sustitución de uracilos o trectos de poli-uracilo. Véanse, por ejemplo, los documentos de patente WO 2015/048577 y US 2016/0237455. Pueden realizarse modificaciones similares en los ácidos nucleicos que codifican Cas, tales como los ARNm Cas. Por ejemplo, los ARNm Cas pueden modificarse mediante el agotamiento de uridina usando codones sinónimos.

Como un ejemplo, los nucleótidos en el extremo 5' o 3' de un ARN guía pueden incluir enlaces fosforotioato (por ejemplo, las bases pueden tener un grupo fosfato modificado que es un grupo fosforotioato). Por ejemplo, un ARN guía puede incluir enlaces fosforotioato entre los 2, 3 o 4 nucleótidos terminales en el extremo 5' o 3' del ARN guía. Como otro ejemplo, los nucleótidos en el extremo 5' y/o 3' de un ARN guía pueden tener modificaciones 2'-O-metilo. Por ejemplo, un ARN guía puede incluir modificaciones 2'-O-metilo en los 2, 3 o 4 nucleótidos terminales en el extremo 5' y/o 3' del ARN guía (por ejemplo, el extremo 5'). Véase, por ejemplo, el documento de patente WO 2017/173054 A1 y Finn et al. (2018) Cell Reports 22:1-9. Otras posibles modificaciones se describen con más detalle en otra parte del presente documento. En un ejemplo específico, un ARN guía incluye análogos de 2'-O-metilo y enlaces internucleotídicos de fosforotioato 3' en los tres primeros residuos de ARN terminales en 5' y 3'. Dichas modificaciones químicas pueden proporcionar, por ejemplo, una mayor estabilidad y protección frente a las exonucleasas a los ARN guía, permitiéndoles persistir dentro de las células durante más tiempo que los ARN guía no modificados. Dichas modificaciones químicas también pueden, por ejemplo, proteger frente a respuestas inmunitarias intracelulares innatas que pueden degradar activamente el ARN o desencadenar cascadas inmunitarias que provoquen la muerte celular.

En algunos ARN guía (por ejemplo, ARN guía únicos), al menos un bucle (por ejemplo, dos bucles) del ARN guía se modifica mediante la inserción de una secuencia de ARN distinta que se une a uno o más adaptadores (es decir, proteínas o dominios adaptadores). Dichas proteínas adaptadoras pueden usarse para reclutar además uno o más dominios funcionales heterólogos, tales como dominios de activación transcripcional (por ejemplo, para su uso en el cribado CRISPRa en las células biosensoras SAM/tau). Ejemplos de proteínas de fusión que comprenden dichas proteínas adaptadoras (es decir, proteínas adaptadoras químéricas) se desvelan en otra parte del presente documento. Por ejemplo, un bucle de unión a MS2 ggccAACAUAGAGGAUCACCCAUGUGCAGggcc (SEQ ID NO: 33) puede reemplazar los nucleótidos +13 a +16 y los nucleótidos +53 a +56 del armazón de ARN_g (esqueleto) establecido en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31) o el esqueleto de ARN_g por el sistema CRISPR/Cas9 de *S. pyogenes* descrito en el documento de patente WO 2016/049258 y Konermann et al. (2015) Nature 517(7536):583-588. Véanse también los documentos de patente US 2019-0284572 y WO 2019/183123. La numeración de ARN guía usada en el presente documento se refiere a la numeración de nucleótidos en la secuencia de armazón de ARN guía (es decir, la secuencia aguas abajo del segmento dirigido al ADN del ARN guía). Por ejemplo, el primer nucleótido del armazón de ARN guía es +1, el segundo nucleótido del armazón es +2, y así sucesivamente.

Los residuos correspondientes a los nucleótidos +13 a +16 en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31) son la secuencia de bucle en la región que abarca los nucleótidos +9 a +21 en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31), una región denominada en el presente documento tetrabucle. Los residuos correspondientes a los nucleótidos +53 a +56 en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31) son la secuencia de bucle en la región que abarca los nucleótidos +48 a +61 en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31), una región denominada en el presente documento tallo-bucle 2. Otras secuencias de tallo-bucle en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31) comprenden el tallo-bucle 1 (nucleótidos +33 a + 41) y el tallo-bucle 3 (nucleótidos +63 a + 75). La estructura resultante es un almacén de ARN_g en el que cada una de las secuencias de tetrabucle y tallo-bucle 2 se ha sustituido por un bucle de unión a MS2. El tetrabucle y el tallo-bucle 2 sobresalen de la proteína Cas9 de tal manera que la adición de un bucle de unión a MS2 no debería interferir con ningún residuo de Cas9. Además, la proximidad de los sitios tetrabucle y tallo-bucle 2 al ADN indica que la localización en estos lugares podría resultar en un alto grado de interacción entre el ADN y cualquier proteína reclutada, tal como un activador transcripcional. Así, en algunos ARN_g, los nucleótidos correspondientes a +13 a +16 y/o los nucleótidos correspondientes a +53 a +56 del almacén de ARN guía establecido en SEQ ID NO: 17 o SEQ ID NO: 19 (o SEQ ID NO: 30 o 31) o los residuos correspondientes cuando se alinean óptimamente con cualquiera de estos almacenes/esqueletos se sustituyen por las distintas secuencias de ARN capaces de unirse a una o más proteínas o dominios adaptadores. Alternativa o adicionalmente, pueden añadirse secuencias de unión al adaptador al extremo 5' o al extremo 3' de un ARN guía. Un almacén de ARN guía a modo de ejemplo que comprende bucles de unión a MS2 en las regiones tetrabucle y bucle madre 2 puede comprender, consistir esencialmente en, o consistir en la secuencia establecida en SEQ ID NO: 34. Un ARN guía único genérico a modo de ejemplo que comprende bucles de unión a MS2 en las regiones tetrabucle y tallo-bucle 2 puede comprender, consistir esencialmente en, o consistir en la secuencia establecida en SEQ ID NO: 35.

Los ARN guía pueden proporcionarse en cualquier forma. Por ejemplo, el ARN_g puede proporcionarse en forma de ARN, ya sea como dos moléculas (ARN_{cr} y ARN_{tracr} separados) o como una molécula (ARN_g), y opcionalmente en forma de un complejo con una proteína Cas. El ARN_g también puede proporcionarse en forma de ADN que codifica el ARN_g. El ADN que codifica el ARN_g puede codificar una única molécula de ARN (ARN_g) o moléculas de ARN separadas (por ejemplo, ARN_{cr} y ARN_{tracr} separados). En este último caso, el ADN que codifica el ARN_g puede proporcionarse como una molécula de ADN o como moléculas de ADN separadas que codifican el ARN_{cr} y el ARN_{tracr}, respectivamente.

Cuando un ARN_g se proporciona en forma de ADN, el ARN_g puede expresarse de forma transitoria, condicional o constitutiva en la célula. Los ADN que codifican los ARN_g pueden integrarse de forma estable en el genoma de la célula y unirse operativamente a un promotor activo en la célula. Alternativamente, los ADN que codifican los ARN_g pueden unirse operativamente a un promotor en una construcción de expresión. Por ejemplo, el ADN que codifica el ARN_g puede estar en un vector que comprende un ácido nucleico heterólogo, tal como un ácido nucleico que codifica una proteína Cas. Alternativamente, puede estar en un vector o un plásmido que está separado del vector que comprende el ácido nucleico que codifica la proteína Cas. Los promotores que pueden usarse en dichas construcciones de expresión son promotores activos en células de mamífero, por ejemplo, promotores activos en una o más de una célula humana, una célula de mamífero no humano, una célula de roedor, una célula de ratón, una célula de rata, una célula pluripotente, una célula madre embrionaria (ES), una célula madre adulta, una célula progenitora de desarrollo restringido, una célula madre pluripotente inducida (iPS), o un embrión en estadio unicelular. Dichos promotores pueden ser, por ejemplo, promotores condicionales, promotores inducibles, promotores constitutivos o promotores específicos de tejido. Dichos promotores también pueden ser, por ejemplo, bidireccionales. Ejemplos específicos de promotores adecuados incluyen un promotor de ARN polimerasa III, tal como un promotor U6 humano, un promotor U6 polimerasa III de rata o un promotor U6 polimerasa III de ratón.

Alternativamente, los ARN_g pueden prepararse mediante otros métodos. Por ejemplo, los ARN_g pueden prepararse mediante transcripción *in vitro* usando, por ejemplo, ARN polimerasa T7 (véanse, por ejemplo, los documentos de patente WO 2014/089290 y WO 2014/065596. Los ARN guía también pueden ser una molécula producida sintéticamente preparada mediante síntesis química. Por ejemplo, un ARN guía puede sintetizarse químicamente para incluir análogos de 2'-O-metil y enlaces internucleotídicos de fosforotioato 3' en los tres primeros residuos de ARN de los extremos 5' y 3'.

Los ARN guía (o los ácidos nucleicos que codifican ARN guía) pueden estar en composiciones que comprenden uno o más ARN guía (por ejemplo, 1, 2, 3, 4 o más ARN guía) y un portador que aumenta la estabilidad del ARN guía (por ejemplo, prolongando el periodo en condiciones de almacenamiento dadas (por ejemplo, -20 °C, 4 °C o temperatura ambiente) durante la cual los productos de degradación permanecen por debajo de un umbral, tal como por debajo del 0,5 % en peso del ácido nucleico o proteína de partida; o aumentando la estabilidad *in vivo*). Ejemplos no limitantes de dichos portadores incluyen microesferas de poli(ácido láctico) (PLA), microesferas de poli(ácido D,L-láctico-co-glicólico) (PLGA), liposomas, micelas, micelas inversas, cocleatos lipídicos y microtúbulos lipídicos. Dichas composiciones pueden comprender además una proteína Cas, tal como una proteína Cas9, o un ácido nucleico que codifica una proteína Cas.

B. Secuencias diana del ARN guía

Los ADN diana para los ARN guía incluyen secuencias de ácido nucleico presentes en un ADN al que se unirá un segmento de ADN diana de un ARNg, siempre que existan condiciones suficientes para la unión. Las condiciones de unión ADN/ARN adecuadas incluyen las condiciones fisiológicas normalmente presentes en una célula. Otras condiciones de unión ADN/ARN adecuadas (por ejemplo, condiciones en un sistema sin células) son conocidas en la técnica (véase, por ejemplo, *Molecular Cloning: A Laboratory Manual*, 3.^a ed. (Sambrook et al., Harbor Laboratory Press 2001)). La cadena del ADN diana que es complementaria y se hibrida con el ARNg puede denominarse "cadena complementaria", y la cadena del ADN diana que es complementaria a la "cadena complementaria" (y, por lo tanto, no es complementaria a la proteína Cas o al ARNg) puede denominarse "cadena no complementaria" o "cadena molde".

El ADN diana incluye tanto la secuencia de la cadena complementaria a la que se hibrida el ARN guía como la secuencia correspondiente de la cadena no complementaria (por ejemplo, adyacente al motivo adyacente al protoespaciador (PAM)). El término "secuencia diana del ARN guía", como se usa en el presente documento, se refiere específicamente a la secuencia de la cadena no complementaria correspondiente a (es decir, el complemento inverso de) la secuencia con la que se hibrida el ARN guía en la cadena complementaria. Es decir, la secuencia diana del ARN guía se refiere a la secuencia de la cadena no complementaria adyacente a la PAM (por ejemplo, aguas arriba o 5' de PAM en el caso de Cas9). Una secuencia diana de ARN guía es equivalente al segmento dirigido al ADN de un ARN guía, pero con timinas en lugar de uracilos. Como un ejemplo, una secuencia diana de ARN guía para una enzima SpCas9 puede referirse a la secuencia aguas arriba de PAM 5'-NGG-3' en la cadena no complementaria. Un ARN guía está diseñado para tener complementariedad con la cadena complementaria de un ADN diana, donde la hibridación entre el segmento dirigido al ADN del ARN guía y la cadena complementaria del ADN diana promueve la formación de un complejo CRISPR. No se requiere necesariamente una complementariedad completa, siempre que haya suficiente complementariedad para causar la hibridación y promover la formación de un complejo CRISPR. Si en el presente documento se hace referencia a un ARN guía como dirigido a una secuencia diana de ARN guía, lo que se quiere decir es que el ARN guía se hibrida con la secuencia de la cadena complementaria del ADN diana que es el complemento inverso de la secuencia diana del ARN guía en la cadena no complementaria.

Un ADN diana o secuencia diana de ARN guía puede comprender cualquier polinucleótido, y puede estar localizada, por ejemplo, en el núcleo o citoplasma de una célula o dentro de un orgánulo de una célula, tal como una mitocondria o cloroplasto. Un ADN diana o secuencia diana de ARN guía puede ser cualquier secuencia de ácido nucleico endógena o exógena a una célula. La secuencia diana de ARN guía puede ser una secuencia que codifique un producto génico (por ejemplo, una proteína) o una secuencia no codificante (por ejemplo, una secuencia reguladora) o puede incluir ambas.

Para los sistemas CRISPRa y SAM, puede ser preferible que la secuencia diana sea adyacente al sitio de inicio de la transcripción de un gen. Por ejemplo, la secuencia diana puede estar dentro de 1000, 900, 800, 700, 600, 500, 400, 300, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5 o 1 par de bases del sitio de inicio de la transcripción. Opcionalmente, la secuencia diana está dentro de la región de 200 pares de bases aguas arriba del sitio de inicio de la transcripción y 1 par de bases aguas abajo del sitio de inicio de la transcripción (-200 a +1).

La unión y el corte específico de sitio de un ADN diana por una proteína Cas puede producirse en lugares determinados por (i) la complementariedad de emparejamiento de bases entre el ARN guía y la cadena complementaria del ADN diana y (ii) un motivo corto, denominado motivo adyacente al protoespaciador (PAM), en la cadena no complementaria del ADN diana. El PAM puede flanquear la secuencia diana de ARN guía. Opcionalmente, la secuencia diana de ARN guía puede estar flanqueada en el extremo 3' por el PAM (por ejemplo, para Cas9). Alternativamente, la secuencia diana de ARN guía puede estar flanqueada en el extremo 5' por la PAM (por ejemplo, para Cpf1). Por ejemplo, el sitio de corte de las proteínas Cas puede estar entre aproximadamente 1 y aproximadamente 10 o entre aproximadamente 2 y aproximadamente 5 pares de bases (por ejemplo, 3 pares de bases) aguas arriba o aguas abajo de la secuencia PAM (por ejemplo, dentro de la secuencia diana de ARN guía). En el caso de SpCas9, la secuencia PAM (es decir, en la cadena no complementaria) puede ser 5'-N₁GG-3', donde N₁ es cualquier nucleótido de ADN, y donde la PAM está inmediatamente a 3' de la secuencia diana de ARN guía en la cadena no complementaria del ADN diana. Como tal, la secuencia correspondiente a la PAM en la cadena complementaria (es decir, el complemento inverso) sería 5'-CCN₂-3', donde N₂ es cualquier nucleótido de ADN y está inmediatamente a 5' de la secuencia a la que hibrida el segmento dirigido al ADN del ARN guía en la cadena complementaria del ADN diana. En algunos de dichos casos, N₁ y N₂ pueden ser complementarios y el par de bases N₁ - N₂ puede ser cualquier par de bases (por ejemplo, N₁ =C y N₂ =G; N₁ =G y N₂ =C; N₁ =A y N₂ =T; o N₁ =T, y N₂ =A). En el caso de Cas9 de *S. aureus*, la PAM puede ser NNGRRT o NNGRR, donde N puede ser A, G, C o T, y R puede ser G o A. En el caso de Cas9 de *C. jejuni*, la PAM puede ser, por ejemplo, NNNNACAC o NNNNRYAC, donde N puede ser A, G, C o T, y R puede ser G o A. En algunos casos (por ejemplo, para FnCpf1), la secuencia PAM puede estar aguas arriba del extremo 5' y tener la secuencia 5'-TTN-3'.

Un ejemplo de una secuencia diana de ARN guía es una secuencia de ADN de 20 nucleótidos inmediatamente anterior a un motivo NGG reconocido por una proteína SpCas9. Por ejemplo, dos ejemplos de secuencias diana de ARN guía más PAM son GN₁₉NGG (SEQ ID NO: 25) o N₂₀NGG (SEQ ID NO: 26). Véase, por ejemplo, el documento de patente WO 2014/165825. La guanina en el extremo 5' puede facilitar la transcripción por la ARN polimerasa en las células. Otros ejemplos de secuencias diana de ARN guía más PAM pueden incluir dos nucleótidos de guanina en el extremo

5' (por ejemplo, GGN₂₀NGG; SEQ ID NO: 27) para facilitar la transcripción eficiente por la T7 polimerasa *in vitro*. Véase, por ejemplo, el documento de patente WO 2014/065596. Otras secuencias diana de ARN guía más PAM pueden tener entre 4-22 nucleótidos de longitud de SEQ ID NOS: 25-27, incluyendo el 5' G o GG y el 3' GG o NGG. Otras secuencias diana de ARN guía más PAM pueden tener entre 14 y 20 nucleótidos de longitud de SEQ ID NO: 25-27.

La formación de un complejo CRISPR hibridado a un ADN diana puede dar lugar al corte de una o ambas cadenas del ADN diana dentro o cerca de la región correspondiente a la secuencia diana del ARN guía (es decir, la secuencia diana del ARN guía en la cadena no complementaria del ADN diana y el complemento inverso en la cadena complementaria a la que se hibrida el ARN guía). Por ejemplo, el sitio de corte puede estar dentro de la secuencia diana del ARN guía (por ejemplo, en una ubicación definida relativa a la secuencia PAM). El "sitio de corte" incluye la posición de un ADN diana en la que una proteína Cas produce una rotura de cadena sencilla o una rotura de cadena doble. El sitio de corte puede estar en una sola cadena (por ejemplo, cuando se usa una melladora) o en ambas cadenas de un ADN bicatenario. Los sitios de corte pueden estar en la misma posición en ambas cadenas (produciendo extremos romos; por ejemplo, Cas9)) o pueden estar en diferentes sitios en cada cadena (produciendo extremos protuberantes (es decir, salientes); por ejemplo, Cpf1). Los extremos protuberantes pueden producirse, por ejemplo, usando dos proteínas Cas, cada una de las cuales produce una rotura de cadena simple en un sitio de corte diferente en una cadena diferente, produciendo así una rotura de cadena doble. Por ejemplo, una primera melladora puede crear una rotura monocatenaria en la primera cadena de ADN bicatenario (ADN_{dc}), y una segunda melladora puede crear una rotura monocatenaria en la segunda cadena de ADN_{dc} de forma que se creen secuencias salientes. En algunos casos, la secuencia diana de ARN guía o el sitio de corte de la melladora en la primera cadena están separados de la secuencia diana de ARN guía o del sitio de corte de la melladora en la segunda cadena por al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100, 250, 500 o 1.000 pares de bases.

IV. Métodos de cribado de modificadores genéticos de la siembra o agregación de Tau

Las líneas celulares biosensoras Cas/tau desveladas en el presente documento pueden usarse en métodos de cribado de modificadores genéticos de la siembra o agregación de tau. Dichos métodos comprenden proporcionar una población de células biosensoras Cas/tau como se describe en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la siembra o agregación de tau en las células diana.

En particular, dichos métodos comprenden proporcionar una población de células biosensoras Cas/tau (es decir, una población de células que comprende una proteína Cas, un primer dominio de repetición de tau unido a un primer indicador, y un segundo dominio de repetición de tau unido a un segundo indicador), introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes, y cultivar la población de células para permitir la edición y expansión del genoma. La pluralidad de ARN guía únicos forman complejos con la proteína Cas, y la proteína Cas corta la pluralidad de genes dando lugar a la inactivación de la función génica para producir una población de células modificada genéticamente. A continuación, la población de células modificada genéticamente se pone en contacto con un agente de siembra de tau para producir una población de células sembrada. La población de células sembrada se cultiva para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación. Por último, se determina la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía. El enriquecimiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía aumenta la agregación de tau.

De forma similar, las líneas celulares biosensoras de SAM/tau desveladas en el presente documento pueden usarse en métodos de cribado de modificadores genéticos de la siembra o agregación de tau. Dichos métodos comprenden proporcionar una población de células biosensoras SAM/tau como se describe en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la siembra o agregación de tau en las células diana.

En particular, dichos métodos comprenden proporcionar una población de células biosensoras SAM/tau (es decir, una población de células que comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, una proteína adaptadora quimérica que comprende una proteína adaptadora fusionada a uno o más dominios de activación transcripcional, un primer dominio de repetición de tau unido a un primer indicador, y un segundo dominio de repetición de tau unido a un segundo indicador), introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes, y cultivar la población de células para permitir la activación transcripcional y la expansión. La pluralidad de ARN guía únicos forma complejos con la proteína Cas quimérica y la proteína adaptadora quimérica, y los complejos activan la transcripción de la pluralidad de genes dando lugar a un aumento de la expresión génica y a una población de células modificada. A continuación, la población de células modificada se pone en contacto

con un agente de siembra de tau para producir una población sembrada de células. La población de células sembrada se cultiva para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación. Por último, se determina la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía. El enriquecimiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía aumenta la agregación de tau.

Las células biosensoras Cas/tau usadas en el método pueden ser cualquiera de las células biosensoras Cas/tau desveladas en otra parte del presente documento. Similarmente, las células biosensoras SAM/tau usadas en el método pueden ser cualquiera de las células biosensoras SAM/tau desveladas en otra parte del presente documento. El primer dominio de repetición de tau y el segundo dominio de repetición de tau pueden ser diferentes o pueden ser similares o iguales. El dominio de repetición de tau puede ser cualquiera de los dominios de repetición de tau desvelados en el presente documento. Por ejemplo, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau puede ser un dominio de repetición de tau natural o puede comprender una mutación pro-agregación (por ejemplo, una mutación patógena, pro-agregación), tal como una mutación de tau P301S. El primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender un dominio de cuatro repeticiones de tau. Como un ejemplo específico, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender, consistir esencialmente en, o consistir en SEQ ID NO: 11 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 11. En un ejemplo específico, el ácido nucleico que codifica el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 12 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 12, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 11.

El primer dominio de repetición de tau puede unirse al primer indicador y el segundo dominio de repetición de tau puede unirse al segundo indicador por cualquier medio. Por ejemplo, el indicador puede fusionarse al dominio de repetición de tau (por ejemplo, como parte de una proteína de fusión). Las proteínas indicadoras son un par de proteínas indicadoras que producen una señal detectable cuando el primer dominio de repetición de tau unido al primer indicador se agrega con el segundo dominio de repetición de tau unido al segundo indicador. Específicamente, la primera y la segunda proteínas indicadoras son proteínas fluorescentes que forman un par de transferencia de energía por resonancia de fluorescencia (FRET). FRET es un fenómeno físico en el que un fluoróforo donante en su estado de excitación transfiere de forma no radiativa su energía de excitación a un fluoróforo aceptor vecino, provocando así que el aceptor emita su fluorescencia característica. Los ejemplos de pares FRET (fluoróforo donante y aceptor) son bien conocidos. Véase, por ejemplo, Bajar et al. (2016) *Sensors* (Basel) 16(9): 1488. Como un ejemplo específico de un par FRET, el primer indicador puede ser proteína fluorescente cian (CFP) y el segundo indicador puede ser proteína fluorescente amarilla (YFP). Como un ejemplo específico, CFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 13 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 13. Como otro ejemplo específico, la YFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 15 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 15.

Para las células biosensoras Cas/tau, la proteína Cas puede ser cualquier proteína Cas descrita en otra parte del presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 21 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 21.

Una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 22 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 22, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 21.

Para las células biosensoras SAM/tau, la proteína Cas puede ser cualquier proteína Cas desvelada en otra parte del presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas química puede comprender la proteína Cas inactiva por nucleasa fusionada a un dominio de activación transcripcional VP64. Por ejemplo, la proteína Cas química puede comprender de extremo N a extremo C: la proteína Cas inactiva por nucleasa; una señal de localización nuclear; y el dominio activador transcripcional VP64. Como un ejemplo específico,

la proteína adaptadora puede ser una proteína de la cubierta de MS2, y el uno o más dominios de activación transcripcional en la proteína adaptadora quimérica puede comprender un dominio de activación transcripcional p65 y un dominio de activación transcripcional HSF1. Por ejemplo, la proteína adaptadora quimérica puede comprender de extremo N a extremo C: la proteína de la cubierta de MS2; una señal de localización nuclear; el dominio de activación transcripcional p65; y el dominio de activación transcripcional HSF1. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 38 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 38, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 36. En un ejemplo específico, el ácido nucleico que codifica la proteína adaptadora quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 39 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 39, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 37.

Una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células.

Como se desvela en otra parte del presente documento, las células pueden ser de cualquier tipo. Por ejemplo, las células pueden ser células eucariotas, células de mamífero o células humanas (por ejemplo, células HEK293T o células neuronales).

La pluralidad de ARN guía únicos puede introducirse en la población de células por cualquier medio conocido. En algunos métodos, los ARN guía se introducen en la población de células mediante transducción vírica, tal como transducción retrovírica, adenovírica o lentivírica. En un ejemplo específico, los ARN guía pueden introducirse mediante transducción lentivírica. Cada uno de la pluralidad de ARN guía únicos puede estar en un vector vírico separado. La población de células puede infectarse con cualquier multiplicidad de infección. Por ejemplo, la multiplicidad de infección puede ser entre aproximadamente 0,1 y aproximadamente 1,0, entre aproximadamente 0,1 y aproximadamente 0,9, entre aproximadamente 0,1 y aproximadamente 0,8, entre aproximadamente 0,1 y aproximadamente 0,7, entre aproximadamente 0,1 y aproximadamente 0,6, entre aproximadamente 0,1 y aproximadamente 0,5, entre aproximadamente 0,1 y aproximadamente 0,4, o entre aproximadamente 0,1 y aproximadamente 0,3. Alternativamente, la multiplicidad de infección puede ser inferior a aproximadamente 1,0, inferior a aproximadamente 0,9, inferior a aproximadamente 0,8, inferior a aproximadamente 0,7, inferior a aproximadamente 0,6, inferior a aproximadamente 0,5, inferior a aproximadamente 0,4, inferior a aproximadamente 0,3 o inferior a aproximadamente 0,2. En un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

Los ARN guía pueden introducirse en la población de células junto con un marcador de selección o un gen indicador para seleccionar células que tengan los ARN guía, y el método puede comprender además la selección de células que comprendan el marcador de selección o el gen indicador. En otra parte del presente documento se proporcionan ejemplos de marcadores de selección y genes indicadores. Como un ejemplo, el marcador de selección puede ser uno que imparta resistencia a un fármaco, tal como neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y blasticidina S desaminasa. Otro marcador de selección a modo de ejemplo es la proteína de resistencia a la bleomicina, codificada por el gen *Sh ble* (gen de la bleomicina de *Streptoalloteichus hindustanus*), que confiere resistencia a la zeocina (fleomicina D1). Por ejemplo, las células pueden seleccionarse con un fármaco (por ejemplo, puromicina) de modo que solo las células transducidas con una construcción de ARN guía se conserven para ser usadas para llevar a cabo el cribado. Por ejemplo, el fármaco puede ser puromicina o zeocina (fleomicina D1).

En algunos métodos, la pluralidad de ARN guía únicos se introduce a una concentración seleccionada de forma que la mayoría de las células reciban solo uno de los ARN guía únicos. Por ejemplo, si los ARN guía se introducen mediante transducción vírica, las células pueden infectarse con una multiplicidad de infección baja para garantizar que la mayoría de las células reciban una única construcción vírica con una alta probabilidad. Como un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

La población de células en la que se introduce la pluralidad de ARN guía únicos puede ser cualquier número adecuado de células. Por ejemplo, la población de células puede comprender más de aproximadamente 50, más de aproximadamente 100, más de aproximadamente 200, más de aproximadamente 300, más de aproximadamente 400, más de aproximadamente 500, más de aproximadamente 600, más de aproximadamente 700, más de aproximadamente 800, más de aproximadamente 900 o más de aproximadamente 1000 células por ARN guía único. En un ejemplo específico, la población de células comprende más de aproximadamente 300 células o más de aproximadamente 500 células por ARN guía único.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de genes. Por ejemplo, la pluralidad de ARN guía únicos puede dirigirse a aproximadamente 50 o más genes, aproximadamente 100 o más genes, aproximadamente 200 o más genes, aproximadamente 300 o más genes, aproximadamente 400 o más genes, aproximadamente 500 o más genes, aproximadamente 1000 o más genes, aproximadamente 2000 o más genes, aproximadamente 3000 o más genes, aproximadamente 4000 o más genes, aproximadamente 5000 o más genes, aproximadamente 10000 o más genes, o aproximadamente 20000 o más genes. En algunos métodos, los ARN guía pueden seleccionarse para dirigirse a genes en una vía de señalización particular. En algunos métodos, la biblioteca de ARN guía únicos es una biblioteca de todo el genoma.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de secuencias dentro de cada gen diana individual. En algunos métodos, se dirige a una pluralidad de secuencias diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, aproximadamente 2 a aproximadamente 10, aproximadamente 2 a aproximadamente 9, aproximadamente 2 a aproximadamente 8, aproximadamente 2 a aproximadamente 7, aproximadamente 2 a aproximadamente 6, aproximadamente 2 a aproximadamente 5, aproximadamente 2 a aproximadamente 4 o aproximadamente 2 a aproximadamente 3 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5 o al menos aproximadamente 6 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Como un ejemplo específico, aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana. Como otro ejemplo específico, aproximadamente 3 a aproximadamente 6 o aproximadamente 4 a aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana.

Los ARN guía pueden dirigirse a cualquier localización deseada en los genes diana. En algunos métodos CRISPRn que usan las células biosensoras Cas/tau, cada ARN guía se dirige a un exón constitutivo si es posible. En algunos métodos, cada ARN guía se dirige a un exón constitutivo en 5' si es posible. En algunos métodos, cada ARN guía se dirige a un primer exón, un segundo exón o un tercer exón (desde el extremo 5' del gen) si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, cada ARN guía puede dirigirse a una secuencia diana de ARN guía dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción, si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, en donde cada ARN guía puede comprender uno o más elementos de unión al adaptador a los que se puede unir específicamente la proteína adaptadora quimérica. En un ejemplo, cada ARN guía comprende dos elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica, opcionalmente en donde un primer elemento de unión al adaptador está dentro de un primer bucle de cada uno de los uno o más ARN guía, y un segundo elemento de unión al adaptador está dentro de un segundo bucle de cada uno de los uno o más ARN guía. Por ejemplo, el elemento de unión al adaptador puede comprender la secuencia establecida en SEQ ID NO: 33. En un ejemplo específico, cada uno de uno o más ARN guía es un único ARN guía que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de ARN CRISPR transactivador (ARNtracr), y el primer bucle es el tetrabucle correspondiente a los residuos 13-16 de SEQ ID NO: 17, 19, 30 o 31, y el segundo bucle es el tallo-bucle 2 correspondiente a los residuos 53-56 de SEQ ID NO: 17, 19, 30 o 31.

La etapa de cultivo de la población de células para permitir la edición y expansión del genoma puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 10 días, entre aproximadamente 3 días y aproximadamente 9 días, entre aproximadamente 4 días y aproximadamente 8 días, entre aproximadamente 5 días y aproximadamente 7 días, o aproximadamente 6 días. Similarmente, la etapa de cultivo de la población de células para permitir la activación transcripcional y la expansión puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 10 días, entre aproximadamente 3 días y aproximadamente 9 días, entre aproximadamente 4 días y aproximadamente 8 días, entre aproximadamente 5 días y aproximadamente 7 días, o aproximadamente 6 días.

El agente de siembra de tau usado para producir una población de células sembrada es un medio acondicionado. Algunos agentes de siembra adecuados comprenden un dominio de repetición de tau que puede ser, por ejemplo, diferente, similar o igual al primer dominio de repetición de tau y/o al segundo dominio de repetición de tau. La etapa de siembra comprende cultivar la población de células modificada genéticamente en presencia de medio acondicionado obtenido de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado. El medio acondicionado se obtiene de células confluentes positivas a la agregación de tau después de estar en las células confluentes durante 1 a 7 días, por ejemplo, aproximadamente 2 a aproximadamente 6 días, aproximadamente 3 a aproximadamente 5 días, o aproximadamente 4 días. La etapa de siembra puede comprender cultivar la población de células modificada genéticamente en cualquier proporción adecuada entre medio acondicionado y medio recién preparado. Por ejemplo, puede comprender cultivar la población de células modificada genéticamente en aproximadamente 90 % de medio acondicionado y aproximadamente 10 % de medio recién preparado, aproximadamente 85 % de medio acondicionado y aproximadamente 15 % de medio recién preparado, aproximadamente 80 % de medio acondicionado y aproximadamente 20 % de medio recién preparado, aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio recién preparado, aproximadamente 70 % de medio acondicionado y aproximadamente 30 % de medio recién preparado, aproximadamente 65 % de medio acondicionado y

aproximadamente 35 % de medio recién preparado, aproximadamente 60 % de medio acondicionado y
aproximadamente 40 % de medio recién preparado, aproximadamente 55 % de medio acondicionado y
aproximadamente 45 % de medio recién preparado, aproximadamente 50 % de medio acondicionado y
aproximadamente 50 % de medio recién preparado, aproximadamente 45 % de medio acondicionado y
5 aproximadamente 55 % de medio recién preparado, aproximadamente 40 % de medio acondicionado y
aproximadamente 60 % de medio recién preparado, aproximadamente 35 % de medio acondicionado y
aproximadamente 65 % de medio recién preparado, aproximadamente 30 % de medio acondicionado y
aproximadamente 70 % de medio recién preparado, aproximadamente 25 % de medio acondicionado y
10 aproximadamente 75 % de medio recién preparado, aproximadamente 20 % de medio acondicionado y
aproximadamente 80 % de medio recién preparado, aproximadamente 15 % de medio acondicionado y
aproximadamente 85 % de medio recién preparado, o aproximadamente 10 % de medio acondicionado y
aproximadamente 90 % de medio recién preparado. En un ejemplo, puede comprender cultivar la población de células
modificada genéticamente en un medio que comprende al menos aproximadamente 50 % de medio acondicionado y
no más de aproximadamente 50 % de medio recién preparado. En un ejemplo específico, puede comprender cultivar
15 la población de células modificada genéticamente en aproximadamente 75 % de medio acondicionado y
aproximadamente 25 % de medio recién preparado. Opcionalmente, la población de células modificada genéticamente
no se cocultiva con las células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta
de forma estable en un estado agregado.

20 La etapa de cultivar la población de células sembrada para permitir que se formen agregados de tau, en donde se
forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un
subconjunto de la población de células sembrada para producir una población de células positiva a la agregación,
puede ser de cualquier duración adecuada. Por ejemplo, el cultivo puede ser durante entre aproximadamente 1 día y
aproximadamente 7 días, entre aproximadamente 2 días y aproximadamente 6 días, entre aproximadamente 3 días y
25 aproximadamente 5 días, o aproximadamente 4 días. La agregación puede determinarse por cualquier medio
adecuado. Por ejemplo, la población de células positiva a la agregación puede identificarse por citometría de flujo.

La abundancia de los ARN guía puede determinarse por cualquier medio adecuado. En un ejemplo específico, la
abundancia se determina mediante secuenciación de nueva generación. La secuenciación de nueva generación se
30 refiere a tecnologías de secuenciación de ADN de alto rendimiento no basadas en Sanger. Por ejemplo, la
determinación de la abundancia de un ARN guía puede comprender la medición del recuento de lecturas del ARN
guía.

En algunos métodos, un ARN guía se considera enriquecido si la abundancia del ARN guía en relación con la población
35 total de la pluralidad de ARN guía únicos es al menos aproximadamente 1,5 veces mayor en la población de células
positiva a la agregación en relación con la población de células que se cultivan tras la introducción de la biblioteca de
ARN guía. También pueden usarse diferentes umbrales de enriquecimiento. Por ejemplo, un umbral de
enriquecimiento puede fijarse más alto para ser más estricto (por ejemplo, al menos aproximadamente 1,6 veces, al
40 menos aproximadamente 1,7 veces, al menos aproximadamente 1,8 veces, al menos aproximadamente 1,9 veces, al
menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces o al menos aproximadamente 2,5 veces).
Alternativamente, un umbral de enriquecimiento puede fijarse más bajo para ser menos estricto (por ejemplo, al menos
aproximadamente 1,4 veces, al menos aproximadamente 1,3 veces o al menos aproximadamente 1,2 veces).

En un ejemplo, la etapa de determinar la abundancia puede comprender determinar la abundancia de la pluralidad de
45 ARN guía únicos en la población positiva a la agregación en relación con la población de células cultivada después de
la introducción de la biblioteca de ARN guía en un primer momento durante el cultivo y/o un segundo momento durante
el cultivo. Por ejemplo, el primer momento puede estar en un primer pase del cultivo de la población de células, y el
segundo momento puede estar en medio del cultivo de la población de células para permitir la edición y expansión del
genoma. Por ejemplo, el primer momento puede ser después de una cantidad de tiempo suficiente para que los ARN
50 guía formen complejos con la proteína Cas, y para que la proteína Cas corte la pluralidad de genes resultando la
inactivación de la función génica (CRISPRn) o para activar transcripcionalmente la pluralidad de genes (CRISPRa).
Sin embargo, lo ideal es que el primer momento sea en el primer pase celular para determinar la representación de la
biblioteca de ARNg poco después de la infección (es decir, antes de la expansión posterior y la edición del genoma) y
para determinar si cada representación de ARNg evoluciona desde el primer momento al segundo y a cualquier
55 momento adicional hasta un momento final. Esto permite descartar los ARNg/dianas enriquecidos debido a las
ventajas del crecimiento celular durante el transcurso del cribado mediante la verificación de que la abundancia de
ARNg no cambia entre el primer y el segundo momento. Como un ejemplo específico, el primer momento puede ser
después de aproximadamente 1 día, aproximadamente 2 días, aproximadamente 3 días o aproximadamente 4 días
de cultivo y expansión, y el segundo momento puede ser después de aproximadamente 3 días, aproximadamente 4
60 días, aproximadamente 5 días o aproximadamente 6 días de cultivo y expansión. Por ejemplo, el primer momento
puede ser después de aproximadamente 3 días de cultivo y expansión, y el segundo momento puede ser después de
aproximadamente 6 días de cultivo y expansión. En algunos métodos, un gen puede considerarse un modificador
genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del
gen potencia (o se espera que potencie) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen
65 en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor (o mayor que un
umbral de enriquecimiento seleccionado diferente) en la población de células positiva a la agregación en relación con

la población de células cultivada tras la introducción de la biblioteca de ARN guía tanto en el primer momento como en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen potencia (o se espera que potencie) la agregación de tau, si la abundancia de al menos dos ARN guía únicos dirigidos al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor (o mayor que un umbral de enriquecimiento seleccionado diferente) en la población de células positiva a la agregación en relación con la población de células cultivada tras la introducción de la biblioteca de ARN guía en el primer momento o en el segundo momento.

En algunos métodos CRISPRn, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen potencia la agregación de tau. Similarmente, en algunos métodos CRISPRa, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la activación transcripcional del gen potencia la agregación de tau. La primera etapa comprende identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación. La segunda etapa comprende calcular la probabilidad aleatoria de que los ARN guía identificados estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de ARN guía únicos introducidos en la población de células, m es la variedad de ARN guía únicos identificados en la etapa (1), n es la variedad de ARN guía únicos introducidos en la población de células que se dirigen al gen, y n' es la variedad de ARN guía únicos identificados en la etapa (1) que se dirigen al gen. La tercera etapa comprende calcular las puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (1). La puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación dividida por la abundancia relativa del ARN guía en la población de células cultivada tras la introducción de la biblioteca de ARN guía. La abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos. La cuarta etapa comprende seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral. Las posibles puntuaciones de enriquecimiento umbral se han tratado anteriormente. Como un ejemplo específico, la puntuación de enriquecimiento umbral puede fijarse en aproximadamente 1,5 veces.

Variedad cuando se usa en la frase variedad de ARN guía únicos significa el número de secuencias de ARN guía únicos. No se trata de la abundancia, sino más bien del cualitativo "presente" o "no presente". Por variedad de ARN guía únicos se entiende el número de secuencias de ARN guía únicos. La variedad de ARN guía únicos se determina mediante secuenciación de nueva generación (NGS) para identificar todos los ARN guía únicos presentes en una población celular. Se hace usando dos cebadores que reconocen las regiones constantes del vector vírico para amplificar el ARNg que está entre las regiones constantes y un cebador que reconoce una región constante para la secuenciación. Cada ARN guía único presente en la muestra generará recuentos de lecturas usando el cebador de secuenciación. Los resultados de la NGS incluirán la secuencia y también el número de lecturas correspondientes a la secuencia. El número de lecturas se usará para el cálculo de la puntuación de enriquecimiento de cada ARN guía, y la presencia de cada secuencia única nos indicará qué ARN guía está presente. Por ejemplo, si hay tres ARN guía únicos para un gen antes de la selección y los tres se conservan después de la selección, entonces n y n' son 3. Estos números se usan para calcular las estadísticas, pero no los recuentos de lecturas reales. Sin embargo, los recuentos de lecturas para cada ARN guía (en un ejemplo, 100, 200, 50, que corresponden a cada uno de los 3 ARN guía únicos) se usarán para el cálculo de la puntuación de enriquecimiento.

V. Métodos de cribado de modificadores genéticos de la agregación de tau que impiden la agregación de tau

Las líneas celulares biosensoras Cas/tau desveladas en el presente documento también pueden usarse en métodos de cribado de modificadores genéticos de la agregación de tau (por ejemplo, que impiden la agregación de tau o que se espera que impidan la agregación de tau). Dichos métodos pueden comprender proporcionar una población de células biosensoras Cas/tau como se desvela en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la agregación de tau en las células diana.

Como un ejemplo, un método puede comprender proporcionar una población de células biosensoras Cas/tau (por ejemplo, una población de células que comprende una proteína Cas, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador), introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes y cultivar la población de células para permitir la edición y expansión del genoma. La pluralidad de ARN guía únicos forman complejos con la proteína Cas, y la proteína Cas corta la pluralidad de genes dando lugar a la inactivación de la función génica para producir una población de células modificada genéticamente. La población de células modificada genéticamente puede ponerse en contacto con un agente de siembra de tau para producir una población de células sembrada. Por ejemplo, el agente de siembra de tau puede ser un agente de "siembra máxima" como se describe en otra parte del presente documento, tal como lisados celulares de células positivas a la agregación de tau. La población de células sembrada puede cultivarse para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación y en donde no se forman agregados en un segundo subconjunto de la población de células sembrada para producir una población de

células negativa a la agregación. Por último, puede determinarse la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía. El enriquecimiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células después de la siembra y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía impide la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía evite la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía impide la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía impida la agregación de tau. El enriquecimiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la alteración del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva o potencie la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva o potencie la agregación de tau.

De forma similar, las líneas celulares biosensoras de SAM/tau desveladas en el presente documento pueden usarse en métodos de cribado de modificadores genéticos de la agregación de tau (por ejemplo, que evitan la agregación de tau o que se espera que eviten la agregación de tau). Dichos métodos pueden comprender proporcionar una población de células biosensoras SAM/tau como se describe en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la agregación de tau en las células diana.

Como un ejemplo, un método puede comprender proporcionar una población de células biosensoras SAM/tau (por ejemplo, una población de células que comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, una proteína adaptadora quimérica que comprende una proteína adaptadora fusionada a uno o más dominios de activación transcripcional, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador), introducir en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes y cultivar la población de células para permitir la activación transcripcional y la expansión. La pluralidad de ARN guía únicos forma complejos con la proteína Cas quimérica y la proteína adaptadora quimérica, y los complejos activan la transcripción de la pluralidad de genes dando lugar a un aumento de la expresión génica y a una población de células modificada. A continuación, la población de células modificada puede ponerse en contacto con un agente de siembra de tau para producir una población sembrada de células. Por ejemplo, el agente de siembra de tau puede ser un agente de "siembra máxima" como se describe en otra parte del presente documento, tal como lisados celulares de células positivas a la agregación de tau. La población de células sembrada puede cultivarse para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación y en donde no se forman agregados en un segundo subconjunto de la población de células sembrada para producir una población de células negativa a la agregación. Por último, puede determinarse la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía. El enriquecimiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células después de la siembra y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía impide la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía evite la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células

positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía impide la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía evite la agregación de tau. El enriquecimiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía promueva o potencie la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células después de ser sembradas y/o en relación con la población de células que se cultiva después de la introducción de la biblioteca de ARN guía indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía promueva o potencie la agregación de tau.

Las células biosensoras Cas/tau usadas en el método pueden ser cualquiera de las células biosensoras Cas/tau desveladas en otra parte del presente documento. Similarmente, las células biosensoras SAM/tau usadas en el método pueden ser cualquiera de las células biosensoras SAM/tau desveladas en otra parte del presente documento. El primer dominio de repetición de tau y el segundo dominio de repetición de tau pueden ser diferentes o pueden ser similares o iguales. El dominio de repetición de tau puede ser cualquiera de los dominios de repetición de tau desvelados en otra parte del presente documento. Por ejemplo, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau puede ser un dominio de repetición de tau natural o puede comprender una mutación pro-agregación (por ejemplo, una mutación patógena, pro-agregación), tal como una mutación de tau P301S. El primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender un dominio de cuatro repeticiones de tau. Como un ejemplo específico, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender, consistir esencialmente en, o consistir en SEQ ID NO: 11 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 11. En un ejemplo específico, el ácido nucleico que codifica el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 12 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 12, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 11.

El primer dominio de repetición de tau puede unirse al primer indicador y el segundo dominio de repetición de tau puede unirse al segundo indicador por cualquier medio. Por ejemplo, el indicador puede fusionarse al dominio de repetición de tau (por ejemplo, como parte de una proteína de fusión). Las proteínas indicadoras son un par de proteínas indicadoras fluorescentes que producen una señal detectable cuando el primer dominio de repetición de tau unido al primer indicador se agrega con el segundo dominio de repetición de tau unido al segundo indicador. En particular, la primera y segunda proteínas indicadoras son un par de transferencia de energía por resonancia de fluorescencia (FRET). FRET es un fenómeno físico en el que un fluoróforo donante en su estado excitado transfiere de forma no radiativa su energía de excitación a un fluoróforo aceptor vecino, provocando así que el aceptor emita su fluorescencia característica. Los ejemplos de pares FRET (fluoróforos donante y aceptor) son bien conocidos. Véase, por ejemplo, Bajar et al. (2016) *Sensors (Basel)* 16(9):1488. Como un ejemplo específico de un par FRET, el primer indicador puede ser proteína fluorescente cian (CFP) y el segundo indicador puede ser proteína fluorescente amarilla (YFP). Como un ejemplo específico, la CFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 13 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 13. Como otro ejemplo específico, la YFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 15 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 15.

Para las células biosensoras Cas/tau, la proteína Cas puede ser cualquier proteína Cas desvelada en otra parte del presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 21 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 21.

Una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador

pueden integrarse genómicamente en la población de células. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 22 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 22, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 21.

Para las células biosensoras SAM/tau, la proteína Cas puede ser cualquier proteína Cas desvelada en el presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas quimérica puede comprender la proteína Cas inactiva por nucleasa fusionada a un dominio de activación transcripcional VP64. Por ejemplo, la proteína Cas quimérica puede comprender de extremo N a extremo C: la proteína Cas inactiva por nucleasa; una señal de localización nuclear; y el dominio activador transcripcional VP64. Como un ejemplo específico, la proteína adaptadora puede ser una proteína de la cubierta de MS2, y el uno o más dominios de activación transcripcional en la proteína adaptadora puede comprender un dominio de activación transcripcional p65 y un dominio de activación transcripcional HSF1. Por ejemplo, la proteína adaptadora quimérica puede comprender de extremo N a extremo C: la proteína de la cubierta de MS2; una señal de localización nuclear; el dominio de activación transcripcional p65; y el dominio de activación transcripcional HSF1. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 38 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 38, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 36. En un ejemplo específico, el ácido nucleico que codifica la proteína adaptadora quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 39 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 39, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 37.

Una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células.

Como se desvela en otra parte del presente documento, las células pueden ser de cualquier tipo de mamífero. Por ejemplo, las células pueden ser células humanas (por ejemplo, células HEK293T o células neuronales).

La pluralidad de ARN guía únicos puede introducirse en la población de células por cualquier medio conocido. En algunos métodos, los ARN guía se introducen en la población de células mediante transducción vírica, tal como la transducción retrovírica, adenovírica o lentivírica. En un ejemplo específico, los ARN guía pueden introducirse mediante transducción lentivírica. Cada uno de la pluralidad de ARN guía únicos puede estar en un vector vírico separado. La población de células puede infectarse con cualquier multiplicidad de infección. Por ejemplo, la multiplicidad de infección puede estar entre aproximadamente 0,1 y aproximadamente 1,0 aproximadamente, entre aproximadamente 0,1 y aproximadamente 0,9, entre aproximadamente 0,1 y aproximadamente 0,8, entre aproximadamente 0,1 y aproximadamente 0,7, entre aproximadamente 0,1 y aproximadamente 0,6, entre aproximadamente 0,1 y aproximadamente 0,5, entre aproximadamente 0,1 y aproximadamente 0,4, o entre aproximadamente 0,1 y aproximadamente 0,3. Alternativamente, la multiplicidad de infección puede ser inferior a aproximadamente 1,0, inferior a aproximadamente 0,9, inferior a aproximadamente 0,8, inferior a aproximadamente 0,7, inferior a aproximadamente 0,6, inferior a aproximadamente 0,5, inferior a aproximadamente 0,4, inferior a aproximadamente 0,3 o inferior a aproximadamente 0,2. En un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

Los ARN guía pueden introducirse en la población de células junto con un marcador de selección o un gen indicador para seleccionar células que tengan los ARN guía, y el método puede comprender además la selección de células que comprendan el marcador de selección o el gen indicador. En otra parte del presente documento se proporcionan ejemplos de marcadores de selección y genes indicadores. Como un ejemplo, el marcador de selección puede ser uno que imparta resistencia a un fármaco, tal como neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y blasticidina S desaminasa. Otro marcador de selección a modo de ejemplo es la proteína de resistencia a la bleomicina, codificada por el gen *Sh ble* (gen de la bleomicina de *Streptoalloteichus hindustanus*), que confiere resistencia a la zeocina (fleomicina D1). Por ejemplo, las células pueden seleccionarse con un fármaco (por ejemplo, puromicina) de modo que solo las células transducidas con una construcción de ARN guía se conserven para ser usadas para llevar a cabo el cribado. Por ejemplo, el fármaco puede ser puromicina o zeocina (fleomicina D1).

En algunos métodos, la pluralidad de ARN guía únicos se introduce a una concentración seleccionada de forma que la mayoría de las células reciban solo uno de los ARN guía únicos. Por ejemplo, si los ARN guía se introducen mediante transducción vírica, las células pueden infectarse con una multiplicidad de infección baja para garantizar que la

mayoría de las células reciban una única construcción vírica con una alta probabilidad. Como un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

La población de células en la que se introduce la pluralidad de ARN guía únicos puede ser cualquier número adecuado de células. Por ejemplo, la población de células puede comprender más de aproximadamente 50, más de aproximadamente 100, más de aproximadamente 200, más de aproximadamente 300, más de aproximadamente 400, más de aproximadamente 500, más de aproximadamente 600, más de aproximadamente 700, más de aproximadamente 800, más de aproximadamente 900 o más de aproximadamente 1000 células por ARN guía único. En un ejemplo específico, la población de células comprende más de aproximadamente 300 células o más de aproximadamente 500 células por ARN guía único.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de genes. Por ejemplo, la pluralidad de ARN guía únicos puede dirigirse a aproximadamente 50 o más genes, aproximadamente 100 o más genes, aproximadamente 200 o más genes, aproximadamente 300 o más genes, aproximadamente 400 o más genes, aproximadamente 500 o más genes, aproximadamente 1000 o más genes, aproximadamente 2000 o más genes, aproximadamente 3000 o más genes, aproximadamente 4000 o más genes, aproximadamente 5000 o más genes, aproximadamente 10000 o más genes, o aproximadamente 20000 o más genes. En algunos métodos, los ARN guía pueden seleccionarse para dirigirse a genes en una vía de señalización particular. En algunos métodos, la biblioteca de ARN guía únicos es una biblioteca de todo el genoma.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de secuencias dentro de cada gen diana individual. En algunos métodos, se dirige a una pluralidad de secuencias diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, aproximadamente 2 a aproximadamente 10, aproximadamente 2 a aproximadamente 9, aproximadamente 2 a aproximadamente 8, aproximadamente 2 a aproximadamente 7, aproximadamente 2 a aproximadamente 6, aproximadamente 2 a aproximadamente 5, aproximadamente 2 a aproximadamente 4 o aproximadamente 2 a aproximadamente 3 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5 o al menos aproximadamente 6 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Como un ejemplo específico, aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana. Como otro ejemplo específico, aproximadamente 3 a aproximadamente 6 o aproximadamente 4 a aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana.

Los ARN guía pueden dirigirse a cualquier localización deseada en los genes diana. En algunos métodos CRISPRn que usan las células biosensoras Cas/tau, cada ARN guía se dirige a un exón constitutivo si es posible. En algunos métodos, cada ARN guía se dirige a un exón constitutivo en 5' si es posible. En algunos métodos, cada ARN guía se dirige a un primer exón, un segundo exón o un tercer exón (desde el extremo 5' del gen) si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, cada ARN guía puede dirigirse a una secuencia diana de ARN guía dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción, si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, en donde cada ARN guía puede comprender uno o más elementos de unión al adaptador a los que se puede unir específicamente la proteína adaptadora quimérica. En un ejemplo, cada ARN guía comprende dos elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica, opcionalmente en donde un primer elemento de unión al adaptador está dentro de un primer bucle de cada uno de los uno o más ARN guía, y un segundo elemento de unión al adaptador está dentro de un segundo bucle de cada uno de los uno o más ARN guía. Por ejemplo, el elemento de unión al adaptador puede comprender la secuencia establecida en SEQ ID NO: 33. En un ejemplo específico, cada uno de uno o más ARN guía es un único ARN guía que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de ARN CRISPR transactivador (ARNtracr), y el primer bucle es el tetrabucle correspondiente a los residuos 13-16 de SEQ ID NO: 17, 19, 30 o 31, y el segundo bucle es el tallo-bucle 2 correspondiente a los residuos 53-56 de SEQ ID NO: 17, 19, 30 o 31.

La etapa de cultivo de la población de células para permitir la edición y expansión del genoma puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 y aproximadamente 15 días, entre aproximadamente 3 y aproximadamente 13 días, entre aproximadamente 5 y aproximadamente 12 días, entre aproximadamente 7 y aproximadamente 11 días, o entre aproximadamente 7 y aproximadamente 11 días. Como otro ejemplo, el cultivo puede ser entre aproximadamente 2 días y aproximadamente 14 días, entre aproximadamente 3 días y aproximadamente 12 días, entre aproximadamente 5 días y aproximadamente 11 días, entre aproximadamente 7 días y aproximadamente 10 días, o aproximadamente 7 días o aproximadamente 10 días. Similarmente, la etapa de cultivo de la población de células para permitir la activación transcripcional y la expansión puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 15 días, entre aproximadamente 3 días y aproximadamente 13 días, entre aproximadamente 5 días y aproximadamente 12 días, entre aproximadamente 7 días y aproximadamente 11 días, o aproximadamente 7 días o aproximadamente 11 días. Similarmente, la etapa de cultivo de la población de células para permitir la activación transcripcional y la expansión puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 14 días, entre aproximadamente 3 días y aproximadamente 12 días, entre aproximadamente 5 días y aproximadamente 11 días, entre aproximadamente

7 días y aproximadamente 10 días, o aproximadamente 7 días o aproximadamente 10 días.

Puede usarse cualquier agente de siembra de tau adecuado para producir una población de células sembrada. Los agentes de siembra de tau adecuados se desvelan en el presente documento. Algunos agentes de siembra adecuados comprenden un dominio de repetición de tau que puede ser, por ejemplo, diferente, similar o igual al primer dominio de repetición de tau y/o al segundo dominio de repetición de tau. En un ejemplo, la etapa de siembra comprende cultivar la población de células modificada genéticamente en presencia de un lisado celular de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado. Opcionalmente, la población de células modificada genéticamente no se cocultiva con las células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado.

La cantidad o concentración del lisado celular en el medio puede ser cualquier cantidad o concentración adecuada. Por ejemplo, la concentración de lisado celular en el medio puede estar entre 0,1 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 25 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 4,5 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 3,5 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 3 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 2,5 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 2 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 1,5 µg/ml, entre aproximadamente 0,1 µg/ml y aproximadamente 1 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 25 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 4,5 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 3,5 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 3 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 2,5 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 2 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 1,5 µg/ml, entre aproximadamente 0,5 µg/ml y aproximadamente 1 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 25 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 4,5 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 3,5 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 3 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 2,5 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 2 µg/ml, entre aproximadamente 1 µg/ml y aproximadamente 1,5 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 25 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 4,5 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 3,5 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 3 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 2,5 µg/ml, entre aproximadamente 1,5 µg/ml y aproximadamente 2 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 4,5 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 3,5 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 3 µg/ml, entre aproximadamente 2 µg/ml y aproximadamente 2,5 µg/ml y aproximadamente 50 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 25 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 10 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 5 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 4,5 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 4 µg/ml, entre aproximadamente 2,5 µg/ml y aproximadamente 3,5 µg/ml, o entre aproximadamente 2,5 µg/ml y aproximadamente 3 µg/ml de medio (por ejemplo, medio de cultivo recién preparado). Por ejemplo, el lisado celular en el medio de cultivo puede estar en una concentración de entre aproximadamente 1 µg/ml y aproximadamente 5 µg/ml o pueden estar en una concentración de aproximadamente 1,5 µg/ml, aproximadamente 2 µg/ml, aproximadamente 2,5 µg/ml, aproximadamente 3 µg/ml, aproximadamente 3,5 µg/ml, aproximadamente 4 µg/ml, aproximadamente 4,5 µg/ml, o aproximadamente 5 µg/ml. Opcionalmente, el lisado celular puede estar en un tampón, tal como una solución salina tamponada con fosfato. Opcionalmente, el tampón puede comprender inhibidores de la proteasa. Ejemplos de inhibidores de la proteasa incluyen, pero no se limitan a, AEBSF, aprotinina, bestatina, E-64, leupeptina, pepstatina A y ácido etilendiaminotetracético (EDTA). El tampón puede comprender cualquiera de estos inhibidores o cualquier combinación de los mismos (por ejemplo, el tampón puede comprender todos estos inhibidores de la proteasa).

Las células para producir el lisado pueden obtenerse en un tampón, tal como solución salina tamponada con fosfato. Opcionalmente, el tampón puede comprender inhibidores de la proteasa. Ejemplos de inhibidores de la proteasa incluyen, pero no se limitan a, AEBSF, aprotinina, bestatina, E-64, leupeptina, pepstatina A y ácido etilendiaminotetracético (EDTA). El tampón puede comprender cualquiera de estos inhibidores o cualquier combinación de los mismos (por ejemplo, el tampón puede comprender todos estos inhibidores de la proteasa).

El lisado celular puede, por ejemplo, recogerse sonicando las células positivas a la agregación tau (por ejemplo, células recogidas en un tampón e inhibidores de la proteasa como se ha descrito anteriormente) durante cualquier cantidad

adecuada de tiempo. Por ejemplo, las células pueden sonicarse durante entre aproximadamente 1 minuto y aproximadamente 6 minutos, entre aproximadamente 1 minuto y aproximadamente 5 minutos, entre aproximadamente 1 minuto y aproximadamente 4 minutos, entre aproximadamente 1 minuto y aproximadamente 3 minutos, entre aproximadamente 2 minutos y aproximadamente 6 minutos, entre aproximadamente 2 minutos y aproximadamente 5 minutos, entre aproximadamente 2 minutos y aproximadamente 4 minutos, entre aproximadamente 2 minutos y aproximadamente 3 minutos, entre aproximadamente 2 minutos y aproximadamente 6 minutos, entre aproximadamente 3 minutos y aproximadamente 5 minutos, o entre aproximadamente 3 minutos y aproximadamente 4 minutos. Por ejemplo, las células pueden sonicarse durante entre aproximadamente 2 y aproximadamente 4 minutos o durante aproximadamente 3 minutos.

Opcionalmente, el medio comprende lipofectamina o liposomas (por ejemplo, liposomas catiónicos) o fosfolípidos u otro agente de transfección. Opcionalmente, el medio comprende lipofectamina. Opcionalmente, el medio no comprende lipofectamina o liposomas (por ejemplo, liposomas catiónicos) o fosfolípidos u otro agente de transfección. Opcionalmente, el medio no comprende lipofectamina. La cantidad o concentración de lipofectamina o liposomas (por ejemplo, liposomas catiónicos) o fosfolípidos u otro agente de transfección en el medio puede ser cualquier cantidad o concentración adecuada. Por ejemplo, la concentración de lipofectamina o liposomas (por ejemplo, catiónicos) o fosfolípidos u otro agente de transfección en el medio puede estar entre aproximadamente 0,5 µl/ml y aproximadamente 10 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 5 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 4,5 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 4 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 3,5 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 3 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 2,5 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 2 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 1,5 µl/ml, entre aproximadamente 0,5 µl/ml y aproximadamente 1 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 10 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 5 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 4,5 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 4 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 3,5 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 3 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 2,5 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 2 µl/ml, entre aproximadamente 1 µl/ml y aproximadamente 1,5 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 10 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 5 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 4,5 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 4 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 3,5 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 3 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 2,5 µl/ml, entre aproximadamente 1,5 µl/ml y aproximadamente 2 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 10 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 5 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 4,5 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 4 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 3,5 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 3 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 2,5 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 2 µl/ml, entre aproximadamente 2 µl/ml y aproximadamente 1,5 µl/ml, o entre aproximadamente 2 µl/ml y aproximadamente 2,5 µl/ml de medio (por ejemplo, medio fresco). Por ejemplo, la concentración de lipofectamina o liposomas (por ejemplo, liposomas catiónicos) o fosfolípidos u otro agente de transfección en el medio puede estar entre aproximadamente 1,5 µl/ml y aproximadamente 4 µl/ml o puede ser aproximadamente 1,5 µl/ml, aproximadamente 2 µl/ml, aproximadamente 2,5 µl/ml, aproximadamente 3 µl/ml, aproximadamente 3,5 µl/ml, o aproximadamente 4 µl/ml.

La etapa de cultivo de la población de células sembrada para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva para agregación, puede ser de cualquier duración adecuada. Por ejemplo, el cultivo puede ser durante entre aproximadamente 1 día y aproximadamente 7 días, entre aproximadamente 2 días y aproximadamente 6 días, entre aproximadamente 3 días y aproximadamente 5 días, entre aproximadamente 1 día y aproximadamente 3 días, o aproximadamente 2 días. La agregación puede determinarse por cualquier medio adecuado, dependiendo de los indicadores usados. Por ejemplo, en métodos en los que el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET), la población de células positiva a la agregación puede identificarse por citometría de flujo.

La abundancia de los ARN guía puede determinarse por cualquier medio adecuado. En un ejemplo específico, la abundancia se determina mediante secuenciación de nueva generación. La secuenciación de nueva generación se refiere a tecnologías de secuenciación de ADN de alto rendimiento no basadas en Sanger. Por ejemplo, la determinación de la abundancia de un ARN guía puede comprender la medición del recuento de lecturas del ARN guía.

En algunos métodos, un ARN guía se considera enriquecido en células negativas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o la población de células sembrada. En algunos métodos, un ARN guía se considera agotado en células positivas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o la población de células sembrada. También pueden usarse diferentes umbrales de enriquecimiento/agotamiento. Por ejemplo, un umbral de enriquecimiento/agotamiento puede fijarse más alto para ser más estricto (por ejemplo, al menos aproximadamente 1,6 veces, al menos aproximadamente 1,7 veces, al menos

aproximadamente 1,8 veces, al menos aproximadamente 1,9 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces o al menos aproximadamente 2,5 veces). Alternativamente, puede fijarse un umbral de enriquecimiento/agotamiento más bajo para ser menos estricto (por ejemplo, al menos 1,4 veces, al menos 1,3 veces o al menos 1,2 veces).

5 Alternativamente, en algunos métodos, un ARN guía se considera enriquecido en células positivas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o la población de células sembrada. En algunos métodos, un ARN guía se considera agotado en células negativas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o la población de células sembrada. También pueden usarse diferentes umbrales de enriquecimiento/agotamiento. Por ejemplo, un umbral de enriquecimiento/agotamiento puede fijarse más alto para ser más estricto (por ejemplo, al menos aproximadamente 1,6 veces, al menos aproximadamente 1,7 veces, al menos aproximadamente 1,8 veces, al menos aproximadamente 1,9 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces o al menos aproximadamente 2,5 veces). Alternativamente, puede fijarse un umbral de enriquecimiento/agotamiento más bajo para ser menos estricto (por ejemplo, al menos 1,4 veces, al menos 1,3 veces o al menos 1,2 veces).

20 En un ejemplo, la etapa de determinar la abundancia puede comprender determinar la abundancia de la pluralidad de ARN guía únicos en la población positiva a la agregación en relación con la población negativa a la agregación y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células sembrada en un segundo momento. Similarmente, la etapa de determinación de la abundancia puede comprender la determinación de la abundancia de la pluralidad de ARN guía únicos en la población negativa a la agregación en relación con la población positiva a la agregación y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células sembrada en un segundo momento. Por ejemplo, el primer momento puede estar en un primer pase del cultivo de la población de células o en medio del cultivo de la población de células para permitir la edición y expansión del genoma. Por ejemplo, el primer momento puede ser después de una cantidad de tiempo suficiente para que los ARN guía formen complejos con la proteína Cas, y para que la proteína Cas corte la pluralidad de genes resultando en la inactivación de la función génica (CRISPRn) o para activar transcripcionalmente la pluralidad de genes (CRISPRa). Sin embargo, en algunos casos, el primer momento debería ser idealmente en el primer pase celular para determinar la representación de la biblioteca de ARNg poco después de la infección (es decir, antes de la expansión posterior y la edición del genoma) y para determinar si cada representación de ARNg evoluciona desde el primer momento al segundo y a cualquier momento adicional hasta un momento final. Esto permite descartar los ARNg/dianas enriquecidos debido a las ventajas del crecimiento celular durante el transcurso del cribado mediante la verificación de que la abundancia de ARNg no cambia entre el primer y el segundo momento. Como un ejemplo específico, el primer momento puede ser después de aproximadamente 1 día, aproximadamente 2 días, aproximadamente 3 días, o aproximadamente 4 días de cultivo y expansión (por ejemplo, a aproximadamente 3 días de cultivo y expansión), y el segundo momento puede ser después de aproximadamente 5 días, aproximadamente 6 días, aproximadamente 7 días, aproximadamente 8 días, aproximadamente 9 días, aproximadamente 10 días, o aproximadamente 11 días de cultivo y expansión. Por ejemplo, el primer momento puede ser después de aproximadamente 3 días de cultivo y expansión, y el segundo momento puede ser después de aproximadamente 7 días o aproximadamente 10 días de cultivo y expansión.

45 En algunos métodos, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen impide (o se espera que impida) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación, la población de células cultivada en el primer momento y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen impide (o se espera que impida) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen impide (o se espera que impida) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación, la población de células cultivada en el primer momento y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen impide (o se espera que impida) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y la población de células sembrada en el segundo

momento.

Alternativamente, en un ejemplo, la etapa de determinar la abundancia puede comprender determinar la abundancia de la pluralidad de ARN guía únicos en la población negativa a la agregación en relación con la población positiva a la agregación y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células sembrada en un segundo momento. Similarmente, la etapa de determinación de la abundancia puede comprender la determinación de la abundancia de la pluralidad de ARN guía únicos en la población positiva a la agregación en relación con la población negativa a la agregación y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células sembrada en un segundo momento. Por ejemplo, el primer momento puede estar en un primer pase del cultivo de la población de células o en medio del cultivo de la población de células para permitir la edición y expansión del genoma. Por ejemplo, el primer momento puede ser después de una cantidad de tiempo suficiente para que los ARN guía formen complejos con la proteína Cas, y para que la proteína Cas corte la pluralidad de genes resultando en la inactivación de la función génica (CRISPRn) o para activar transcripcionalmente la pluralidad de genes (CRISPRa). Sin embargo, en algunos casos, el primer momento debería ser idealmente en el primer pase celular para determinar la representación de la biblioteca de ARNg poco después de la infección (es decir, antes de la expansión posterior y la edición del genoma) y para determinar si cada representación de ARNg evoluciona desde el primer momento al segundo y a cualquier momento adicional hasta un momento final. Esto permite descartar los ARNg/dianas enriquecidos debido a las ventajas del crecimiento celular durante el transcurso del cribado mediante la verificación de que la abundancia de ARNg no cambia entre el primer y el segundo momentos. Como un ejemplo específico, el primer momento puede ser después de aproximadamente 1 día, aproximadamente 2 días, aproximadamente 3 días, o aproximadamente 4 días de cultivo y expansión (por ejemplo, a aproximadamente 3 días de cultivo y expansión), y el segundo momento puede ser después de aproximadamente 5 días, aproximadamente 6 días, aproximadamente 7 días, aproximadamente 8 días, aproximadamente 9 días, aproximadamente 10 días, o aproximadamente 11 días de cultivo y expansión. Por ejemplo, el primer momento puede ser después de aproximadamente 3 días de cultivo y expansión, y el segundo momento puede ser después de aproximadamente 7 días o unos 10 días de cultivo y expansión.

En algunos métodos, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia la agregación de tau (o se espera que promueva o potencie la agregación de tau), si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación, la población de células cultivada en el primer momento y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia la agregación de tau (o se espera que promueva o potencie la agregación de tau), si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia la agregación de tau (o se espera que promueva o potencie la agregación de tau), si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación, la población de células cultivada en el primer momento y la población de células sembrada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia la agregación de tau (o se espera que promueva o potencie la agregación de tau), si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces inferior en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y la población de células sembrada en el segundo momento.

En algunos métodos CRISPRn, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación del gen impide (o se espera que impida) la agregación de tau. Similarmente, en algunos métodos CRISPRa, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la activación transcripcional del gen impide (o se espera que impida) la agregación de tau. La primera etapa comprende identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células negativa a la agregación. La segunda etapa comprende calcular la probabilidad aleatoria de que los ARN guía identificados estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de ARN guía únicos introducidos en la población de células, m es la variedad de ARN guía únicos identificados en la etapa (1), n es la variedad de ARN guía únicos introducidos en la población de células que se dirigen al gen y n' es la variedad de ARN guía únicos identificados en la etapa (1) que se dirigen al gen. La tercera etapa comprende calcular puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (1). La puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células negativa a la agregación dividida por la abundancia relativa del ARN guía en las células positivas a la agregación o la población de células cultivada tras la introducción de la biblioteca de ARN guía o la población de células sembrada. La abundancia

relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos. La cuarta etapa comprende seleccionar el gen si un ARN guía que se dirige al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral. Las posibles puntuaciones de enriquecimiento umbral se han tratado anteriormente. Como un ejemplo específico, la puntuación de enriquecimiento umbral puede fijarse en aproximadamente 1,5 veces.

Alternativamente, en algunos métodos CRISPRn, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau. Similarmente, en algunos métodos CRISPRa, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la activación transcripcional del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau. La primera etapa comprende identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación. La segunda etapa comprende calcular la probabilidad aleatoria de que los ARN guía identificados estén presentes usando la fórmula $nCn' * (x-n)C(m-n) / xCm$, donde x es la variedad de ARN guía únicos introducidos en la población de células, m es la variedad de ARN guía únicos identificados en la etapa (1), n es la variedad de ARN guía únicos introducidos en la población de células que se dirigen al gen y n' es la variedad de ARN guía únicos identificados en la etapa (1) que se dirigen al gen. La tercera etapa comprende calcular puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (1). La puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación dividida por la abundancia relativa del ARN guía en las células negativas a la agregación o la población de células cultivada tras la introducción de la biblioteca de ARN guía o la población de células sembrada. La abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos. La cuarta etapa comprende seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral. Las posibles puntuaciones de enriquecimiento umbral se han tratado anteriormente. Como un ejemplo específico, la puntuación de enriquecimiento umbral puede fijarse en aproximadamente 1,5 veces.

Variedad cuando se usa en la frase variedad de ARN guía únicos significa el número de secuencias de ARN guía únicos. No se trata de la abundancia, sino más bien del cualitativo "presente" o "no presente". Por variedad de ARN guía únicos se entiende el número de secuencias de ARN guía únicos. La variedad de ARN guía únicos se determina mediante secuenciación de nueva generación (NGS) para identificar todos los ARN guía únicos presentes en una población celular. Se hace usando dos cebadores que reconocen las regiones constantes del vector vírico para amplificar el ARNg que está entre las regiones constantes y un cebador que reconoce una región constante para la secuenciación. Cada ARN guía único presente en la muestra generará recuentos de lecturas usando el cebador de secuenciación. Los resultados de la NGS incluirán la secuencia y también el número de lecturas correspondientes a la secuencia. El número de lecturas se usará para el cálculo de la puntuación de enriquecimiento de cada ARN guía, y la presencia de cada secuencia única nos indicará qué ARN guía está presente. Por ejemplo, si hay tres ARN guía únicos para un gen antes de la selección y los tres se conservan después de la selección, entonces n y n' son 3. Estos números se usan para calcular las estadísticas, pero no los recuentos de lecturas reales. Sin embargo, los recuentos de lecturas para cada ARN guía (en un ejemplo, 100, 200, 50, que corresponden a cada uno de los 3 ARN guía únicos) se usarán para el cálculo de la puntuación de enriquecimiento.

VI. Métodos de cribado de modificadores genéticos de la desagregación de tau

Las líneas celulares biosensoras Cas/tau desveladas en el presente documento pueden usarse en métodos de cribado de modificadores genéticos de la desagregación de tau (por ejemplo, que promuevan la desagregación de tau). Dichos métodos pueden comprender proporcionar una población de células biosensoras Cas/tau positivas a la desagregación de tau como se desvela en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la desagregación de tau en las células diana.

Como un ejemplo, un método puede comprender proporcionar una población de células biosensoras Cas/tau (por ejemplo, una población de células que comprende una proteína Cas, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador), en donde las células son células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, introduciendo en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes, y cultivando la población de células para permitir la edición y expansión del genoma. La pluralidad de ARN guía únicos forman complejos con la proteína Cas, y la proteína Cas corta la pluralidad de genes dando lugar a la inactivación de la función génica para producir una población de células modificada genéticamente. El cultivo da como resultado una población de células positiva a la agregación y una población de células negativa a la agregación, que pueden ser identificadas. Por último, puede determinarse la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía en uno o más momentos. El enriquecimiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la desagregación de tau, en

donde la inactivación del gen al que se dirige el ARN guía promueve la desagregación de tau, o es un modificador genético candidato de la desagregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva la desagregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la desagregación de tau, en donde la inactivación del gen al que se dirige el ARN guía promueve la desagregación de tau, o es un modificador genético candidato de la desagregación de tau (por ejemplo, para pruebas adicionales mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva la desagregación de tau. El enriquecimiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la inactivación del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva o potencie la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la alteración del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la inactivación del gen al que se dirige el ARN guía promueva o potencie la agregación de tau.

De forma similar, las líneas celulares biosensoras SAM/tau desveladas en el presente documento pueden usarse en métodos de cribado de modificadores genéticos de la desagregación de tau (por ejemplo, que promuevan la desagregación de tau). Dichos métodos pueden comprender proporcionar una población de células SAM/tau biosensoras positivas a la desagregación de tau como se desvela en otra parte del presente documento, introducir una biblioteca que comprende una pluralidad de ARN guía únicos y evaluar la desagregación de tau en las células diana.

Como un ejemplo, un método puede comprender proporcionar una población de células biosensoras SAM/tau (por ejemplo, una población de células que comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, una proteína adaptadora quimérica que comprende una proteína adaptadora fusionada a uno o más dominios de activación transcripcional, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador), en donde las células son células positivas a la agregación de tau en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, introduciendo en la población de células una biblioteca que comprende una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes, y cultivando la población de células para permitir la activación y la expansión transcripcional. La pluralidad de ARN guía únicos forma complejos con la proteína Cas quimérica y la proteína adaptadora quimérica, y los complejos activan la transcripción de la pluralidad de genes dando lugar a un aumento de la expresión génica y a una población de células modificada. El cultivo da como resultado una población de células positiva a la agregación y una población de células negativa a la agregación, que pueden ser identificadas. Por último, puede determinarse la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos. El enriquecimiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la desagregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía promueve la desagregación de tau, o es un modificador genético candidato de la desagregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía promueva la desagregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o en relación con la población de células que se cultivan tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para su posterior análisis mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía

promueva o potencie la agregación de tau. Similarmente, el agotamiento de un ARN guía en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o en relación con la población de células que se cultiva tras la introducción de la biblioteca de ARN guía en uno o más momentos indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía promueve o potencia la agregación de tau, o es un modificador genético candidato de la agregación de tau (por ejemplo, para pruebas posteriores mediante cribados secundarios), en donde se espera que la activación transcripcional del gen al que se dirige el ARN guía promueva o potencie la agregación de tau.

Las células biosensoras Cas/tau usadas en el método pueden ser cualquiera de las células biosensoras Cas/tau desveladas en otra parte del presente documento. Similarmente, las células biosensoras SAM/tau usadas en el método pueden ser cualquiera de las células biosensoras SAM/tau desveladas en otra parte del presente documento. El primer dominio de repetición de tau y el segundo dominio de repetición de tau pueden ser diferentes o pueden ser similares o iguales. El dominio de repetición de tau puede ser cualquiera de los dominios de repetición de tau desvelados en otra parte del presente documento. Por ejemplo, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau puede ser un dominio de repetición de tau natural o puede comprender una mutación pro-agregación (por ejemplo, una mutación patógena, pro-agregación), tal como una mutación de tau P301S. El primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender un dominio de cuatro repeticiones de tau. Como un ejemplo específico, el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau pueden comprender, consistir esencialmente en, o consistir en SEQ ID NO: 11 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 11. En un ejemplo específico, el ácido nucleico que codifica el dominio de repetición de tau puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 12 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 12, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 11.

El primer dominio de repetición de tau puede unirse al primer indicador y el segundo dominio de repetición de tau puede unirse al segundo indicador por cualquier medio. Por ejemplo, el indicador puede fusionarse al dominio de repetición de tau (por ejemplo, como parte de una proteína de fusión). Las proteínas indicadoras son un par de proteínas indicadoras fluorescentes que producen una señal detectable cuando el primer dominio de repetición de tau unido al primer indicador se agrega con el segundo dominio de repetición de tau unido al segundo indicador. En particular, la primera y segunda proteínas indicadoras son un par de transferencia de energía por resonancia de fluorescencia (FRET). FRET es un fenómeno físico en el que un fluoróforo donante en su estado excitado transfiere de forma no radiativa su energía de excitación a un fluoróforo aceptor vecino, provocando así que el aceptor emita su fluorescencia característica. Los ejemplos de pares FRET (fluoróforos donante y aceptor) son bien conocidos. Véase, por ejemplo, Bajar et al. (2016) *Sensors* (Basel) 16(9): 1488, incorporado en el presente documento por referencia en su totalidad a todos los efectos. Como un ejemplo específico de un par FRET, el primer indicador puede ser proteína fluorescente cian (CFP) y el segundo indicador puede ser proteína fluorescente amarilla (YFP). Como un ejemplo específico, la CFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 13 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 13. Como otro ejemplo específico, la YFP puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 15 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 15.

Para las células biosensoras Cas/tau, la proteína Cas puede ser cualquier proteína Cas desvelada en otra parte del presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 21 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 21.

Una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 22 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 22, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 21.

Para las células biosensoras SAM/tau, la proteína Cas puede ser cualquier proteína Cas desvelada en otra parte del presente documento. Como un ejemplo, la proteína Cas puede ser una proteína Cas9. Por ejemplo, la proteína Cas9 puede ser una proteína Cas9 de *Streptococcus pyogenes*. Como un ejemplo específico, la proteína Cas química puede comprender la proteína Cas inactiva por nucleasa fusionada a un dominio de activación transcripcional VP64. Por ejemplo, la proteína Cas química puede comprender de extremo N a extremo C: la proteína Cas inactiva por nucleasa; una señal de localización nuclear; y el dominio activador transcripcional VP64. Como un ejemplo específico,

la proteína adaptadora puede ser una proteína de la cubierta de MS2, y el uno o más dominios de activación transcripcional en la proteína adaptadora quimérica puede comprender un dominio de activación transcripcional p65 y un dominio de activación transcripcional HSF1. Por ejemplo, la proteína adaptadora quimérica puede comprender de extremo N a extremo C: la proteína de la cubierta de MS2; una señal de localización nuclear; el dominio de activación transcripcional p65; y el dominio de activación transcripcional HSF1. En un ejemplo específico, el ácido nucleico que codifica la proteína Cas quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 38 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 38, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 36. En un ejemplo específico, el ácido nucleico que codifica la proteína adaptadora quimérica puede comprender, consistir esencialmente en, o consistir en SEQ ID NO: 39 o una secuencia al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica a SEQ ID NO: 39, opcionalmente en donde el ácido nucleico codifica una proteína que comprende, consiste esencialmente en, o consiste en SEQ ID NO: 37.

Una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden expresarse de forma estable en la población de células. Por ejemplo, los ácidos nucleicos que codifican una o más o todas las proteínas Cas quiméricas, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador pueden integrarse genómicamente en la población de células.

Como se desvela en otra parte del presente documento, las células pueden ser de cualquier tipo. Por ejemplo, las células pueden ser células eucariotas, células de mamífero o células humanas (por ejemplo, células HEK293T o células neuronales).

La pluralidad de ARN guía únicos puede introducirse en la población de células por cualquier medio conocido. En algunos métodos, los ARN guía se introducen en la población de células mediante transducción vírica, tal como transducción retrovírica, adenovírica o lentivírica. En un ejemplo específico, los ARN guía pueden introducirse mediante transducción lentivírica. Cada uno de la pluralidad de ARN guía únicos puede estar en un vector vírico separado. La población de células puede infectarse con cualquier multiplicidad de infección. Por ejemplo, la multiplicidad de infección puede estar comprendida entre aproximadamente 0,1 y aproximadamente 1,0, entre aproximadamente 0,1 y aproximadamente 0,9, entre aproximadamente 0,1 y aproximadamente 0,8, entre aproximadamente 0,1 y aproximadamente 0,7, entre aproximadamente 0,1 y aproximadamente 0,6, entre aproximadamente 0,1 y aproximadamente 0,5, entre aproximadamente 0,1 y aproximadamente 0,4, o entre aproximadamente 0,1 y aproximadamente 0,3. Alternativamente, la multiplicidad de infección puede ser inferior a aproximadamente 1,0, inferior a aproximadamente 0,9, inferior a aproximadamente 0,8, inferior a aproximadamente 0,7, inferior a aproximadamente 0,6, inferior a aproximadamente 0,5, inferior a aproximadamente 0,4, inferior a aproximadamente 0,3 o inferior a aproximadamente 0,2. En un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

Los ARN guía pueden introducirse en la población de células junto con un marcador de selección o un gen indicador para seleccionar células que tengan los ARN guía, y el método puede comprender además la selección de células que comprendan el marcador de selección o el gen indicador. En otra parte del presente documento se proporcionan ejemplos de marcadores de selección y genes indicadores. Como un ejemplo, el marcador de selección puede ser uno que imparta resistencia a un fármaco, tal como neomicina fosfotransferasa, higromicina B fosfotransferasa, puromicina-N-acetiltransferasa y blasticidina S desaminasa. Otro marcador de selección a modo de ejemplo es la proteína de resistencia a la bleomicina, codificada por el gen *Sh ble* (gen de la bleomicina *del Streptoalloteichus hindustanus*), que confiere resistencia a la zeocina (fleomicina D1). Por ejemplo, las células pueden seleccionarse con un fármaco (por ejemplo, puromicina) de modo que solo las células transducidas con una construcción de ARN guía se conserven para ser usadas para llevar a cabo el cribado. Por ejemplo, el fármaco puede ser puromicina o zeocina (fleomicina D1).

En algunos métodos, la pluralidad de ARN guía únicos se introduce a una concentración seleccionada de forma que la mayoría de las células reciban solo uno de los ARN guía únicos. Por ejemplo, si los ARN guía se introducen mediante transducción vírica, las células pueden infectarse con una multiplicidad de infección baja para garantizar que la mayoría de las células reciban una única construcción vírica con una alta probabilidad. Como un ejemplo específico, la multiplicidad de infección puede ser inferior a aproximadamente 0,3.

La población de células en la que se introduce la pluralidad de ARN guía únicos puede ser cualquier número adecuado de células. Por ejemplo, la población de células puede comprender más de aproximadamente 50, más de aproximadamente 100, más de aproximadamente 200, más de aproximadamente 300, más de aproximadamente 400, más de aproximadamente 500, más de aproximadamente 600, más de aproximadamente 700, más de aproximadamente 800, más de aproximadamente 900 o más de aproximadamente 1000 células por ARN guía único. En un ejemplo específico, la población de células comprende más de aproximadamente 300 células o más de aproximadamente 500 células por ARN guía único.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de genes. Por ejemplo, la pluralidad de ARN guía únicos puede dirigirse a aproximadamente 50 o más genes, aproximadamente 100 o más genes, aproximadamente 200 o más genes, aproximadamente 300 o más genes, aproximadamente 400 o más genes, aproximadamente 500 o más genes, aproximadamente 1000 o más genes, aproximadamente 2000 o más genes, aproximadamente 3000 o más genes, aproximadamente 4000 o más genes, aproximadamente 5000 o más genes, aproximadamente 10000 o más genes, o aproximadamente 20000 o más genes. En algunos métodos, los ARN guía pueden seleccionarse para dirigirse a genes en una vía de señalización particular. En algunos métodos, la biblioteca de ARN guía únicos es una biblioteca de todo el genoma.

La pluralidad de ARN guía únicos puede dirigirse a cualquier número de secuencias dentro de cada gen diana individual. En algunos métodos, se dirige a una pluralidad de secuencias diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, aproximadamente 2 a aproximadamente 10, aproximadamente 2 a aproximadamente 9, aproximadamente 2 a aproximadamente 8, aproximadamente 2 a aproximadamente 7, aproximadamente 2 a aproximadamente 6, aproximadamente 2 a aproximadamente 5, aproximadamente 2 a aproximadamente 4 o aproximadamente 2 a aproximadamente 3 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Por ejemplo, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5 o al menos aproximadamente 6 secuencias diana únicas pueden ser diana de media en cada uno de la pluralidad de genes diana. Como un ejemplo específico, aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana. Como otro ejemplo específico, aproximadamente 3 a aproximadamente 6 o aproximadamente 4 a aproximadamente 6 secuencias diana pueden ser diana de media en cada uno de la pluralidad de genes diana.

Los ARN guía pueden dirigirse a cualquier localización deseada en los genes diana. En algunos métodos CRISPRn que usan las células biosensoras Cas/tau, cada ARN guía se dirige a un exón constitutivo si es posible. En algunos métodos, cada ARN guía se dirige a un exón constitutivo en 5' si es posible. En algunos métodos, cada ARN guía se dirige a un primer exón, un segundo exón o un tercer exón (desde el extremo 5' del gen) si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, cada ARN guía puede dirigirse a una secuencia diana de ARN guía dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción, si es posible. En algunos métodos CRISPRa que usan las células biosensoras SAM/tau, en donde cada ARN guía puede comprender uno o más elementos de unión al adaptador a los que se puede unir específicamente la proteína adaptadora quimérica. En un ejemplo, cada ARN guía comprende dos elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica, opcionalmente en donde un primer elemento de unión al adaptador está dentro de un primer bucle de cada uno de los uno o más ARN guía, y un segundo elemento de unión al adaptador está dentro de un segundo bucle de cada uno de los uno o más ARN guía. Por ejemplo, el elemento de unión al adaptador puede comprender la secuencia establecida en SEQ ID NO: 33. En un ejemplo específico, cada uno de uno o más ARN guía es un único ARN guía que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de ARN CRISPR transactivador (ARNtracr), y el primer bucle es el tetrabucle correspondiente a los residuos 13-16 de SEQ ID NO: 17, 19, 30 o 31, y el segundo bucle es el tallo-bucle 2 correspondiente a los residuos 53-56 de SEQ ID NO: 17, 19, 30 o 31.

La etapa de cultivar la población de células para permitir la edición y expansión del genoma puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 15 días, entre aproximadamente 3 días y aproximadamente 14 días, entre aproximadamente 5 días y aproximadamente 14 días, entre aproximadamente 7 días y aproximadamente 14 días, entre aproximadamente 9 días y aproximadamente 14 días, entre aproximadamente 10 días y aproximadamente 14 días, entre aproximadamente 11 días y aproximadamente 14 días, o entre aproximadamente 12 días y aproximadamente 14 días. Similarmente, la etapa de cultivo de la población de células para permitir la activación transcripcional y la expansión puede ser de cualquier periodo de tiempo adecuado. Por ejemplo, el cultivo puede ser durante entre aproximadamente 2 días y aproximadamente 15 días, entre aproximadamente 3 días y aproximadamente 14 días, entre aproximadamente 5 días y aproximadamente 14 días, entre aproximadamente 7 días y aproximadamente 14 días, entre aproximadamente 9 días y aproximadamente 14 días, entre aproximadamente 10 días y aproximadamente 14 días, entre aproximadamente 11 días y aproximadamente 14 días, o entre aproximadamente 12 días y aproximadamente 14 días.

El paso de identificar la población de células positiva a la agregación y la población de células negativa a la agregación puede comprender la sincronización de la progresión del ciclo celular. Por ejemplo, esto puede usarse para obtener una población celular predominantemente enriquecida en la fase S o predominantemente enriquecida en la fase G2. Como un ejemplo específico, la sincronización puede lograrse mediante un bloqueo doble con timidina.

La agregación puede determinarse por cualquier medio adecuado, dependiendo de los indicadores usados. Por ejemplo, en los métodos en los que el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET), la población de células positiva a la agregación puede identificarse por citometría de flujo.

La abundancia de los ARN guía puede determinarse por cualquier medio adecuado. En un ejemplo específico, la abundancia se determina mediante secuenciación de nueva generación. La secuenciación de nueva generación se refiere a tecnologías de secuenciación de ADN de alto rendimiento no basadas en Sanger. Por ejemplo, la

determinación de la abundancia de un ARN guía puede comprender la medición de recuentos de lecturas del ARN guía.

5 En algunos métodos, un ARN guía se considera enriquecido en células negativas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o la población de células cultivada en uno o más momentos. En algunos métodos, un ARN guía se considera agotado en células positivas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o la población de células cultivada en uno o más momentos. También pueden usarse diferentes umbrales de enriquecimiento/agotamiento. Por ejemplo, un umbral de enriquecimiento/agotamiento puede fijarse más alto para ser más estricto (por ejemplo, al menos aproximadamente 1,6 veces, al menos aproximadamente 1,7 veces, al menos aproximadamente 1,8 veces, al menos aproximadamente 1,9 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces o al menos aproximadamente 2,5 veces). Alternativamente, puede fijarse un umbral de enriquecimiento/agotamiento más bajo para ser menos estricto (por ejemplo, al menos aproximadamente 1,4 veces, al menos aproximadamente 1,3 veces o al menos aproximadamente 1,2 veces).

20 En algunos métodos, un ARN guía se considera enriquecido en células positivas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y/o la población de células cultivada en uno o más momentos. En algunos métodos, un ARN guía se considera agotado en células negativas a la agregación si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y/o la población de células cultivada en uno o más momentos. También pueden usarse diferentes umbrales de enriquecimiento/agotamiento. Por ejemplo, un umbral de enriquecimiento/agotamiento puede fijarse más alto para ser más estricto (por ejemplo, al menos aproximadamente 1,6 veces, al menos aproximadamente 1,7 veces, al menos aproximadamente 1,8 veces, al menos aproximadamente 1,9 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces o al menos aproximadamente 2,5 veces). Alternativamente, puede fijarse un umbral de enriquecimiento/agotamiento más bajo para ser menos estricto (por ejemplo, al menos aproximadamente 1,4 veces, al menos aproximadamente 1,3 veces o al menos aproximadamente 1,2 veces).

35 En un ejemplo, la etapa de determinar la abundancia puede comprender determinar la abundancia de la pluralidad de ARN guía únicos en la población positiva a la agregación en relación con la población negativa a la agregación y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células cultivada después de la introducción de la biblioteca de ARN guía en un segundo momento. Similarmente, la etapa de determinar la abundancia puede comprender determinar la abundancia de la pluralidad de ARN guía únicos en la población negativa a la agregación en relación con la población positiva a la agregación y/o en relación con la población de células cultivada tras la introducción de la biblioteca de ARN guía en un primer momento y/o en relación con la población de células cultivada tras la introducción de la biblioteca de ARN guía en un segundo momento. Por ejemplo, el primer momento puede ser en un primer pase del cultivo de la población de células o en medio del cultivo de la población de células para permitir la edición y expansión del genoma, y el segundo momento puede estar en medio del cultivo de la población de células para permitir la edición y expansión del genoma. Por ejemplo, el primer momento puede ser después de una cantidad de tiempo suficiente para que los ARN guía formen complejos con la proteína Cas, y para que la proteína Cas corte la pluralidad de genes resultando en la inactivación de la función génica (CRISPRn) o active transcripcionalmente la pluralidad de genes (CRISPRa). Sin embargo, en algunos casos, el primer momento debería ser idealmente en el primer pase celular para determinar la representación de la biblioteca de ARNg poco después de la infección (es decir, antes de la expansión posterior y la edición del genoma) y para determinar si cada representación de ARNg evoluciona desde el primer momento al segundo y a cualquier momento adicional hasta un momento final. Esto permite descartar los ARNg/diana enriquecidos debido a las ventajas del crecimiento celular durante el transcurso del cribado mediante la verificación de que la abundancia de ARNg no cambia entre el primer y segundo momentos. Como un ejemplo específico, el primer momento puede ser después de aproximadamente 1 día, aproximadamente 2 días, aproximadamente 3 días, aproximadamente 4 días, aproximadamente 5 días, aproximadamente 6 días, aproximadamente 7 días, o aproximadamente 8 días de cultivo y expansión (por ejemplo, a aproximadamente 7 días de cultivo y expansión), y el segundo momento puede ser después de aproximadamente 5 días, aproximadamente 6 días, aproximadamente 7 días, aproximadamente 8 días, aproximadamente 9 días, aproximadamente 10 días, aproximadamente 11 días, o aproximadamente 12 días (por ejemplo, aproximadamente 10 días) de cultivo y expansión. Por ejemplo, el primer momento puede ser después de aproximadamente 7 días de cultivo y expansión, y el segundo momento puede ser después de aproximadamente 10 días de cultivo y expansión.

65 En algunos métodos, un gen puede considerarse un modificador genético de la desagregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve (o se espera que promueva) la desagregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación en

relación con la población de células positiva a la agregación, la población de células cultivada en el primer momento y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la desagregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve (o se espera que promueva) la desagregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la desagregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve (o se espera que promueva) la desagregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación, la población de células cultivada en el primer momento y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede entonces considerarse un modificador genético de la desagregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve (o se espera que promueva) la desagregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y la población de células cultivada en el segundo momento.

En algunos métodos, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación, la población de células cultivada en el primer momento y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede entonces considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en relación con la población de células negativa a la agregación y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células negativa a la agregación en relación con la población de células positiva a la agregación, la población de células cultivada en el primer momento y la población de células cultivada en el segundo momento. Alternativa o adicionalmente, un gen puede entonces considerarse un modificador genético de la agregación de tau, en donde la inactivación (CRISPRn) o la activación transcripcional (CRISPRa) del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau, si la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces menor en la población de células con agregación de tau en relación con la población de células positiva a la agregación y la población de células cultivada en el segundo momento.

En algunos métodos CRISPRn, se siguen las siguientes etapas para identificar un gen como modificador genético de la desagregación de tau, en donde la inactivación del gen promueve (o se espera que promueva) la desagregación de tau. Similarmente, en algunos métodos CRISPRa, se siguen las siguientes etapas para identificar un gen como modificador genético de la desagregación de tau, en donde la activación transcripcional del gen promueve (o se espera que promueva) la desagregación de tau. La primera etapa comprende identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células negativa a la agregación. La segunda etapa comprende calcular la probabilidad aleatoria de que los ARN guía identificados estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de ARN guía únicos introducidos en la población de células, m es la variedad de ARN guía únicos identificados en la etapa (1), n es la variedad de ARN guía únicos introducidos en la población de células que se dirigen al gen y n' es la variedad de ARN guía únicos identificados en la etapa (1) que se dirigen al gen. La tercera etapa comprende calcular puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (1). La puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células negativa a la agregación dividida por la abundancia relativa del ARN guía en las células positivas a la agregación o la población de células cultivada después de la introducción de la biblioteca de ARN guía en el primer o segundo momento. La abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos. La cuarta etapa comprende seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral. Las posibles puntuaciones de enriquecimiento umbral se han tratado anteriormente. Como un ejemplo específico, la puntuación de enriquecimiento umbral puede fijarse en aproximadamente 1,5 veces.

En algunos métodos CRISPRn, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación del gen promueve o potencia (o se espera que promueva o potencie) la

agregación de tau. Similarmente, en algunos métodos CRISPRa, se siguen las siguientes etapas para identificar un gen como modificador genético de la agregación de tau, en donde la activación transcripcional del gen promueve o potencia (o se espera que promueva o potencie) la agregación de tau. La primera etapa comprende identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación. La segunda etapa comprende calcular la probabilidad aleatoria de que los ARN guía identificados estén presentes usando la fórmula $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de ARN guía únicos introducidos en la población de células, m es la variedad de ARN guía únicos identificados en la etapa (1), n es la variedad de ARN guía únicos introducidos en la población de células que se dirigen al gen, y n' es la variedad de ARN guía únicos identificados en la etapa (1) que se dirigen al gen. La tercera etapa comprende calcular puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (1). La puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación dividida por la abundancia relativa del ARN guía en las células negativas a la agregación o la población de células cultivada después de la introducción de la biblioteca de ARN guía en el primer o segundo momento. La abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos. La cuarta etapa comprende seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral. Las posibles puntuaciones de enriquecimiento umbral se han tratado anteriormente. Como un ejemplo específico, la puntuación de enriquecimiento umbral puede fijarse en aproximadamente 1,5 veces.

Variedad cuando se usa en la frase variedad de ARN guía únicos significa el número de secuencias de ARN guía únicos. No se trata de la abundancia, sino más bien del cualitativo "presente" o "no presente". Por variedad de ARN guía únicos se entiende el número de secuencias de ARN guía únicos. La variedad de ARN guía únicos se determina mediante secuenciación de nueva generación (NGS) para identificar todos los ARN guía únicos presentes en una población celular. Se hace usando dos cebadores que reconocen las regiones constantes del vector vírico para amplificar el ARNg que está entre las regiones constantes y un cebador que reconoce una región constante para la secuenciación. Cada ARN guía único presente en la muestra generará recuentos de lecturas usando el cebador de secuenciación. Los resultados de la NGS incluirán la secuencia y también el número de lecturas correspondientes a la secuencia. El número de lecturas se usará para el cálculo de la puntuación de enriquecimiento de cada ARN guía, y la presencia de cada secuencia única nos indicará qué ARN guía está presente. Por ejemplo, si hay tres ARN guía únicos para un gen antes de la selección y los tres se conservan después de la selección, entonces n y n' son 3. Estos números se usan para calcular las estadísticas, pero no los recuentos de lecturas reales. Sin embargo, los recuentos de lecturas para cada ARN guía (en un ejemplo, 100, 200, 50, que corresponden a cada uno de los 3 ARN guía únicos) se usarán para el cálculo de la puntuación de enriquecimiento.

Si diferentes versiones de una secuencia están asociadas a un número de acceso en diferentes momentos, se entenderá la versión asociada al número de acceso en la fecha de presentación efectiva de la presente solicitud. Por fecha de presentación efectiva se entiende la fecha de presentación real o la fecha de presentación de una solicitud de prioridad que haga referencia al número de acceso, si procede, la que sea anterior. Similarmente, si se publican diferentes versiones de una publicación, sitio web o similar en diferentes momentos, se entenderá la versión publicada más recientemente en la fecha de presentación efectiva de la solicitud, a menos que se indique lo contrario. Cualquier característica, etapa, elemento, realización o aspecto de la invención puede usarse en combinación con cualquier otro, a menos que se indique específicamente lo contrario. Aunque la presente invención se ha descrito con cierto detalle a modo de ilustración y ejemplo con fines de claridad y comprensión, será evidente que ciertos cambios y modificaciones pueden practicarse dentro del alcance de las reivindicaciones adjuntas.

Breve descripción de las secuencias

Las secuencias de nucleótidos y aminoácidos que figuran en el listado de secuencias adjunto se muestran usando abreviaturas de letras estándar para las bases nucleotídicas y códigos de tres letras para los aminoácidos. Las secuencias de nucleótidos siguen la convención estándar de comenzar en el extremo 5' de la secuencia y avanzar (es decir, de izquierda a derecha en cada línea) hasta el extremo 3'. Solo se muestra una cadena de cada secuencia de nucleótidos, pero la cadena complementaria se entiende incluida por cualquier referencia a la cadena mostrada. Cuando se proporciona una secuencia de nucleótidos que codifica una secuencia de aminoácidos, se entiende que también se proporcionan variantes de codón degenerado de la misma que codifican la misma secuencia de aminoácidos. Las secuencias de aminoácidos siguen la convención estándar de comenzar en el extremo amino de la secuencia y avanzar (es decir, de izquierda a derecha en cada línea) hasta el extremo carboxi.

Tabla 2. Descripción de las secuencias.

SEQ ID NO	Tipo	Descripción
1	Proteína	Dominio de repetición R1 de tau
2	Proteína	Dominio de repetición R2 de tau
3	Proteína	Dominio de repetición R3 de tau
4	Proteína	Dominio de repetición R4 de tau
5	ADN	Secuencia codificante del dominio de repetición R1 de tau

(continuación)

SEQ ID NO	Tipo	Descripción
6	ADN	Secuencia codificante del dominio de repetición R2 de tau
7	ADN	Secuencia codificante del dominio de repetición R3 de tau
8	ADN	Secuencia codificante del dominio de repetición R4 de tau
9	Proteína	Dominio de cuatro repeticiones de tau (R1-R4; aminoácidos 243-375 de tau de longitud completa (P10636-8))
10	ADN	Secuencia codificante del dominio de cuatro repeticiones de tau (R1-R4; secuencia codificante de los aminoácidos 243-375 de tau de longitud completa (P10636-8))
11	Proteína	Dominio de cuatro repeticiones de tau (R1-R4) con mutación P301S
12	ADN	Secuencia codificante del dominio de cuatro repeticiones de tau (R1-R4) con mutación P301S
13	Proteína	eCFP
14	ADN	Secuencia codificante de eCFP
15	Proteína	eYFP
16	ADN	Secuencia codificante de eYFP
17	ARN	Armazón de ARN guía V1
18	ARN	Armazón de ARN guía V2
19	ARN	Armazón de ARN guía V3
20	ARN	Armazón de ARN guía V4
21	Proteína	Cas9-FLAG
22	ADN	Secuencia codificante de Cas9-FLAG
23	ARN	Cola de ARNcr
24	ARN	ARNtracr
25	ADN	Secuencia diana de ARN guía más PAM V1
26	ADN	Secuencia diana de ARN guía más PAM V2
27	ADN	Secuencia diana de ARN guía más PAM V3
28	ARN	ARNtracr v2
29	ARN	ARNtracr v3
30	ARN	Armazón de ARN guía V5
31	ARN	Armazón de ARN guía V6
32	ARN	Armazón de ARN guía V7
33	ARN	Bucle de unión a MS2
34	ARN	Armazón de ARN guía con bucles de unión a MS2
35	ARN	ARNgu genérico con bucles de unión a MS2
36	Proteína	Proteína Cas quimérica dCas9-VP64
37	Proteína	Proteína adaptadora quimérica MCP-p65-HSF1
38	ADN	ADN que codifica la proteína Cas quimérica dCas9-VP64
39	ADN	ADN que codifica la proteína adaptadora quimérica MCP-p65-HSF1
40	Proteína	MCP
41	ADN	ADN que codifica MCP
42	ADN	Lenti dCas9-VP64
43	ADN	Lenti MCP-p65-HSF1_Hygro

Ejemplos

5

Ejemplo 1. Desarrollo de una plataforma CRISPR/Cas9 de cribado de todo el genoma para identificar modificadores genéticos de la agregación de Tau.

La agregación o fibrilación anormal de proteínas es una característica definitoria de muchas enfermedades, incluyendo notablemente varias enfermedades neurodegenerativas, tales como la enfermedad de Alzheimer (EA), la enfermedad de Parkinson (EP), la demencia frontotemporal (DFT), la esclerosis lateral amiotrófica (ELA), la encefalopatía traumática crónica (ETC) y la enfermedad de Creutzfeldt-Jakob (ECJ), entre otras. En muchas de estas enfermedades, la fibrilación de ciertas proteínas en agregados insolubles no solo es un sello distintivo de la enfermedad, sino que también se ha implicado como factor causante de neurotoxicidad. Además, estas enfermedades se caracterizan por la propagación de la patología de agregados a través del sistema nervioso central siguiendo patrones estereotípicos, un proceso que se correlaciona con la progresión de la enfermedad. La identificación de genes y vías genéticas que modifican los procesos de agregación anormal de proteínas, o la propagación de agregados de célula a célula, son, por lo tanto, de gran valor para comprender mejor la etiología de las enfermedades neurodegenerativas, así como para idear estrategias para intervención terapéutica.

20

Con el fin de identificar genes y vías que modifiquen los procesos de agregación anormal de la proteína tau, se desarrolló una plataforma para realizar cribados de todo el genoma con bibliotecas de ARNgu de nucleasas CRISPR

(CRISPRn) para identificar genes que regulan el potencial de las células para ser "sembradas" por agregados de proteínas asociadas a la enfermedad de tau (es decir, genes que, cuando se inactivan, hacen que las células sean más susceptibles a la formación de agregados de tau cuando se exponen a una fuente de proteína fibrilada tau). La identificación de dichos genes puede dilucidar los mecanismos de propagación de agregados de tau de célula a célula y las vías genéticas que rigen la susceptibilidad de las neuronas a formar agregados de tau en el contexto de las enfermedades neurodegenerativas.

El cribado empleó una línea celular humana biosensora de tau que consistía en células HEK293T que expresan de forma estable el dominio de cuatro repeticiones de tau, tau_{4RD}, que comprende el dominio de unión a microtúbulos (MBD) de tau con la mutación patógena P301S, fusionado a CFP o YFP. Es decir, las líneas celulares HEK293T contienen dos transgenes que expresan de forma estable variantes de proteínas asociadas a la enfermedad fusionadas a la proteína fluorescente CFP o a la proteína fluorescente YFP: tau^{4RD}-CFP/tau^{4RD}-YFP (TCY), en donde el dominio de repetición de tau (4RD) comprende la mutación patógena P301S. Véase la FIG. 1. En estas líneas de biosensores, la agregación de la proteína tau-CFP/tau-YFP produce una señal FRET, el resultado de una transferencia de energía fluorescente de la CFP donante a la YFP aceptora. Véase la FIG. 2. Las células FRET-positivas, que contienen agregados de tau, pueden clasificarse y aislarse por citometría de flujo. En el nivel basal, las células no estimuladas expresan los indicadores en un estado estable y soluble con una señal FRET mínima. Tras la estimulación (por ejemplo, transfección liposómica de partículas semilla), las proteínas indicadoras forman agregados, produciendo una señal FRET. Las células que contienen agregados pueden aislarse mediante FACS. Las líneas celulares que contienen agregados que se propagan de forma estable, Ag[+], pueden aislarse mediante dilución clónica en serie de líneas celulares Ag[-].

Se realizaron varias modificaciones en esta línea celular biosensora de tau para hacerla útil para el cribado genético. En primer lugar, estas células biosensoras de tau se modificaron mediante la introducción de un transgén que expresa Cas9 (SpCas9) a través de un vector lentivírico. Las líneas celulares transgénicas clónicas que expresaban Cas9 se seleccionaron con blasticidina y se aislaron mediante dilución clónica en serie para obtener clones unicelulares. Se evaluó el nivel de expresión de Cas9 en los clones mediante qRT-PCR (FIG. 3A) y la actividad de corte del ADN mediante PCR digital (FIG. 3B). Los niveles relativos de expresión de Cas9 también se muestran en la **Tabla 3**.

Tabla 3. Niveles relativos de expresión de Cas9.

Nombre del clon	Cas9D Ct				Cas9D Ct AVG	B2m Ct AVG	Cas9D-B2m delta Ct
	rep1	rep2	rep3	rep4			
3B5-B1	26,22	26,31	26,36	26,45	26,33	22,01	4,33
3B5-G2	23,68	23,85	24,39	23,61	23,88	21,51	2,38
7B5-B2	23,63	23,60	24,12	23,50	23,71	21,38	2,34
3B3-A2	24,05	23,95	24,02	24,47	24,12	21,94	2,19
7B10-C3	22,58	22,71	22,67	23,20	22,79	21,19	1,59
3B5-E1	24,12	24,32	24,75	24,05	24,31	22,81	1,50
3B5-G1	21,16	21,14	21,09	21,43	21,20	21,35	-0,15
7B5-C3	19,98	19,99	19,86	19,97	19,95	21,24	-1,29
7B5-A2	18,84	18,74	19,33	18,99	18,97	22,10	-3,12
7B5-G1	19,01	18,88	19,61	19,18	19,17	22,33	-3,16

Específicamente, se evaluó la eficiencia de mutación de Cas9 mediante PCR digital 3 y 7 días después de la transducción de lentivirus que codificaban ARNg contra dos genes diana seleccionados. La eficiencia de corte se vio limitada por los niveles de Cas9 en los clones de menor expresión. Se necesitaba un clon con un nivel adecuado de expresión de Cas9 para lograr la máxima actividad. Varios clones derivados con menor expresión de Cas9 no fueron capaces de cortar secuencias diana de forma eficiente, mientras que los clones con mayor expresión (incluidos los usados para el cribado) fueron capaces de generar mutaciones en secuencias diana en los genes *PERK* y *SNCA* con una eficiencia aproximada del 80 % tras tres días en cultivo. El corte eficiente se observó ya a los 3 días después de la transducción del ARNg, con una mejora solo marginal a los 7 días. El clon 7B10-C3 se seleccionó como un clon de alto rendimiento para su uso en posteriores cribados de bibliotecas.

En segundo lugar, se desarrollaron reactivos y un método para sensibilizar las células a la actividad de siembra de tau. La propagación de tau de célula a célula puede ser el resultado de la actividad de agregación de tau secretada por células que contienen agregados. Para estudiar la propagación celular de la agregación de tau, se obtuvieron subclones de una línea celular tau-YFP que consistía en células HEK293T que expresan de forma estable el dominio de repetición de tau, tau_{4RD}, que comprende el dominio de unión a microtúbulos (MBD) de tau con la mutación patógena P301S, fusionado a YFP. Véase la FIG. 5. Las células en las que la proteína tau-YFP se presenta de forma estable en un estado agregado (Ag[+]) se obtuvieron tratando estas células tau-YFP con tau fibrilada recombinante mezclada con reactivo de lipofectamina para sembrar la agregación de la proteína tau-YFP expresada de forma estable por estas células. A continuación, las células "sembradas" se diluyeron en serie para obtener clones unicelulares. A continuación, estos clones se expandieron para identificar líneas celulares clónicas en las que los agregados de tau-YFP persisten de forma estable en todas las células con el crecimiento y múltiples pases a lo largo del tiempo. Uno de

estos clones tau-YFP_Ag[+], Clon_18, se usó para producir medio acondicionado recogiendo el medio que había estado en células confluentes tau-YFP_Ag[+] durante cuatro días. A continuación, se aplicó medio acondicionado (MA) sobre células biosensoras sin tratamiento previo tau-CFP/Tau-YFP en una proporción de 3:1 MA:medio recién preparado para poder inducir la agregación de tau en un pequeño porcentaje de estas células receptoras. No se usó lipofectamina. No se usó lipofectamina para tener un ensayo lo más fisiológico posible, sin engañar a las células receptoras para forzar/aumentar la agregación de tau usando lipofectamina. Como se mide usando citometría de flujo para evaluar el porcentaje de células que producen una señal FRET como medida de agregación, el medio acondicionado indujo FRET de forma consistente en aproximadamente el 0,1 % de las células. Véase la FIG. 6. En conclusión, las células tau-YFP_Ag[+] no pueden producir una señal FRET, pero pueden proporcionar una fuente de semillas de tau.

Ejemplo 2. Cribado de CRISPR/Cas9 en todo el genoma para identificar modificadores genéticos de la agregación de Tau

Para revelar genes modificadores de la agregación de tau como ARNgu enriquecidos en células FRET[+], las células biosensoras tau-CFP/tau-YFP que expresan Cas9 sin agregados (Ag[-]) fueron transducidas con dos bibliotecas de ARNgu humano CRISPR de todo el genoma usando un enfoque de administración lentivírica para introducir mutaciones de inactivación en cada gen diana. Véase la FIG. 4. Cada biblioteca de ARNgu CRISPR se dirige a exones constitutivos en 5' para la inactivación funcional con una cobertura promedio de ~3 ARNgu por gen (total de 6 ARNgu por gen en las dos bibliotecas combinadas). La distribución del recuento de lecturas (es decir, la representación de cada ARNgu en la biblioteca) fue normal y similar para cada biblioteca. Los ARNgu se diseñaron para evitar efectos inespecíficos evitando ARNgu con dos o menos emparejamientos incorrectos con secuencias genómicas inespecíficas. Las bibliotecas cubren 19.050 genes humanos y 1.864 miARN con 1.000 ARNgu de control no diana. Las bibliotecas se transdujeron con una multiplicidad de infección (MOI) < 0,3 con una cobertura de > 300 células por ARNgu. Las células biosensoras de tau se cultivaron bajo selección con puromicina para seleccionar células con integración y expresión de un único ARNgu por célula. La selección con puromicina comenzó 24 horas después de la transducción a 1 µg/ml. En el cribado primario se usaron cinco repeticiones de cribado independientes.

Se recogieron muestras de toda la población celular transducida tras los pases celulares en el día 3 y día 6 después de la transducción. Tras el pase del día 6, las células se cultivaron en medio acondicionado para sensibilizarlas a la actividad de siembra. En el día 10, se usó clasificación celular asistida por fluorescencia (FACS) para aislar específicamente la subpoblación de células FRET[+]. Véase la FIG. 7. El cribado consistió en cinco experimentos repetidos. El aislamiento del ADN y la amplificación por PCR de las construcciones de ARNgu integradas permitieron una caracterización mediante secuenciación de nueva generación (NGS) del repertorio de ARNgu en cada momento.

El análisis estadístico de los datos NGS permitió la identificación de ARNgu enriquecidos en la subpoblación FRET[+] del día 10 de los cinco experimentos en comparación con el repertorio de ARNgu en los momentos anteriores día 3 y día 6. Los conceptos de abundancia relativa y enriquecimiento para el análisis NGS se ejemplifican en la FIG. 8. La primera estrategia para identificar posibles modificadores de tau fue usar la secuenciación de ADN para producir recuentos de lecturas de ARNgu en cada muestra usando el algoritmo DESeq para encontrar los ARNgu que son más abundantes en el día 10 frente al día 3 o en el día 10 frente al día 6, pero no en el día 6 frente al día 3 (cambio en veces [fc] ≥ 1,5 y prueba binomial negativa p < 0,01). Fc ≥ 1,5 significa el cociente de (promedio de los recuentos del día 10) / (promedio de los recuentos del día 3 o del día 6) ≥ 1,5. P < 0,01 significa la probabilidad de que no haya diferencia estadística entre los recuentos del día 10 y del día 3 o día 6 < 0,01. El algoritmo DESeq es un algoritmo ampliamente usado para el "análisis de expresión diferencial para datos de recuento de secuencias". Véase, por ejemplo, Anders et al. (2010) Genome Biology 11:R106.

Específicamente, se usaron dos comparaciones en cada biblioteca para identificar los ARNgu significativos: día 10 frente a día 3, y día 10 frente a día 6. Para cada una de estas cuatro comparaciones, se usó el algoritmo DESeq, y el umbral de corte para ser considerado significativo fue el cambio en veces ≥ 1,5, así como la prueba binomial negativa p < 0,01. Una vez identificadas las guías significativas en cada una de estas comparaciones para cada biblioteca, se consideró que un gen era significativo si cumplía uno de los dos criterios siguientes: (1) al menos dos ARNgu correspondientes a ese gen se consideraron significativos en una comparación (día 10 frente a día 3 o día 10 frente a día 6); y (2) al menos un ARNgu fue significativo en ambas comparaciones (día 10 frente a día 3 y día 10 frente a día 6). Usando este algoritmo, los presentes inventores identificaron cinco genes que eran significativos de la primera biblioteca y cuatro genes de la segunda biblioteca. Véase la Tabla 4.

Tabla 4. Genes identificados mediante la estrategia n.º 1.

Biblioteca n.º 1					
Día 10 frente a día 3		Día 10 frente a día 6		Día 6 frente a día 3	
Gen	ARNg significativos	Gen	ARNg significativos	Gen	ARNg significativos
Gen diana 1	1	Gen diana 1	1	Gen diana 1	0
Gen diana 2	3	Gen diana 2	1	Gen diana 2	0

(continuación)

Biblioteca n.º 1					
Día 10 frente a día 3		Día 10 frente a día 6		Día 6 frente a día 3	
Gen	ARNg significativos	Gen	ARNg significativos	Gen	ARNg significativos
Gen diana 15	1	Gen diana 15	1	Gen diana 15	0
Gen diana 16	1	Gen diana 16	1	Gen diana 16	0
Gen diana 17	2	Gen diana 17	0	Gen diana 17	0
Biblioteca n.º 2					
Día 10 frente a día 3		Día 10 frente a día 6		Día 6 frente a día 3	
Gen	ARNg significativos	Gen	ARNg significativos	Gen	ARNg significativos
Gen diana 2	1	Gen diana 2	1	Gen diana 2	0
Gen diana 18	1	Gen diana 18	1	Gen diana 18	0
Gen diana 19	1	Gen diana 19	1	Gen diana 19	0
Gen diana 20	1	Gen diana 20	1	Gen diana 20	0

5 Sin embargo, la primera estrategia requiere ciertos niveles de homogeneidad en el recuento de lecturas dentro de cada grupo experimental, lo que podría ser demasiado estricto. Para el mismo ARNg, muchos factores podrían producir variabilidad en el recuento de lecturas entre las muestras dentro de cada grupo experimental (muestras del día 3, día 6 o día 10), tal como los recuentos víricos iniciales en la biblioteca de cribado, la infección o la eficacia de la edición génica y la tasa de crecimiento relativa tras la edición génica. Así, también se usó una segunda estrategia basada en la aparición positiva (recuento de lecturas > 30) de guías por gen en cada muestra en el día 10 (post-selección) en lugar del recuento exacto de lecturas. Se calculó un valor de p estadístico formal para la observación positiva de un número de guías en la muestra post-selección (n') dado el tamaño de la biblioteca (x), el número de guías por gen (n) y el número total de guías positivas en la muestra post-selección (m) (el "número" se refiere al tipo de ARNg (es decir, secuencias de ARN guía único), no al recuento de lecturas) ($p_{n'} = nCn' * (x-n')C(m-n) / xCm$). La probabilidad de que n' guías estén presentes por casualidad es: $p_{n'} = nCn' * (x-n')C(m-n) / xCm$. La probabilidad de que n' guías o más para el gen g estén presentes por casualidad se calculó como:

$$p_g = \sum_{i=n'}^n p_i$$

20 El enriquecimiento global de los recuentos de lecturas de un gen post-selección en comparación con preselección se usó como parámetro adicional para identificar genes positivos: (abundancia relativa = [recuento de lecturas de un gen] / [recuento de lecturas de todos los genes] y enriquecimiento post-selección = [abundancia relativa post-selección] / [abundancia relativa preselección]).

25 Más específicamente, la segunda estrategia es un método de análisis nuevo y más sensible para la selección positiva CRISPR. El objetivo de la selección positiva CRISPR es usar la secuenciación del ADN para identificar los genes para los que la perturbación por ARNg está correlacionada con el fenotipo. Para reducir el ruido de fondo, en estos experimentos se suelen usar múltiples ARNg para el mismo gen junto con repeticiones del experimento. Sin embargo, actualmente los métodos de análisis estadístico comúnmente usados, que requieren un cierto grado de homogeneidad/acuerdo entre los ARNg para el mismo gen, así como entre las repeticiones técnicas, no funcionan bien. Esto se debe a que estos métodos no pueden manejar una gran variación entre ARNg y repeticiones para el mismo gen, debido a muchas razones posibles (por ejemplo, diferente infección o eficiencia de edición génica, recuentos víricos iniciales en la biblioteca de cribado y la presencia de otros ARNg con el mismo fenotipo). Por el contrario, los presentes inventores desarrollaron un método que es robusto a grandes variaciones. Se basa en las apariciones positivas de guías por gen en un experimento individual en lugar del recuento exacto de lecturas de cada ARNg. Se calculan valores de p estadísticos formales para la observación positiva de un número de ARNg a lo largo de repeticiones de experimentos dado el tamaño de la biblioteca, el número de ARNg por gen y el número total de ARNg positivos en cada experimento. También se usa como parámetro el enriquecimiento relativo de lecturas de secuencias de ARNg antes y después de la selección del fenotipo. El método de los presentes inventores funciona mejor que los métodos ampliamente usados hasta la fecha, incluyendo DESeq, MAGECK y otros. Específicamente, este método incluye las siguientes etapas:

- (1) Para cada experimento, identificar cualquier guía presente en las células con fenotipo positivo.
- (2) A nivel de genes, calcular la probabilidad aleatoria de que las guías estén presentes en cada experimento: $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de guías antes de la selección del fenotipo, m es la variedad de guías después de la selección del fenotipo, n es la variedad de guías para un gen antes de la selección del fenotipo, y n' es la variedad de guías para el gen después de la selección del fenotipo. La probabilidad global de estar presente en múltiples experimentos (generando un único valor de p con respecto a valores de p generados a partir de varios experimentos) se calcula mediante la prueba de probabilidad combinada de Fisher (referencia: Fisher, R.A.; Fisher, R. A (1948) "Questions and answers #14" The American

Statistician). Es decir, primero se calcula una estadística de prueba ϕ usando los valores de p de los múltiples experimentos:

$$\phi = -2 \sum_{k=1}^K p_k$$

5 donde p_k es el valor de p calculado para el experimento k , y K es el número total de experimentos. Entonces, el valor de p combinado con respecto a los K experimentos es igual a la probabilidad de observar el valor de ϕ bajo la distribución de la chi al cuadrado con el grado de libertad de $2 \cdot K$. Alternativamente, la posibilidad global de estar presente en múltiples experimentos se calcula multiplicando la posibilidad calculada anteriormente obtenida de cada experimento.

10 (3) Calcular el enriquecimiento promedio de las guías a nivel de genes: Puntuación de enriquecimiento = abundancia relativa post-selección/ abundancia relativa preselección. Abundancia relativa = recuento de lecturas de guías para un gen/ recuento de lecturas de todas las guías.

15 (4) Seleccionar genes significativamente por debajo de la probabilidad aleatoria de estar presentes, así como por encima de cierta puntuación de enriquecimiento.

Catorce de los genes diana identificados por los dos enfoques diferentes (cualquier enfoque o ambos) como enriquecidos en las células FRET[+] fueron seleccionados como principales candidatos para una mayor validación tras una inspección visual basada en los datos de recuento de lecturas. Véase la **Tabla 5**. Treinta ARN_g individuales fueron probados en cribados secundarios para su validación. En la **FIG. 9** se muestra un esquema de los cribados secundarios y en la **FIG. 10** se muestran los resultados. La inactivación del gen diana 2 o del gen diana 8, por múltiples ARN_g probados, aumentó la susceptibilidad de una célula para formar agregados de tau en respuesta a una fuente de actividad de siembra de tau (medio acondicionado). La inducción de la señal FRET aumentó en 15-20 veces en células con inactivación de cualquiera de estas dos dianas. La inactivación de estos dos genes diana aumentó la formación de agregados de tau en respuesta al medio acondicionado, pero no al medio recién preparado. Véase la **FIG. 11**.

20
25

Tabla 5. Dianas identificados.

Gen diana	ARNgu
Gen diana 1	g1 - Lib-A
Gen diana 1	g5 - Lib-B
Gen diana 2	g1 - Lib-A
Gen diana 2	g2 - Lib-A
Gen diana 2	g3 - Lib-A
Gen diana 2	g6 - Lib-B
Gen diana 3	g2 - Lib-A
Gen diana 3	g5 - Lib-B
Gen diana 4	g3 - Lib-A
Gen diana 4	g5 - Lib-B
Gen diana 5	g1 - Lib-A
Gen diana 5	g4 - Lib-B
Gen diana 6	g2 - Lib-A
Gen diana 6	g5 - Lib-B
Gen diana 7	g1 - Lib-A
Gen diana 7	g5 - Lib-B
Gen diana 8	g5 - Lib-B
Gen diana 8	g6 - Lib-B
Gen diana 9	g2 - Lib-A
Gen diana 9	g6 - Lib-B
Gen diana 10	g1 - Lib-A
Gen diana 10	g6 - Lib-B
Gen diana 11	g3 - Lib-A
Gen diana 11	g4 - Lib-B
Gen diana 12	g1 - Lib-A
Gen diana 12	g6 - Lib-B
Gen diana 13	g5 - Lib-B
Gen diana 13	g6 - Lib-B
Gen diana 14	g1 - Lib-A
Gen diana 14	g4 - Lib-B

5 A continuación, se realizaron otros experimentos con los genes diana 2 y 8 para validar aún más que la modificación dirigida de cada gen promueve la agregación de tau. Véase la **FIG. 12**. Se probaron dos ARNgu diferentes contra el gen diana 2 y se usó un ARNgu contra el gen diana 8. Se usó un ARNgu no diana como control negativo. Se realizaron cuatro transducciones lentivíricas independientes para cada ARN guía en el día 0. El día 6, se realizó la siembra de tau con medio acondicionado con o sin lipofectamina y se recogieron muestras para qRT-PCR. Los datos de la qRT-PCR se muestran en la **FIG. 13**. Cada uno de los dos ARNgu que se dirigen al gen diana 2 redujo la expresión del ARNm del gen diana 2, y el ARNgu que se dirige al gen diana 8 redujo la expresión del gen diana 8. El día 10 se realizó un análisis FACS para evaluar la inducción de la señal FRET. La agregación de Tau se incrementó por cada uno de los dos ARNgu que se dirigían al gen diana 2 y el ARNgu que se dirigía al gen diana 8. Véase la **FIG. 15**. En el día 13, se recogieron muestras para el análisis por transferencia Western. Los resultados de la transferencia Western se muestran en la **FIG. 14**. De forma similar a los experimentos de qRT-PCR que evalúan la expresión de ARNm, la expresión de la proteína 2 se redujo por los dos ARNgu que se dirigían al gen diana 2, y la expresión de la proteína 8 se redujo por el ARNgu que se dirigía al gen diana 8.

La validación adicional de los genes diana 2 y 8 como modificadores de la agregación de tau se realizó aislando clones atenuados individuales del gen diana 2 y clones atenuados individuales del gen diana 8 para su validación. Las células biosensoras tau-CFP/tau-YFP que expresan Cas9 sin agregados (Ag[-]) fueron transducidas con lentivirus que expresaban el ARNgu 1 del gen diana 2, el ARNgu 5 del gen diana 8, o un ARNgu no diana. A continuación, se llevó a cabo una dilución clónica en serie para seleccionar clones individuales. Los niveles de ARNm del gen diana 2 y de ARNm del gen diana 8 se evaluaron mediante qRT-PCR, y los niveles de proteína 2 y proteína 8 se evaluaron mediante transferencia Western. Cada clon de ARNgu del gen diana 2 tenía expresión de ARNm del gen diana 2 (datos no mostrados) y expresión de proteína 2 reducidas (**FIG. 16**), y cada clon de ARNgu del gen diana 8 tenía expresión de ARNm del gen diana 8 (datos no mostrados) y expresión de proteína 8 reducida (**FIG. 16**).

A continuación, cada clon se sembró con medio acondicionado durante 4 días y se realizó un análisis FRET para evaluar la agregación de tau. Los clones atenuados validan los genes diana 2 y 8 como modificadores de la agregación de tau. Véase la **FIG. 17**.

A continuación, los clones individuales se caracterizaron mediante secuenciación de nueva generación para determinar qué modificaciones se habían realizado en los loci del gen diana 2 y gen diana 8. Las modificaciones se

resumen en la **Tabla 6** a continuación. Casi todos los clones mutantes contienen algún porcentaje de alelos naturales. El porcentaje de células FRET[+] (actividad de agregación de tau) se correlacionó con el porcentaje de inserciones/deleciones causadas por la unión de extremos no homólogos en los sitios de corte (es decir, la agregación de tau se correlacionó inversamente con el porcentaje de alelos naturales -a menor porcentaje de alelos naturales, mayor porcentaje de células FRET[+]). Véanse la **FIG. 16** y la **Tabla 6**.

Tabla 6. Caracterización de clones del gen diana 2 y del gen diana 8.

Gen (diana)	Clon	Amplicón secuenciado	Frecuencia alélica (lecturas ≥ 5 %)		
			WT	INDEL 1	IINDEL 2
Gen diana 2	MP1-3	TG8_g5	98,80 %		
		TG2_g1	11,30 %	49,9 % (+1 pb)	33,9 % ($\Delta 16$ pb)
	MP1-10	TG8_g5	98,60 %		
		TG2_g1	16,50 %	79,1 % (+1 pb)	
	MP1-13	TG8_g5	98,80 %		
		TG2_g1	14,90 %	35,9 % ($\Delta 6$ pb)	44,3 % (+1 pb)
GP 1-23	TG8_g5	98,70 %			
	TG2_g1	20,20 %	71,4 % (+1 pb)		
Gen diana 8	MP5-7	TG8_g5	0,00 %	54,8 % ($\Delta 3$ pb+6pb)	29,7 % (C \rightarrow T)
		TG2_g1	99,5 %		
	MP5-9	TG8_g5	34,80 %	55,0 % ($\Delta 20$ pb)	6,8 % (+1 pb)
		TG2_g1	99,30 %		

También se evaluó la fosforilación de tau en cada clon mediante transferencia Western. Se encontró que tau estaba hiperfosforilada en S262 y S356 en el clon 5.7 de ARN_g del gen diana 8. Véase la **FIG. 18**. Aunque los clones en los que tau no estaba hiperfosforilada parecían potenciar la inducción de FRET en este experimento, el clon 5.7 era bastante inestable, y el gen diana 8 está implicado en muchos procesos biológicos, por lo que no se pudieron extraer conclusiones generales. Otros experimentos en neuronas corticales primarias de ratón mostraron que la tinción fosfo-tau (S356) aumenta en el compartimento somatodendrítico en neuronas corticales primarias de ratón tratadas con ARN_g de Cas9 y del gen diana 2 mediante administración lentivírica y se mantuvo durante 14 días en cultivo (datos no mostrados).

Esta validación confirmó el valor del enfoque de cribado primario en la identificación de genes que pueden regular la susceptibilidad de las células a la siembra de tau cuando se exponen a una fuente externa de actividad de siembra de tau. Las dianas identificadas mediante el cribado podrían ser, por tanto, dianas relevantes en la propagación célula a célula de la patología de tau en el contexto de la enfermedad neurodegenerativa, y se seguirán explorando.

Ejemplo 3. Desarrollo de una plataforma de cribado CRISPR/Cas9 de todo el genoma para identificar modificadores genéticos de la agregación de tau usando una biblioteca CRISPR/Cas9 de activación transcripcional.

Para identificar aún más los genes y las vías que modifican los procesos de agregación anormal de la proteína tau, se desarrolló una plataforma para realizar cribados de todo el genoma con activación transcripcional (hSAM) con bibliotecas de ARN_g CRISPRa para identificar genes que regulan el potencial de las células para ser "sembradas" por agregados de proteína asociados a la enfermedad de tau (es decir, genes que, cuando se activan transcripcionalmente, hacen que las células sean más susceptibles a la formación de agregados de tau cuando se exponen a una fuente de proteína fibrilada tau). La identificación de dichos genes puede dilucidar los mecanismos de propagación de agregados de tau de célula a célula y las vías genéticas que rigen la susceptibilidad de las neuronas a formar agregados de tau en el contexto de las enfermedades neurodegenerativas.

El cribado empleó una línea celular humana biosensora de tau que consistía en células HEK293T que expresan de forma estable el dominio de cuatro repeticiones de tau, tau_{4RD}, que comprende el dominio de unión a microtúbulos (MBD) de tau con la mutación patógena P301S, fusionado a CFP o YFP. Es decir, las líneas celulares HEK293T contienen dos transgenes que expresan de forma estable variantes de proteínas asociadas a la enfermedad fusionadas a la proteína fluorescente CFP o a la proteína fluorescente YFP: tau^{4RD}-CFP/tau^{4RD}-YFP (TCY), en donde el dominio de repetición de tau (4RD) comprende la mutación patógena P301S. Véase la **FIG. 1**. En estas líneas de biosensores, la agregación de la proteína tau-CFP/tau-YFP produce una señal FRET, el resultado de una transferencia de energía fluorescente de la CFP donante a la YFP aceptora. Véase la **FIG. 2**. Las células FRET-positivas, que contienen agregados de tau, pueden clasificarse y aislarse por citometría de flujo. En el nivel basal, las células no estimuladas expresan los indicadores en un estado estable y soluble con una señal FRET mínima. Tras la estimulación (por ejemplo, transfección liposómica de partículas semilla), las proteínas indicadoras forman agregados, produciendo una señal FRET. Las células que contienen agregados pueden aislarse mediante FACS. Las líneas celulares con agregados que se propagan de forma estable, Ag[+], pueden aislarse mediante dilución clónica en serie de líneas celulares Ag[-].

Se realizaron varias modificaciones en esta línea celular biosensora de tau para hacerla útil para el cribado genético.

En primer lugar, esta línea celular biosensora se modificó transgénicamente para expresar los componentes del sistema de activación transcripcional CRISPR/Cas SAM: dCas9-VP64 y MS2-P65-HSF1. Las construcciones lentivíricas dCas9-VP64 y MS2-P65-HSF1 se proporcionan en SEQ ID NOS: 42 y 43, respectivamente. Se seleccionó un clon como clon de alto rendimiento para su uso en posteriores cribados de bibliotecas. Se validó la eficacia de este clon en la activación de genes diana seleccionados.

En segundo lugar, se desarrollaron reactivos y un método para sensibilizar las células a la actividad de siembra de tau como en el Ejemplo 1. Alternativamente, si se necesita más agregación, puede usarse lisado celular de un clon Ag[+] como en el Ejemplo 5.

Ejemplo 4. Cribado CRISPR/Cas9 de todo el genoma para identificar modificadores genéticos de la agregación de tau usando una biblioteca CRISPR/Cas9 de activación transcripcional.

Para revelar más genes modificadores de la agregación de tau como ARN_gu enriquecidos en células FRET[+], las células biosensoras tau-CFP/tau-YFP que expresan SAM sin agregados (Ag[-]) fueron transducidas con una biblioteca de ARN_gu humano hSAM CRISPR de todo el genoma usando un enfoque de administración lentivírica para activar transcripcionalmente cada gen diana. Los ARN_gu de la biblioteca se dirigen a sitios dentro de 200 pb aguas arriba del sitio de inicio de la transcripción con una cobertura promedio de ~3 ARN_gu por gen. Los ARN_gu se diseñaron para evitar efectos inespecíficos evitando los ARN_gu con dos o menos emparejamientos incorrectos con secuencias genómicas inespecíficas. La biblioteca cubre 18.946 genes humanos. La biblioteca se transdujo con una multiplicidad de infección (MOI) < 0,3 con una cobertura de > 300 células por ARN_gu. Las células biosensoras de tau se cultivaron bajo selección con zeocina para seleccionar células con integración y expresión de un único ARN_gu por célula. En el cribado primario se usaron cinco repeticiones de cribado independientes.

Se recogieron muestras de toda la población celular transducida tras los pases celulares en el día 3 y día 6 después de la transducción. Tras el pase del día 6, las células se cultivaron en medio acondicionado para sensibilizarlas a la actividad de siembra. En el día 10, se usó clasificación celular asistida por fluorescencia (FACS) para aislar específicamente la subpoblación de células FRET[+]. El aislamiento del ADN y la amplificación por PCR de las construcciones de ARN_gu integradas permitieron una caracterización mediante secuenciación de nueva generación (NGS) del repertorio de ARN_gu en cada momento.

Dado que el análisis de datos y el análisis estadístico reflejaron el enfoque usado en el Ejemplo 2, no se repiten aquí todos los detalles del Ejemplo 2. El análisis estadístico de los datos NGS permitió la identificación de ARN_gu enriquecidos en la subpoblación FRET[+] del día 10 de los experimentos múltiples en comparación con el repertorio de ARN_gu en los momentos anteriores día 3 y día 6. La primera estrategia para identificar posibles modificadores de tau fue usar la secuenciación de ADN para producir recuentos de lecturas de ARN_gu en cada muestra usando el algoritmo DESeq para encontrar los ARN_gu que son más abundantes en el día 10 frente al día 3 o en el día 10 frente al día 6, pero no en el día 6 frente al día 3 (cambio en veces (fc) $\geq 1,5$ y prueba binomial negativa $p < 0,01$). $F_c \geq 1,5$ significa el cociente de (promedio de los recuentos del día 10) / (promedio de los recuentos del día 3 o del día 6) $\geq 1,5$. $P < 0,01$ significa la probabilidad de que no haya diferencia estadística entre los recuentos del día 10 y del día 3 o 6 $< 0,01$. El algoritmo DESeq es un algoritmo ampliamente usado para el "análisis de expresión diferencial para datos de recuento de secuencias". Véase, por ejemplo, Anders et al. (2010) *Genome Biology* 11:R106.

Específicamente, se usaron dos comparaciones en cada biblioteca para identificar los ARN_gu significativos: día 10 frente a día 3, y día 10 frente a día 6. Para cada una de estas cuatro comparaciones, se usó el algoritmo DESeq, y el umbral de corte para ser considerado significativo fue cambio en veces $\geq 1,5$, así como la prueba binomial negativa $p < 0,01$. Una vez identificadas las guías significativas en cada una de estas comparaciones para cada biblioteca, se consideró que un gen era significativo si cumplía uno de los dos criterios siguientes: (1) al menos dos ARN_gu correspondientes a ese gen se consideraron significativos en una comparación (día 10 frente a día 3 o día 10 frente a día 6); y (2) al menos un ARN_gu fue significativo en ambas comparaciones (día 10 frente a día 3 y día 10 frente a día 6).

Sin embargo, la primera estrategia requiere ciertos niveles de homogeneidad en el recuento de lecturas dentro de cada grupo experimental, lo que podría ser demasiado estricto. Para el mismo ARN_gu, muchos factores podrían producir variabilidad en el recuento de lecturas entre las muestras dentro de cada grupo experimental (muestras del día 3, día 6 o día 10), tal como los recuentos víricos iniciales en la biblioteca de cribado, la infección o la eficacia de la edición génica y la tasa de crecimiento relativa tras la edición génica. Así, también se usó una segunda estrategia basada en la aparición positiva (recuento de lecturas > 30) de guías por gen en cada muestra en el día 10 (post-selección) en lugar del recuento exacto de lecturas. Se calculó un valor de p estadístico formal para la observación positiva de un número de guías en la muestra post-selección (n') dado el tamaño de la biblioteca (x), el número de guías por gen (n) y el número total de guías positivas en la muestra post-selección (m) (el "número" se refiere al tipo de ARN_gu (es decir, secuencias únicas de ARN guía), no al recuento de lecturas) ($p_{n'} = nCn' * (x-n')C(m-n) / xCm$). La probabilidad de que n' guías estén presentes por casualidad es: $p_{n'} = nCn' * (x-n')C(m-n) / xCm$. La probabilidad de que n' guías o más para el gen g estén presentes por casualidad se calculó como:

$$p_g = \sum_{i=n'}^n p_i$$

El enriquecimiento global de los recuentos de lecturas de un gen post-selección en comparación con preselección se usó como parámetro adicional para identificar genes positivos: (Abundancia relativa = [recuento de lecturas de un gen] / [recuento de lecturas de todos los genes] y enriquecimiento post-selección = [abundancia relativa después de la selección] / [abundancia relativa preselección]).

Más específicamente, la segunda estrategia es un método de análisis nuevo y más sensible para la selección positiva CRISPR. El objetivo de la selección positiva CRISPR es usar la secuenciación del ADN para identificar los genes para los que la activación transcripcional por ARN_g está correlacionada con el fenotipo. Para reducir el ruido de fondo, en estos experimentos se suelen usar múltiples ARN_g para el mismo gen junto con repeticiones del experimento. Sin embargo, actualmente los métodos de análisis estadístico comúnmente usados, que requieren un cierto grado de homogeneidad/acuerdo entre los ARN_g para el mismo gen, así como entre las repeticiones técnicas, no funcionan bien. Esto se debe a que estos métodos no pueden manejar una gran variación entre ARN_g y repeticiones para el mismo gen, debido a muchas razones posibles (por ejemplo, diferente infección o eficiencia de edición génica, recuentos víricos iniciales en la biblioteca de cribado y la presencia de otros ARN_g con el mismo fenotipo). Por el contrario, los presentes inventores desarrollaron un método que es robusto a grandes variaciones. Se basa en las apariciones positivas de guías por gen en un experimento individual en lugar del recuento exacto de lecturas de cada ARN_g. Se calculan valores de p estadísticos formales para la observación positiva de un número de ARN_g a lo largo de repeticiones de experimentos dado el tamaño de la biblioteca, el número de ARN_g por gen y el número total de ARN_g positivos en cada experimento. También se usa como parámetro el enriquecimiento relativo de lecturas de secuencias de ARN_g antes y después de la selección del fenotipo. El método de los presentes inventores funciona mejor que los métodos ampliamente usados hasta la fecha, incluyendo DESeq, MAGECK y otros. Específicamente, este método incluye las siguientes etapas:

- (1) Para cada experimento, identificar las guías presentes en las células con fenotipo positivo.
- (2) A nivel de genes, calcular la probabilidad aleatoria de que las guías estén presentes en cada experimento: $nCn' * (x-n')C(m-n) / xCm$, donde x es la variedad de guías antes de la selección del fenotipo, m es la variedad de guías después de la selección del fenotipo, n es la variedad de guías para un gen antes de la selección del fenotipo, y n' es la variedad de guías para el gen después de la selección del fenotipo. La probabilidad global de estar presente en múltiples experimentos (generando un único valor de p con respecto a valores de p generados a partir de varios experimentos) se calcula mediante la prueba de probabilidad combinada de Fisher (referencia: Fisher, R.A.; Fisher, R. A (1948) "Questions and answers #14" The American Statistician). Es decir, primero se calcula una estadística de prueba ϕ usando los valores de p de los múltiples experimentos:

$$\phi = -2 \sum_{k=1}^K p_k$$

donde p_k es el valor de p calculado para el experimento k, y K es el número total de experimentos. Entonces, el valor de p combinado con respecto a los K experimentos es igual a la probabilidad de observar el valor de ϕ bajo la distribución de la chi al cuadrado con el grado de libertad de 2*K. Alternativamente, la posibilidad global de estar presente en múltiples experimentos se calcula multiplicando la posibilidad calculada anteriormente obtenida de cada experimento.

(3) Calcular el enriquecimiento promedio de las guías a nivel de genes: Puntuación de enriquecimiento = abundancia relativa post-selección/ abundancia relativa preselección. Abundancia relativa = recuento de lecturas de guías para un gen/ recuento de lecturas de todas las guías.

(4) Seleccionar genes significativamente por debajo de la probabilidad aleatoria de estar presentes, así como por encima de cierta puntuación de enriquecimiento.

Un total de 34 genes diana (dirigidos por 42 ARN_g diferentes) fueron identificados por los dos enfoques diferentes (cualquier enfoque o ambos) como enriquecidos en las células FRET[+].

Ejemplo 5. Preparación y validación de lisados celulares de tau-YFP Clon 18 Ag[+] para siembra de tau

A continuación, los presentes inventores querían identificar genes diana que, cuando se inactivan, reducen la agregación de tau (ya sea bloqueando la captación de semillas, inhibiendo la formación de oligómeros o fibrillas, promoviendo su desmontaje, o por algún otro mecanismo). Para desarrollar dicho cribado para ARN_g que impiden la agregación de tau, los presentes inventores necesitaban una fuente de actividad de siembra que fuera potente y que pudiera ser fácilmente generada en grandes cantidades. Como se describe en este ejemplo, los presentes inventores descubrieron que el lisado de células enteras de tau-YFP Ag[+] clon18 puede inducir la agregación de tau y la señal FRET, mientras que el lisado de tau-YFP Ag[-] no lo hace. El lisado de células enteras es mucho más potente para sembrar la actividad de tau (sin lipofectamina) que los medios acondicionados o las fibrillas de tau recombinante. Para inducir la agregación (FRET) en un mayor porcentaje de células (por ejemplo, > 50 %), los presentes inventores intentaron usar lisado de células enteras en combinación con lipofectamina. También necesitaban desarrollar un

método para recoger el lisado celular en un tampón que no fuera tóxico para las células. En los experimentos descritos en el presente documento, los presentes inventores usaron solución salina tamponada con fosfato (PBS) + inhibidores de la proteasa para recoger las células y sonicación para lisar las células. El lisado celular sonicado durante 3 minutos funcionó mejor.

5 Las células se prepararon para la sonicación de la siguiente manera: (1) tomar 6x placas T175 de células tau-YFP Ag[-] y 6x placas T175 de células tau-YFP Clon 18 Ag[+] (~50 millones de células/T175); (2) enjuagar el matraz con 10 ml de PBS, luego raspar las células en 5 ml de PBS; (3) recoger 2xT175 en un tubo de 15 ml, luego enjuagar estos 2xT175 con un total de 5 ml de PBS y añadirlo al tubo de 15 ml para un volumen final de 15 ml (debería haber 3 tubos de células tau-YFP Ag[-] y 3 tubos de células tau-YFP Clon 18 Ag[+]); (4) centrifugar los tubos a 1000 rpm durante 5 minutos, aspirar, añadir 4 ml de PBS/tubo + 40 µl de HALT™ (seis inhibidores de la proteasa de amplio espectro AEBSF, aprotinina, bestatina, E-64, leupeptina y pepstatina A estabilizados en sulfóxido de dimetilo (DMSO) de alta calidad) + 40 µl de ácido etilendiaminotetracético (EDTA)/tubo (inhibidores de la proteasa); y (5) transferir las suspensiones celulares a tubos de 50 ml (el soporte para tubos de sonicación solo puede contener tubos de 50 ml) y congelar a -80 °C hasta la sonicación.

La sonicación para crear lisados celulares se realizó de la siguiente manera: (1) descongelar las suspensiones celulares en un baño de agua caliente; (2) sonicar durante 1 minuto, 3 minutos o 6 minutos; (3) centrifugar las muestras sonicadas a 1000 rpm durante 5 minutos, tomar alícuotas de 300 µl y congelar a -80 °C.

20 Las células biosensoras tau-CFP/tau-YFP/Cas9 Clon 7B10C3 Ag[-] (C3) y las células de control FRET/FACS tau-CFP, tau-YFP, HEK-HZ, tau-CFP/tau-YFP/Cas9 Clon 7B10C3-B2 Ag[+] (B2) se descongelaron y se sembraron en placas de 12 pocillos y se trataron con 10 µg, 30 µg o 100 µg del lisado celular producido anteriormente. Los resultados se muestran en la FIG. 19. Se observó una respuesta dependiente de la dosis, y la sonicación de 3 minutos funcionó mejor.

En un experimento posterior, se sembraron 150.000 células 7B10C3 Ag[-] por pocillo en placas de 24 pocillos por duplicado. Se usó 1 ml de medio recién preparado por pocillo. Se añadió 0 (control), 10, 25 o 50 µg de lisado celular del clon 18 tau-YFP Ag[+] de 3 min +/- lipofectamina (10 µl por 1 ml de medio). Con la lipofectamina, todas las células eran Ag[+], pero todas las células estaban muertas después de 24 horas en las muestras de lisado de 25 y 50 µg y después de 48 horas en las muestras de lisado de 10 µg.

35 A continuación, los presentes inventores probaron a disminuir la cantidad de lisado usado. Se sembraron 150.000 células 7B10C3 Ag[-] por pocillo en placas de 24 pocillos por duplicado. Se usó 1 ml de medio recién preparado por pocillo. Se añadió 0 (control), 1, 5 o 10 µg de lisado celular del clon 18 tau-YFP Ag[+] de 3 min +/- lipofectamina (10 µl por 1 ml de medio). Con la lipofectamina, todas las células eran Ag[+], pero las células parecían poco saludables después de 24 horas y estaban muertas después de 48 horas en las muestras de lisado de 5 y 10 µg. En las muestras de lisado de 1 µg, todas las células eran Ag[+], pero tenían un aspecto poco saludable a las 24 horas. A los 3 días, las células seguían vivas y Ag[+].

40 A continuación, los presentes inventores probaron a disminuir la cantidad de lisado usado y a probar diferentes concentraciones de lipofectamina. Se sembraron 150.000 células 7B10C3 Ag[-] por pocillo en placas de 24 pocillos por duplicado. Se usó 1 ml de medio recién preparado por pocillo. Se añadió 0 (control), 0,1, 0,5 o 1 µg de lisado celular del clon 18 tau-YFP Ag[+] de 3 min +/- lipofectamina (1, 5 o 10 µl por 1 ml de medio). Con 10 µl de lipofectamina, todas las células estaban muertas después de 24 horas en todas las muestras. Con 5 µl de lipofectamina y 0,5 µg de lisado, las células tenían buen aspecto a las 48 horas y la mayoría eran Ag[+]. Similarmente, con 5 µl de lipofectamina y 1 µg de lisado, las células tenían buen aspecto a las 48 horas y todas eran Ag[+]. Sin embargo, la lipofectamina de 5 µl era tóxica en las muestras de 0 µg y 0,1 µg, ya que las células tenían un aspecto poco saludable después de 48 horas. Con 1 µl de lipofectamina, las células se veían bien después de 48 horas y eran -30 % Ag[+] en las muestras de lisado de 0,1 µg, eran -40 % Ag[+] en las muestras de lisado de 0,5 µg y eran -50 % Ag[+] en las muestras de lisado de 1 µg. Véase la FIG. 20. Se realizó un experimento similar con 3 µg, 5 µg o 10 µg de lisado celular y 1 µl, 2 µl o 3 µl de lipofectamina. Los resultados se muestran en la FIG. 21. Se realizó otro experimento con 0,5 µg, 0,7 µg o 1 µg de lisado celular y 3,5 µl o 4 µl de lipofectamina. El porcentaje de células positivas a la agregación para cada condición fue el siguiente: 0,5 µg de lisado/3,5 µl de lipofectamina: 73,5 %; 0,7 µg de lisado/3,5 µl de lipofectamina: 71,7 %; 1 µg de lisado/3,5 µl de lipofectamina: 75,7 %; 0,5 µg de lisado/4 µl de lipofectamina: 76,4 %; 0,7 µg de lisado/4 µl de lipofectamina: 76,7 %; y 1 µg de lisado/4 µl de lipofectamina: 78,0 %. En experimentos posteriores, 2 µl y 2,5 µl de lipofectamina no fueron tóxicos para las células, pero 3 µl, 3,5 µl, 4 µl, 4,5 µl y 5 µl sí lo fueron.

60 A continuación, se amplió a matraces T25, sembrando 2,8 millones de células 7B10C3 + lisado + lipofectamina. La clasificación FACS/FRET se realizó tras 2 días de cultivo. Se probaron dos condiciones: (1) 1,5 µg de lisado + 2 µl de lipofectamina por ml de medio recién preparado; y (2) 2 µg de lisado + 2 µl de lipofectamina por ml de medio recién preparado. En el primer experimento, 965.880 células eran FRET[-] (es decir, Ag[-]) y 984.760 células eran FRET[+] (es decir, Ag[+]). En el segundo experimento, 547.960 células eran FRET[-] (es decir, Ag[-]) y 855.900 células eran FRET[+] (es decir, Ag[+]). Los presentes inventores a la conclusión de que para los cribados de todo el genoma usarían 2,5 µl de lipofectamina por ml de medio recién preparado y compararían dos dosis diferentes de lisados celulares del clon 18 tau-YFP Ag[+] (1,5 µg y 2 µg), lo que potencialmente podría aumentar el porcentaje de células Ag[+]

manteniéndolas sanas.

Ejemplo 6. Cribado CRISPR/Cas9 de todo el genoma para identificar modificadores genéticos que impiden la agregación de tau.

El objetivo del cribado del Ejemplo 2 era identificar genes modificadores que, cuando se inactivaran, promovieran la formación de agregados de tau cuando se estimulan con un material de siembra de tau débil. Por el contrario, en el cribado descrito en este ejemplo, los presentes inventores trataron sus células biosensoras con un material de siembra de tau potente que normalmente provoca la agregación de tau y la inducción de FRET en la mayoría de las células. Este material de "siembra máxima" consiste en lisado celular completo sonificado de células del clon 18 tau-YFP Ag[+], aplicado con el reactivo de transfección lipofectamina. Los presentes inventores tomaron este enfoque para identificar genes diana que, cuando se inactivaran, redujeran la agregación de tau (ya sea bloqueando la captación de semillas, inhibiendo la formación de oligómeros o fibrillas, promoviendo su desensamblaje, o por algún otro mecanismo). Véase la FIG. 22.

Las células biosensoras tau-CFP/tau-YFP que expresan Cas9 sin agregados (Ag[-]) fueron transducidas con dos bibliotecas de ARN_g humano CRISPR de todo el genoma usando un enfoque de administración lentivírica para introducir mutaciones de inactivación en cada gen diana. Cada biblioteca de ARN_g CRISPR se dirige a exones constitutivos en 5' para la inactivación funcional con una cobertura promedio de ~3 ARN_g por gen (total de 6 ARN_g por gen en las dos bibliotecas combinadas). La distribución del recuento de lecturas (es decir, la representación de cada ARN_g en la biblioteca) fue normal y similar para cada biblioteca. Los ARN_g se diseñaron para evitar efectos inespecíficos evitando ARN_g con dos o menos emparejamientos incorrectos con secuencias genómicas inespecíficas. Las bibliotecas cubren 19.050 genes humanos y 1.864 miARN con 1.000 ARN_g de control no diana. Las bibliotecas se transdujeron con una multiplicidad de infección (MOI) < 0,3 con una cobertura de > 300 células por ARN_g. Las células biosensoras de tau se cultivaron bajo selección con puromicina para seleccionar células con integración y expresión de un único ARN_g por célula. La selección con puromicina comenzó 24 horas después de la transducción a 1 µg/ml. En el cribado primario se usaron cinco repeticiones de cribado independientes.

En cada duplicado, se recogieron muestras de células biosensoras 7B10C3 el día 3 y el día 7 después de la transducción de las bibliotecas CRISPR encapsidadas lentivíricamente. En el pase del día 7, se añadió el material de "siembra máxima" a las células, y después de 48 horas (en el día 9), se realizó FACS para separar y recoger las poblaciones FRET[-] y FRET[+]. El cribado consistió en cinco experimentos repetidos. El aislamiento del ADN y la amplificación por PCR de las construcciones de ARN_g integradas permitieron una caracterización por secuenciación de nueva generación (NGS) del repertorio de ARN_g en cada momento. Hubo 40 muestras en total: 5 cribados repetidos * 2 bibliotecas * 4 muestras para cada una (día 3, día 7, día 9 FRET[-] y día 9 FRET[+]). Para la "siembra máxima", se usaron 2,5 µl de lipofectamina por ml de medio recién preparado, y se probaron diferentes cantidades de lisado celular por ml de medio recién preparado (2 µg, 4 µg y 5 µg).

El análisis de datos y el análisis estadístico reflejaron el enfoque usado en el Ejemplo 2, por lo que los detalles del Ejemplo 2 no se repiten aquí. Los presentes inventores compararon (análisis emparejado) FRET(-) frente a FRET(+) frente al Día 7. También realizaron un análisis emparejado por comparación con el día 3 (más estricto). Las muestras FRET[-] y FRET[+] del día 9 se analizaron en busca de ARN_g cuyo enriquecimiento o agotamiento pudiera indicar un efecto sobre el fenotipo de agregación de tau. Los ARN_g que están enriquecidos en células FRET[-] del día 9 en relación con FRET[+] del día 9, día 3 y día 7, y/o están agotados en células FRET[+] del día 9 en relación con FRET[-] del día 9, día 3 y día 7, indican genes diana que, cuando se inactivan, pueden reducir o proteger contra la agregación de tau. Los ARN_g que están enriquecidos en las células FRET[+] del día 9 en relación con las células FRET[-] del día 9, del día 3 y del día 7, y/o están agotados en las células FRET[-] del día 9 en relación con las células FRET[+] del día 9, del día 3 y del día 7, indican genes diana que, cuando se inactivan, pueden promover o aumentar la agregación de tau. Los ARN_g que se agotan tanto en el día 9 FRET[+], tal como en el día 9 FRET[-] en relación con los momentos anteriores probablemente representan genes esenciales que reducen la viabilidad celular con el tiempo cuando se inactivan.

Los presentes inventores identificaron 142 ARN_g significativos que estaban enriquecidos o agotados en las células FRET[-] (en comparación con las células FRET[+] y las células del día 7). De estos, 46 ARN_g se agotaron en las células FRET[-], 77 ARN_g se enriquecieron en las células FRET[-] y 20 ARN_g se enriquecieron en las células FRET[-] en comparación con el día 7 (no significativo en comparación con el día 3). Véanse las FIG. 23 y 24.

A continuación, se probaron 405 ARN_g individuales en cribados secundarios para su validación. En la FIG. 25 se muestra un esquema de los cribados secundarios. Se realizaron cuatro experimentos. La cantidad de lisado celular usada en cada uno fue de 5 µg/ml de medio recién preparado. Se probaron cuatro cantidades diferentes de lipofectamina (por ml de medio recién preparado): 1,5 µl, 2 µl, 2,5 µl y 3,5 µl. Esta validación confirma el valor del enfoque de cribado primario en la identificación de genes que pueden actuar como modificadores positivos y negativos de la agregación de tau.

Ejemplo 7. Cribado CRISPR/Cas9 en todo el genoma para identificar modificadores genéticos que impiden la agregación de tau usando una biblioteca CRISPR/Cas9 de activación transcripcional.

De forma similar al Ejemplo 6, en el cribado descrito en este ejemplo, los presentes inventores trataron sus células biosensoras con un potente material de siembra de tau que normalmente causa agregación de tau e inducción de FRET en una mayoría de células. Este material de "siembra máxima" consiste en lisado celular completo sonificado de células del clon 18 tau-YFP Ag[+], aplicado con el reactivo de transfección lipofectamina. Adoptaron este enfoque para identificar genes diana que, cuando se activan transcripcionalmente, reducen la agregación de tau (ya sea bloqueando la captación de semillas, inhibiendo la formación de oligómeros o fibrillas, promoviendo su desensamblaje o mediante algún otro mecanismo). Véase la FIG. 26.

Las células biosensoras tau-CFP/tau-YFP que expresan SAM sin agregados (Ag[-]) fueron transducidas con una biblioteca de ARN_g humano CRISPR hSAM de todo el genoma usando un enfoque de administración lentivírica para activar transcripcionalmente cada gen diana. Los ARN_g de la biblioteca se dirigen a sitios dentro de 200 pb aguas arriba del sitio de inicio de la transcripción con una cobertura promedio de ~3 ARN_g por gen. Los ARN_g se diseñaron para evitar efectos inespecíficos evitando los ARN_g con dos o menos emparejamientos incorrectos con secuencias genómicas inespecíficas. La biblioteca cubre 18.946 genes humanos. La biblioteca se transdujo con una multiplicidad de infección (MOI) < 0,3 con una cobertura de > 300 células por ARN_g. Las células biosensoras de tau se cultivaron bajo selección con zeocina para seleccionar células con integración y expresión de un único ARN_g por célula. En el cribado primario se usaron cinco repeticiones de cribado independientes.

En el día 10 (PM) posterior a la transducción de las bibliotecas CRISPR encapsidadas lentivíricamente, se añadió el material de "siembra máxima" a las células, y en el día 13 (AM) se realizó FACS para separar y recoger las poblaciones FRET[-] y FRET[+]. Véase la FIG. 26. El cribado consistió en cinco experimentos duplicados. El aislamiento del ADN y la amplificación por PCR de las construcciones de ARN_g integradas permitieron una caracterización por secuenciación de nueva generación (NGS) del repertorio de ARN_g en cada momento. Para la "siembra máxima", se usaron tres cantidades de lipofectamina por ml de medio recién preparado (2,5 µl, 3,5 µl y 4 µl), y se probaron diferentes cantidades de lisado celular por ml de medio recién preparado (3 µg, 4 µg y 5 µg).

Las muestras FRET[-] y FRET[+] del día 13 se analizaron en busca de ARN_g cuyo enriquecimiento o agotamiento pudiera indicar un efecto sobre el fenotipo de agregación de tau. Los ARN_g que están enriquecidos en las células FRET[-] del día 13 en relación con las FRET[+] del día 13 y el día 10, y/o están agotados en las células FRET[+] del día 13 en relación con las FRET[-] del día 13 y el día 10, indican genes diana que, cuando se activan transcripcionalmente, pueden reducir o proteger contra la agregación de tau. Los ARN_g que están enriquecidos en las células FRET[+] del día 13 en relación con las células FRET[-] del día 13 y el día 10, y/o están agotados en las células FRET[-] del día 13 en relación con las células FRET[+] del día 13 y el día 10, indican genes diana que, cuando se activan transcripcionalmente, pueden promover o aumentar la agregación de tau. Los datos se analizan y prueban en cribados secundarios para validación como en el Ejemplo 6. El análisis estadístico reproduce el enfoque usado en el Ejemplo 2. Comparamos (análisis pareado) FRET(-) frente a FRET(+) frente día 10.

Ejemplo 8. Cribado CRISPR/Cas9 de todo el genoma para identificar modificadores genéticos de la desagregación de tau

En los ejemplos anteriores, los presentes inventores han llevado a cabo dos series de cribados de todo el genoma con el objetivo de identificar reguladores positivos y negativos de la agregación de tau. En su primer cribado, los presentes inventores han identificado genes modificadores que, cuando se inactivan, promueven la formación de agregados de tau cuando se estimulan con un material de siembra de tau débil. Este material de "siembra mínima" consiste en un medio acondicionado recogido de células tau-YFP agregadas[+], aplicado directamente a células 7B10C3 para desencadenar células -0,1 % de FRET[+] después de 3 días in vitro (DIV). A diferencia, en el segundo cribado han tratado las células biosensoras con un material de siembra de tau potente o máximo, que normalmente causa agregación de tau e inducción de FRET en la mayoría de células. Este material de "siembra máxima" consiste en lisado celular completo sonificado de células tau-YFP agregadas[+], aplicado con el reactivo de transfección lipofectamina.

Aquí, transdujeron células 7B10C3-B2 o DC11-B6 agregadas[+] con bibliotecas CRISPR encapsidadas lentivíricamente. 7B10C3-B2 y DC11-B6 son clones que contienen agregados de tau que se propagan de forma estable. Este cribado se realizó en células biosensoras de tau HEK293 que expresan Cas9 (clon 7B10C3) o dCas9-SAM (clon DC11), en las que la agregación de tau produce una señal FRET que puede detectarse y usarse como medio de clasificación celular por FACS. Los clones 7B10C3 y DC11 se trataron posteriormente con fibrillas de tau para derivar subclones que contuvieran agregados de tau que se propagaran de forma estable. Dos de estos subclones estables con agregados positivos agregados[+], denominados 7B10C3-B2 y DC11-B6, se seleccionaron para su expansión y se usaron para el cribado.

Después de dos semanas en cultivo, aislaron y secuenciaron por secuenciación de nueva generación (NGS) para revelar ARN_g únicos (ARN_g) agotados/enriquecidos en células FRET[-] con la hipótesis de que los ARN_g que causan una pérdida de agregación de tau a través de la inactivación de sus dianas específicas se enriquecerán en la población FRET[-]. La población de células FRET[-] estaba compuesta predominantemente por células sin agregados. Sin embargo, observaron algunas células FRET[-] que mostraban motitas, compuestas por minúsculos agregados que

no eran suficientes para emitir FRET y por lo tanto no eran reconocidas como FRET[+] por FACS.

5 Para reducir los falsos positivos en la población FRET[-] y minimizar el número de células motitas[+] observadas predominantemente en la fase G1 después de la mitosis (véase la **FIG. 28**), sincronizaron la progresión del ciclo celular mediante el bloqueo doble con timidina, un inhibidor de la síntesis de ADN (véase la **FIG. 29**). Esta novedosa aplicación de este método de sincronización les permitió obtener una población celular predominantemente enriquecida en fase S y sincronizar así la acumulación de agregados tras la mitosis.

10 Se realizan cinco cribados repetidos para cada biblioteca (dos bibliotecas CRISPRn y una biblioteca CRISPRa). En cada repetición, se recogen muestras de células biosensoras 7B10C3-B2 y DC11-B6 el día 7 y el día 10 después de la transducción de las bibliotecas CRISPR encapsidadas lentivíricamente. En el día 12, las células se cultivan en presencia de timidina durante 21 horas. Tras eliminar la timidina del primer bloque (día 13), las células se lavan y se cultivan en medio recién preparado durante 8 horas para liberarlas del bloque y se incuban de nuevo con timidina para el segundo bloque. Como resultado de esta sincronización, las células progresan de forma sincronizada a través de la fase G2 y mitótica y se detienen al comienzo de la fase S. En el día 14, las células se liberan mediante el segundo bloque de timidina y se cultivan en medio recién preparado durante 3 horas para que progresen de forma sincrónica a través de la fase G2. Se realiza FACS para separar y recoger las poblaciones FRET[-] y FRET[+]. Véase la **FIG. 27**.

20 Se recoge ADN genómico de cada muestra y se amplifica el repertorio de ARN_g de cada una mediante PCR. Hay 60 muestras en total: 5 cribados replicados * 3 bibliotecas * 4 muestras para cada una (día 7, día 10, día 14 FRET[-], y día 14 FRET[+],

25 El análisis de datos y el análisis estadístico reflejan el enfoque usado en el Ejemplo 2, por lo que los detalles del Ejemplo 2 no se repiten aquí. Los ARN_g que están enriquecidos en las células FRET[-] del día 14 en relación con las células FRET[+] del día 14, día 7 y día 10, y/o están agotados en las células FRET[+] del día 14 en relación con las células FRET[-] del día 14, día 7 y día 10, pueden indicar genes diana que, cuando se inactivan (o se activan transcripcionalmente en el caso del cribado CRISPRa), inducen la desagregación de tau. Los ARN_g que están enriquecidos en las células FRET[+] del día 14 en relación con las células FRET[-] del día 14, del día 7 y del día 10, y/o están agotados en las células FRET[-] del día 14 en relación con las células FRET[+] del día 14, del día 7 y del día 10, pueden indicar genes diana que, cuando se inactivan (o se activan transcripcionalmente en el caso del cribado CRISPRa), promueven o potencian la agregación de tau.

REIVINDICACIONES

1. Un método de cribado de modificadores genéticos de la agregación de tau, que comprende:

5 (a) proporcionar una población de células que comprende una proteína Cas, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador,

en donde las células son células de mamífero,

10 en donde el primer indicador y el segundo indicador son proteínas fluorescentes, y en donde el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET);

(b) introducir en la población de células una biblioteca que comprenda una pluralidad de ARN guía únicos que se dirigen a una pluralidad de genes;

15 (c) cultivar la población de células para permitir la edición y expansión del genoma, en donde la pluralidad de ARN guía únicos forman complejos con la proteína Cas, y la proteína Cas corta la pluralidad de genes resultando en la inactivación de la función génica para producir una población de células modificada genéticamente;

(d) poner en contacto la población de células modificada genéticamente con un agente de siembra de tau para producir una población de células sembrada,

20 en donde la etapa (d) comprende cultivar la población de células modificada genéticamente en presencia de medio acondicionado obtenido de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, en donde el medio acondicionado se obtuvo después de estar en células positivas a la agregación de tau confluentes durante 1 a 7 días;

25 (e) cultivar la población de células sembrada para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto de la población de células sembrada para producir una población de células positiva a la agregación; y

(f) determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células modificada genéticamente en la etapa (c),

30 en donde el enriquecimiento de un ARN guía en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células cultivada en la etapa (c) indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la alteración del gen al que se dirige el ARN guía potencia la agregación de tau.

35 2. El método de la reivindicación 1, en donde la proteína Cas es una proteína Cas9, opcionalmente en donde la proteína Cas es una proteína Cas9 de *Streptococcus pyogenes*, opcionalmente en donde la proteína Cas comprende SEQ ID NO: 21, y opcionalmente en donde la proteína Cas está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 22.

40 3. El método de cualquier reivindicación precedente, en donde la proteína Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se expresan de forma estable en la población de células, o en donde los ácidos nucleicos que codifican la proteína Cas, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador están integrados genómicamente en la población de células.

45 4. El método de cualquier reivindicación precedente, en donde cada ARN guía se dirige a un exón constitutivo, opcionalmente en donde cada ARN guía se dirige a un exón constitutivo en 5', o en donde cada ARN guía se dirige a un primer exón, un segundo exón o un tercer exón.

50 5. Un método de cribado de modificadores genéticos de la agregación de tau, que comprende:

(a) proporcionar una población de células que comprende una proteína Cas quimérica que comprende una proteína Cas inactiva por nucleasa fusionada a uno o más dominios de activación transcripcional, una proteína adaptadora quimérica que comprenda una proteína adaptadora fusionada a uno o más dominios de activación transcripcional, un primer dominio de repetición de tau unido a un primer indicador y un segundo dominio de repetición de tau unido a un segundo indicador,

en donde las células son células de mamífero,

60 en donde el primer indicador y el segundo indicador son proteínas fluorescentes, y en donde el primer indicador y el segundo indicador son un par de transferencia de energía por resonancia de fluorescencia (FRET);

(b) introducir en la población de células una biblioteca que comprenda una pluralidad de ARN guía únicos dirigidos a una pluralidad de genes;

65 (c) cultivar la población de células para permitir la activación transcripcional y la expansión, en donde la pluralidad de ARN guía únicos forman complejos con la proteína Cas quimérica y la proteína adaptadora quimérica, y los complejos activan la transcripción de la pluralidad de genes dando lugar a un aumento de la expresión génica para

- producir una población de células modificada genéticamente;
 (d) poner en contacto la población de células modificada genéticamente con un agente de siembra de tau para producir una población de células sembrada,
 5 en donde la etapa (d) comprende cultivar la población de células modificada genéticamente en presencia de medio acondicionado obtenido de células positivas a la agregación de tau cultivadas en las que un dominio de repetición de tau se presenta de forma estable en un estado agregado, en donde el medio acondicionado se obtuvo después de estar en células positivas a la agregación de tau confluentes durante 1 a 7 días;
 (e) cultivar la población de células sembrada para permitir que se formen agregados de tau, en donde se forman agregados del primer dominio de repetición de tau y del segundo dominio de repetición de tau en un subconjunto
 10 de la población de células sembrada para producir una población de células positiva a la agregación; y
 (f) determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células modificada genéticamente en la etapa (c),
 15 en donde el enriquecimiento de un ARN guía en la población de células positiva a la agregación identificada en la etapa (e) en relación con la población de células cultivada en la etapa (c) indica que el gen al que se dirige el ARN guía es un modificador genético de la agregación de tau, en donde la activación transcripcional del gen al que se dirige el ARN guía aumenta la agregación de tau.
- 20 6. El método de la reivindicación 5, en donde la proteína Cas es una proteína Cas9, opcionalmente en donde la proteína Cas es una proteína Cas9 de *Streptococcus pyogenes*, y
- 25 en la que la proteína Cas quimérica comprende la proteína Cas inactiva por nucleasa fusionada a un dominio de activación transcripcional VP64, opcionalmente en donde la proteína Cas quimérica comprende de extremo N a extremo C: la proteína Cas inactiva por nucleasa; una señal de localización nuclear; y el dominio activador transcripcional VP64, opcionalmente en donde la proteína Cas quimérica comprende SEQ ID NO: 36, y opcionalmente en donde la proteína Cas quimérica está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 38, y
- 30 en donde la proteína adaptadora es una proteína de la cubierta de MS2, y en donde el uno o más dominios de activación transcripcional en la proteína adaptadora quimérica comprenden un dominio de activación transcripcional p65 y un dominio de activación transcripcional HSF1, opcionalmente en donde la proteína adaptadora quimérica comprende desde el extremo N hasta el extremo C: la proteína de la cubierta de MS2; una señal de localización nuclear; el dominio de activación transcripcional p65; y el dominio de activación transcripcional HSF1, opcionalmente en donde la proteína adaptadora quimérica comprende SEQ ID NO: 37, y opcionalmente en
 35 donde la proteína adaptadora quimérica está codificada por una secuencia codificante que comprende la secuencia establecida en SEQ ID NO: 39.
7. El método de la reivindicación 5 o 6, en donde la proteína Cas quimérica, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo
 40 indicador se expresan de forma estable en la población de células, o en donde los ácidos nucleicos que codifican la proteína Cas quimérica, la proteína adaptadora quimérica, el primer dominio de repetición de tau unido al primer indicador y el segundo dominio de repetición de tau unido al segundo indicador se integran genómicamente en la población de células.
- 45 8. El método de cualquiera de las reivindicaciones 5-7, en donde cada ARN guía se dirige a una secuencia diana de ARN guía dentro de 200 pb aguas arriba de un sitio de inicio de la transcripción, y
- 50 en donde cada ARN guía comprende uno o más elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica, opcionalmente en donde cada ARN guía comprende dos elementos de unión al adaptador a los que puede unirse específicamente la proteína adaptadora quimérica, opcionalmente en donde un primer elemento de unión al adaptador está dentro de un primer bucle de cada uno de los uno o más ARN guía, y un segundo elemento de unión al adaptador está dentro de un segundo bucle de cada uno de los uno o más ARN guía, opcionalmente en donde el elemento de unión al adaptador comprende la secuencia establecida en SEQ ID NO: 33, y
 55 opcionalmente, en donde cada uno de uno o más ARN guía es un ARN guía único que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de ARN CRISPR transactivador (ARNtracr), y el primer bucle es el tetrabucle correspondiente a los residuos 13-16 de SEQ ID NO: 17, y el segundo bucle es el tallo-bucle 2 correspondiente a los residuos 53-56 de SEQ ID NO: 17.
- 60 9. El método de cualquier reivindicación precedente, en donde la etapa (c) es de aproximadamente 3 días a aproximadamente 9 días, opcionalmente en donde la etapa (c) es de aproximadamente 6 días.
10. El método de cualquier reivindicación precedente, en donde el medio acondicionado se obtuvo después de estar en células confluentes positivas a la agregación de tau durante aproximadamente 4 días.
- 65 11. El método de cualquier reivindicación precedente, en donde la etapa (d) comprende cultivar la población de células modificada genéticamente en aproximadamente 75 % de medio acondicionado y aproximadamente 25 % de medio

recién preparado.

12. El método de cualquier reivindicación precedente, en donde la población de células modificada genéticamente no se cocultiva con las células positivas a la agregación tau.

13. El método de cualquier reivindicación precedente, en donde la etapa (e) es de aproximadamente 2 días a aproximadamente 6 días, opcionalmente en donde la etapa (e) es de aproximadamente 4 días.

14. El método de cualquier reivindicación precedente, en donde la abundancia se determina mediante secuenciación de nueva generación.

15. El método de cualquier reivindicación precedente, en donde un ARN guía se considera enriquecido si la abundancia del ARN guía en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c).

16. El método de cualquier reivindicación precedente, en donde la etapa (f) comprende determinar la abundancia de cada uno de la pluralidad de ARN guía únicos en la población de células positiva a la agregación en la etapa (e) en relación con la población de células cultivada en la etapa (c) en un primer momento en la etapa (c) y/o un segundo momento en la etapa (c),

opcionalmente, en donde el primer momento de la etapa (c) está en un primer pase del cultivo de la población de células, y el segundo momento se encuentra a mitad del cultivo de la población de células para permitir la edición y expansión del genoma o la activación y expansión transcripcional, y

opcionalmente, en donde el primer momento de la etapa (c) es después de aproximadamente tres días de cultivo, y el segundo momento de la etapa (c) es después de aproximadamente seis días de cultivo.

17. El método de la reivindicación 16, en donde un gen se considera un modificador genético de la agregación de tau, en donde la inactivación o la activación transcripcional del gen potencia la agregación de tau, si:

(1) la abundancia de un ARN guía que se dirige al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación de la etapa (e) en relación con la población de células cultivada de la etapa (c) tanto en el primer momento de la etapa (c) como en el segundo momento de la etapa (c); y/o

(2) la abundancia de al menos dos ARN guía únicos dirigidos al gen en relación con la población total de la pluralidad de ARN guía únicos es al menos 1,5 veces mayor en la población de células positiva a la agregación de la etapa (e) en relación con la población de células cultivada de la etapa (c) en el primer momento de la etapa (c) o en el segundo momento de la etapa (c).

18. El método de cualquier reivindicación precedente, en donde los siguientes pasos se realizan en la etapa (f) para identificar un gen como modificador genético de la agregación de tau, en donde la inactivación o activación transcripcional del gen potencia la agregación de tau:

(1) identificar cuál de la pluralidad de ARN guía únicos está presente en la población de células positiva a la agregación producida en la etapa (e);

(2) calcular la probabilidad aleatoria de que los ARN guía identificados en la etapa (f)(1) estén presentes usando la fórmula $nCn' * (x-n)C(m-n) / xCm$,

en donde x es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b),

en donde m es la variedad de ARN guía únicos identificados en la etapa (f)(1),

en donde n es la variedad de ARN guía únicos introducidos en la población de células en la etapa (b) que se dirigen al gen, y

en donde n' es la variedad de ARN guía únicos identificados en la etapa (f)(1) que se dirigen al gen;

(3) calcular las puntuaciones promedio de enriquecimiento para los ARN guía identificados en la etapa (f)(1),

en donde la puntuación de enriquecimiento para un ARN guía es la abundancia relativa del ARN guía en la población de células positiva a la agregación producida en la etapa (e) dividida por la abundancia relativa del ARN guía en la población de células cultivada en la etapa (c), y

en donde la abundancia relativa es el recuento de lecturas del ARN guía dividido por el recuento de lecturas de la población total de la pluralidad de ARN guía únicos; y

(4) seleccionar el gen si un ARN guía dirigido al gen está significativamente por debajo de la probabilidad aleatoria de estar presente y por encima de una puntuación de enriquecimiento umbral.

19. El método de cualquier reivindicación precedente, en donde el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau es un dominio de repetición de tau humano, y/o

- 5 en donde el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprenden una mutación proagregación, opcionalmente en donde el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprenden una mutación de tau P301S; y/o
 en donde el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprenden un dominio de cuatro repeticiones de tau; y/o
 en donde el primer dominio de repetición de tau y/o el segundo dominio de repetición de tau comprenden SEQ ID NO: 11; y/o
 10 en donde el primer dominio de repetición de tau y el segundo dominio de repetición de tau son iguales.
20. El método de cualquier reivindicación precedente, en donde el primer dominio de repetición de tau y el segundo dominio de repetición de tau son el mismo y cada uno comprende un dominio de cuatro repeticiones de tau que comprende una mutación P301S de tau.
- 15 21. El método de cualquier reivindicación precedente, en donde el primer indicador es la proteína fluorescente cian (CFP) y el segundo indicador es la proteína fluorescente amarilla (YFP).
22. El método de cualquier reivindicación precedente, en donde las células son células humanas, opcionalmente en donde las células son células HEK293T.
- 20 23. El método de cualquier reivindicación precedente, en donde la pluralidad de ARN guía únicos se introducen a una concentración seleccionada de tal manera que la mayoría de las células reciben solo uno de los ARN guía únicos.
24. El método de cualquier reivindicación precedente, en donde la pluralidad de ARN guía únicos se dirigen a 100 o más genes, 1000 o más genes, o 10000 o más genes, o
 25 en donde la biblioteca es una biblioteca de todo el genoma.
25. El método de cualquier reivindicación precedente, en donde a una pluralidad de secuencias diana se dirigen de media en cada uno de la pluralidad de genes diana,
 30 opcionalmente, en donde a al menos tres secuencias diana se dirigen de media en cada uno de la pluralidad de genes diana,
 opcionalmente, en donde a aproximadamente tres a aproximadamente seis secuencias diana se dirigen de media en cada uno de los genes seleccionados.
- 35 26. El método de cualquier reivindicación precedente, en donde la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción vírica,
 40 opcionalmente, en donde cada uno de la pluralidad de ARN guía únicos está en un vector vírico separado,
 opcionalmente, en donde la pluralidad de ARN guía únicos se introduce en la población de células mediante transducción lentivírica, y
 opcionalmente, en donde la población de células se infecta con una multiplicidad de infección inferior a aproximadamente 0,3.
- 45 27. El método de cualquier reivindicación precedente, en donde la pluralidad de ARN guía únicos se introduce en la población de células junto con un marcador de selección, y la etapa (b) comprende además seleccionar células que comprenden el marcador de selección,
 50 opcionalmente, en donde el marcador de selección confiere resistencia a un fármaco, opcionalmente, en donde el marcador de selección confiere resistencia a la puomicina o a la zeocina, y
 opcionalmente, en donde el marcador de selección se selecciona de neomicina fosfotransferasa, higromicina B fosfotransferasa, puomicina-N-acetiltransferasa y blasticidina S desaminasa.
- 55 28. El método de cualquier reivindicación precedente, en donde la población de células en la que se introduce la pluralidad de ARN guía únicos en la etapa (b) comprende más de aproximadamente 300 células por ARN guía único.

Nota:

El documento completo que incluye las tablas de referencia y el listado de secuencias se puede descargar de la página web de la EPO (<https://register.epo.org>)

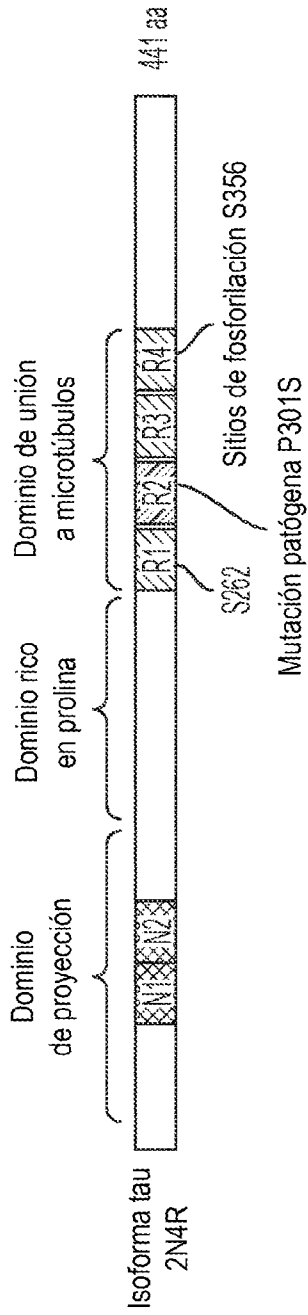
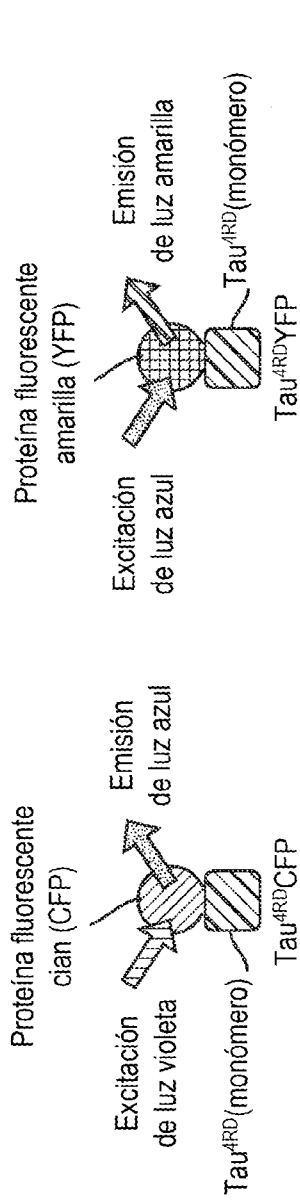
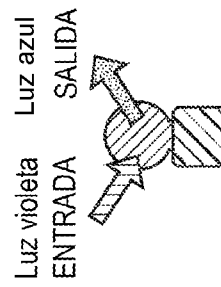


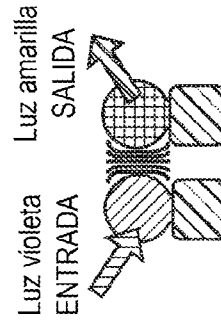
FIG. 1



Sin agregación = SIN FRET



Oligomerización de tau



Agregación de tau = FRET

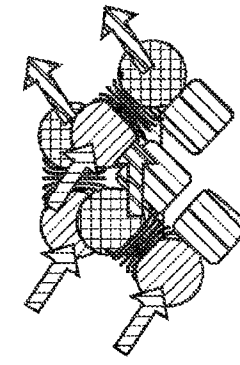


FIG. 2

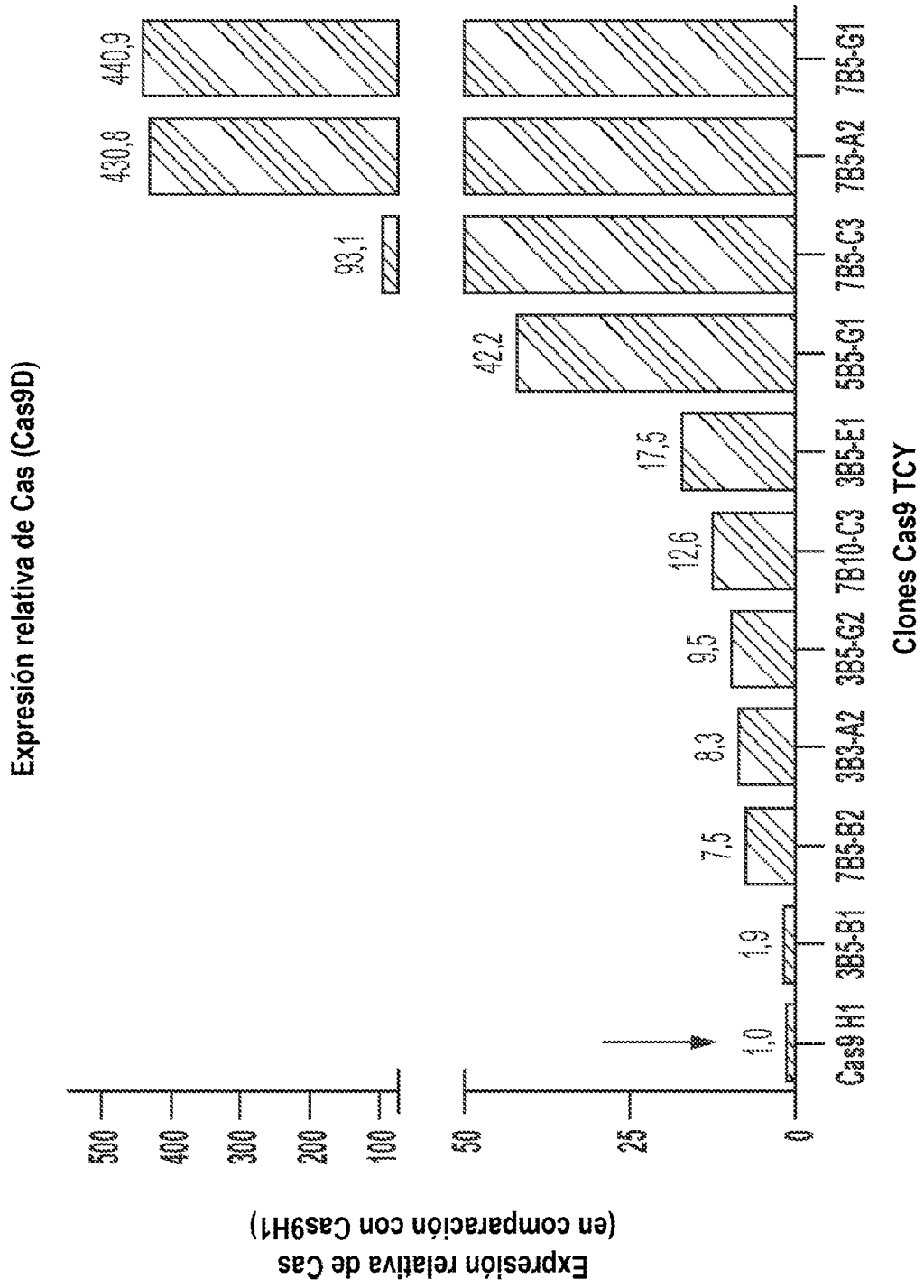


FIG. 3A

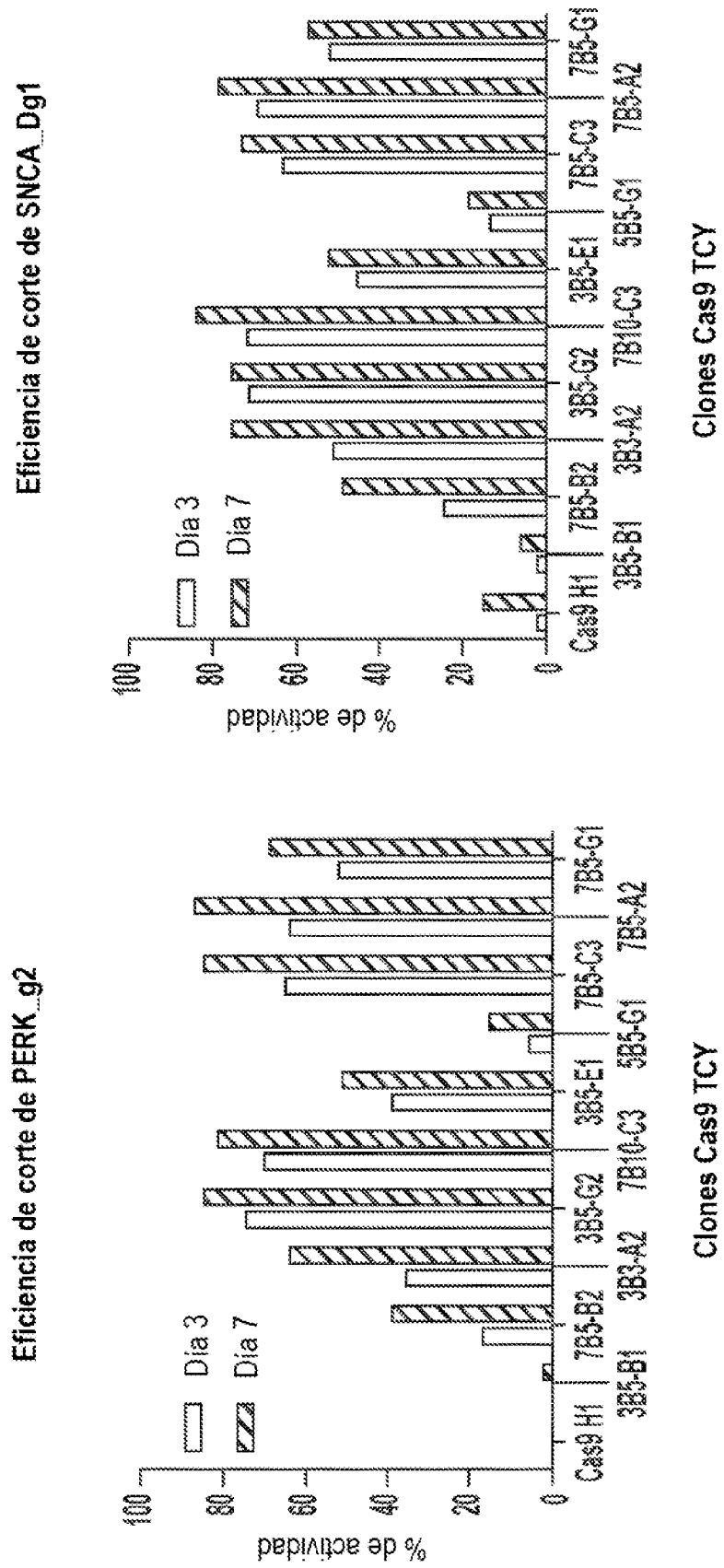


FIG. 3B

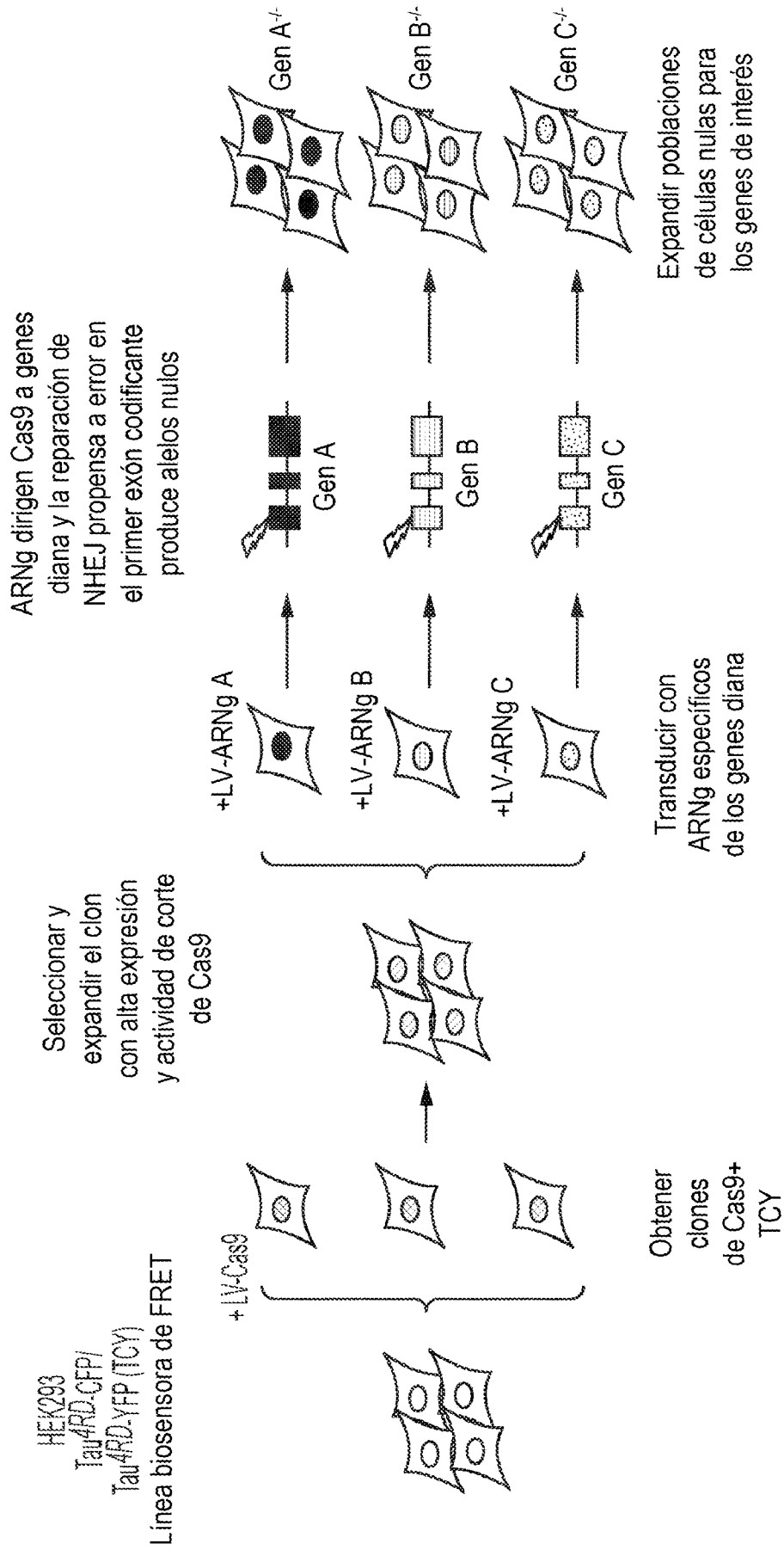


FIG. 4

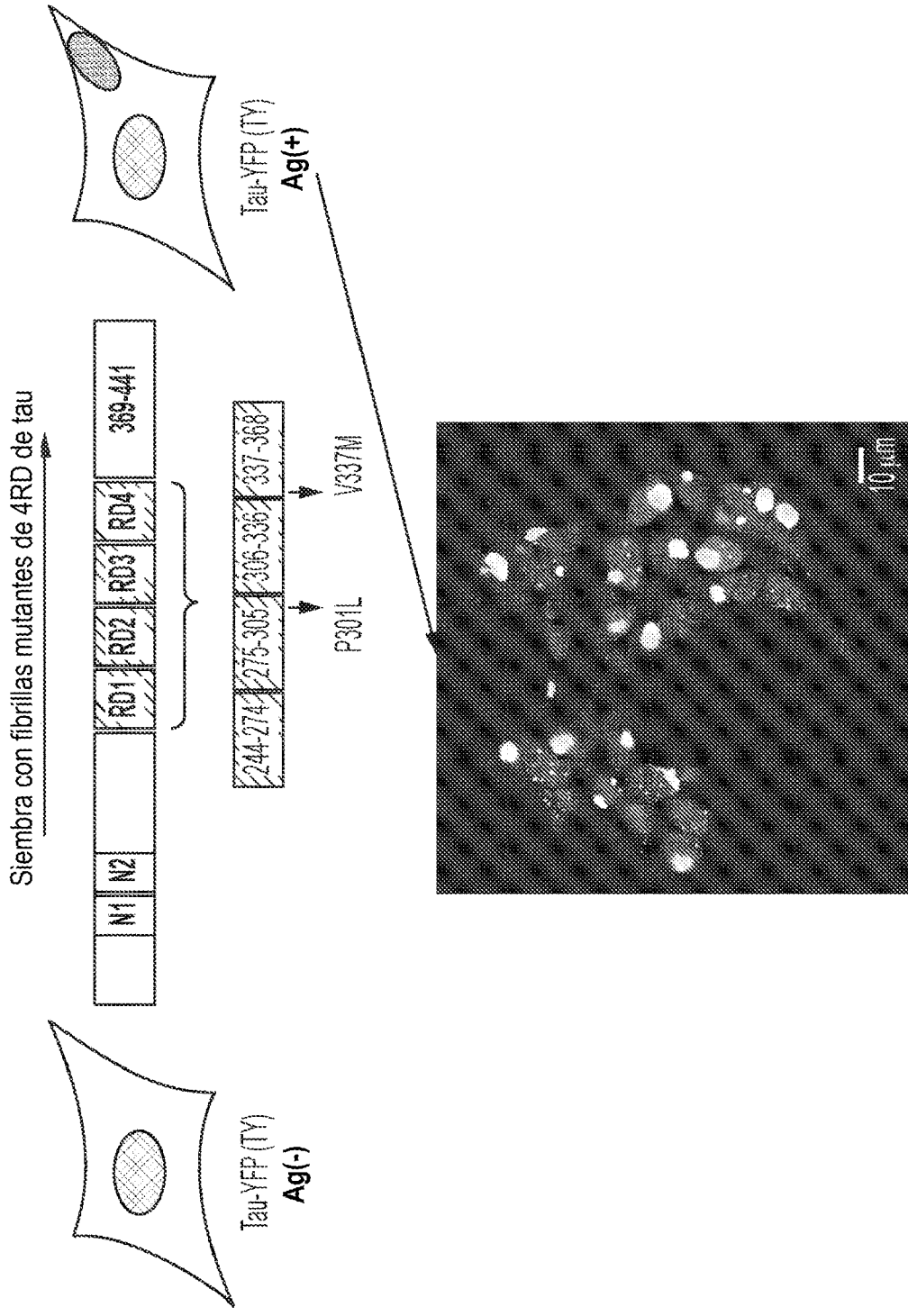


FIG. 5

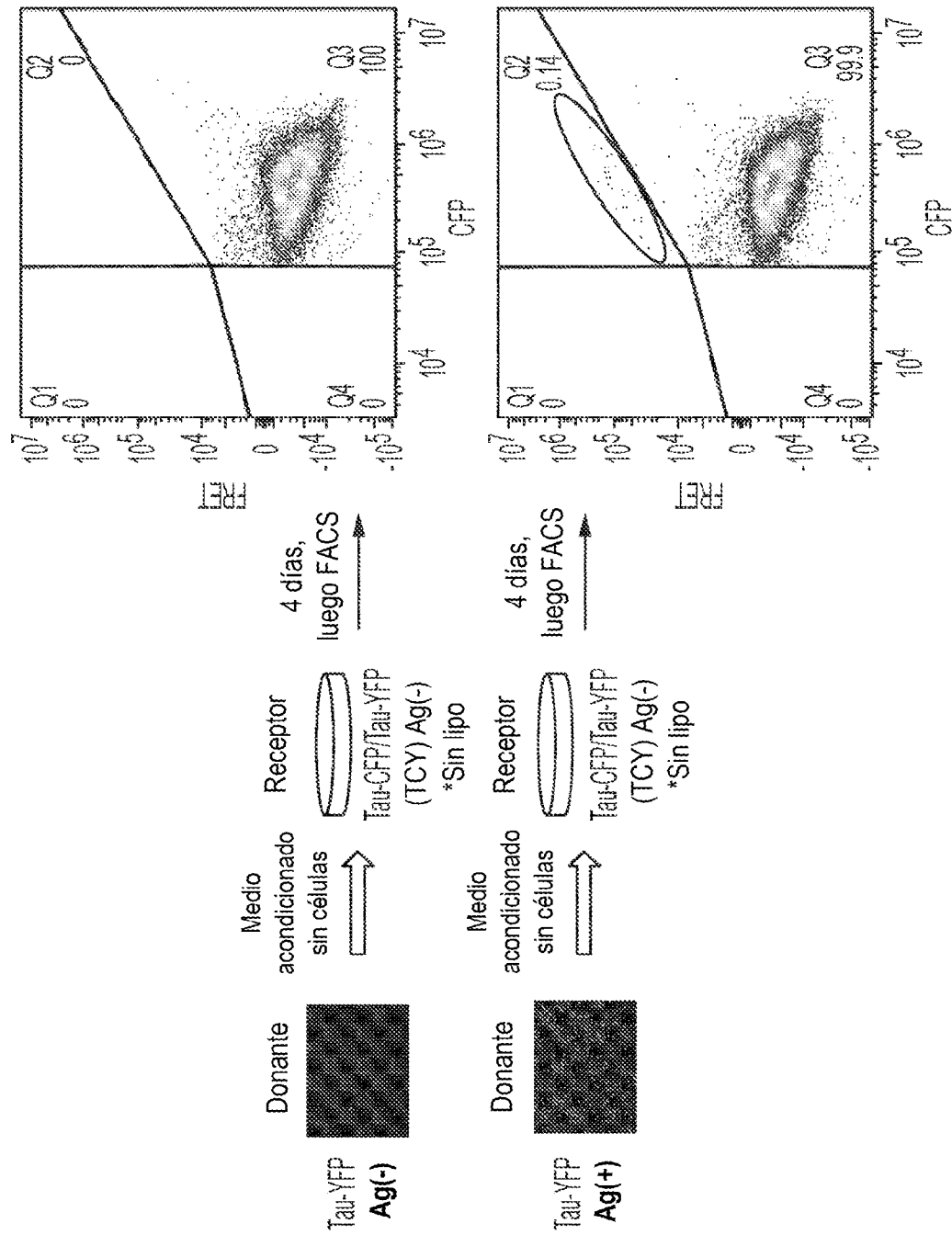


FIG. 6

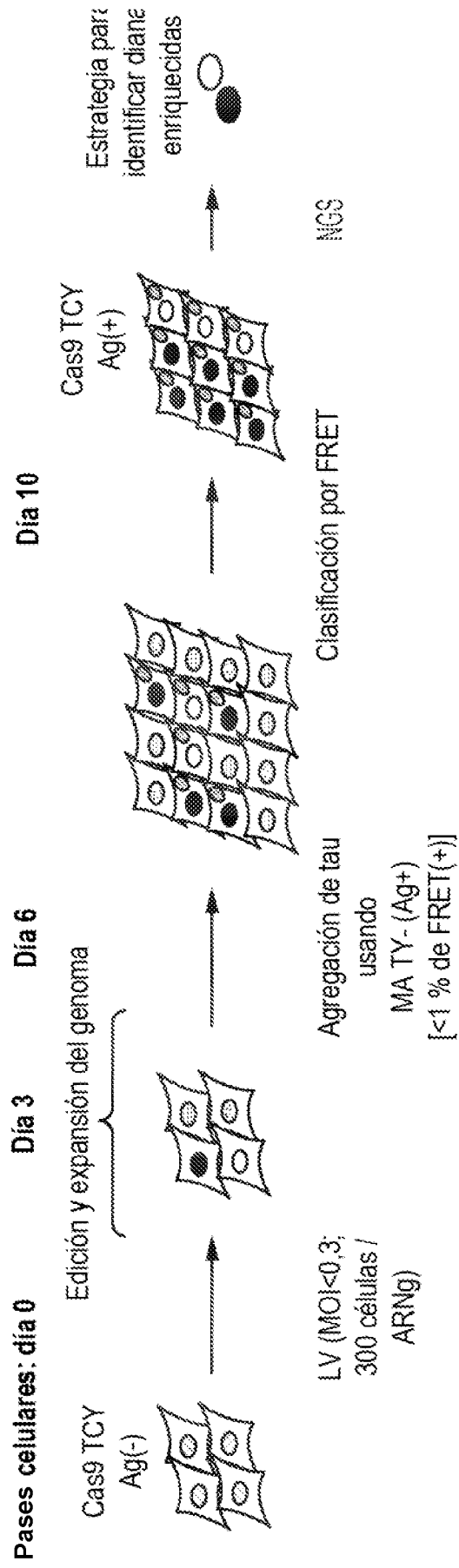


FIG. 7

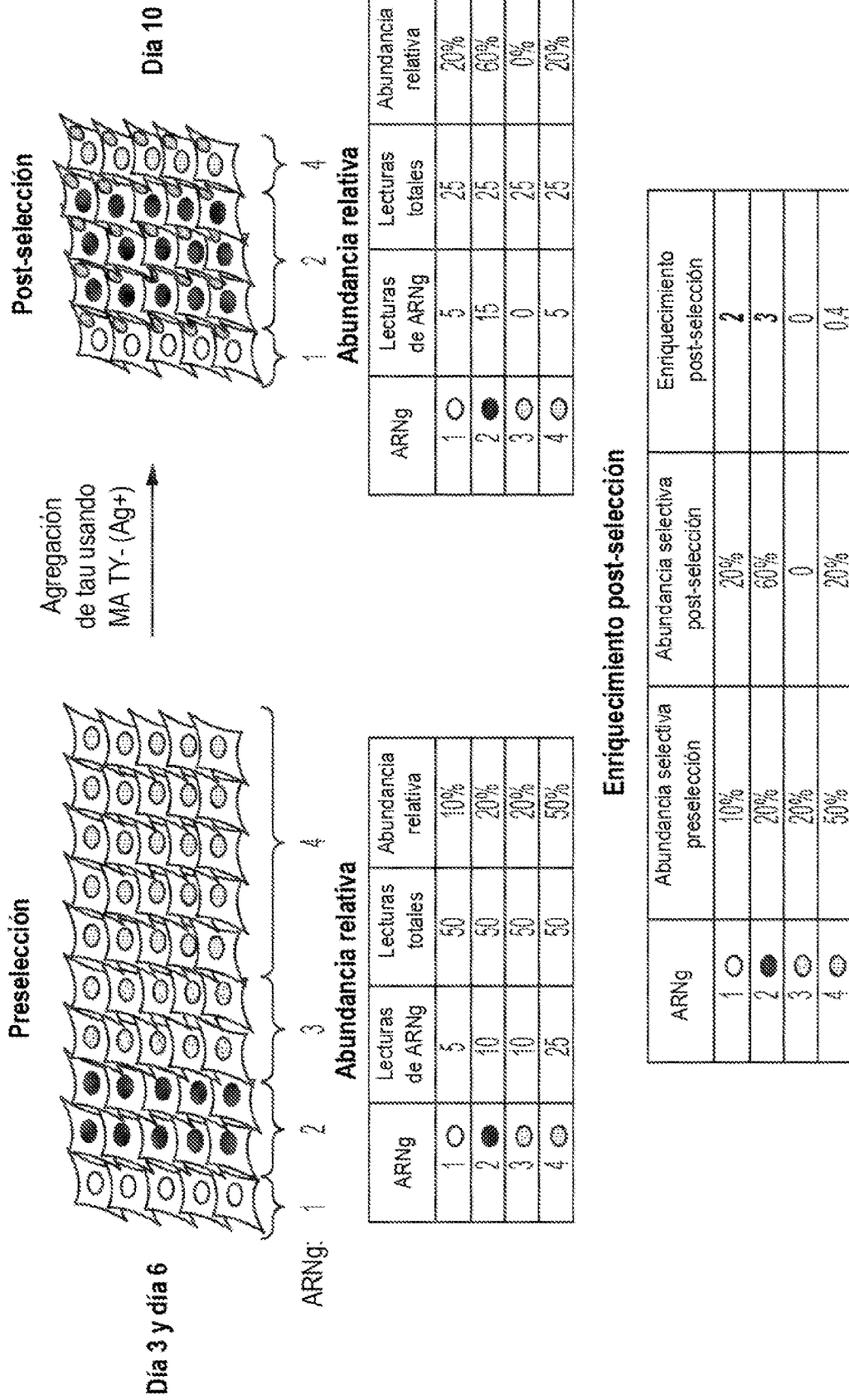


FIG. 8

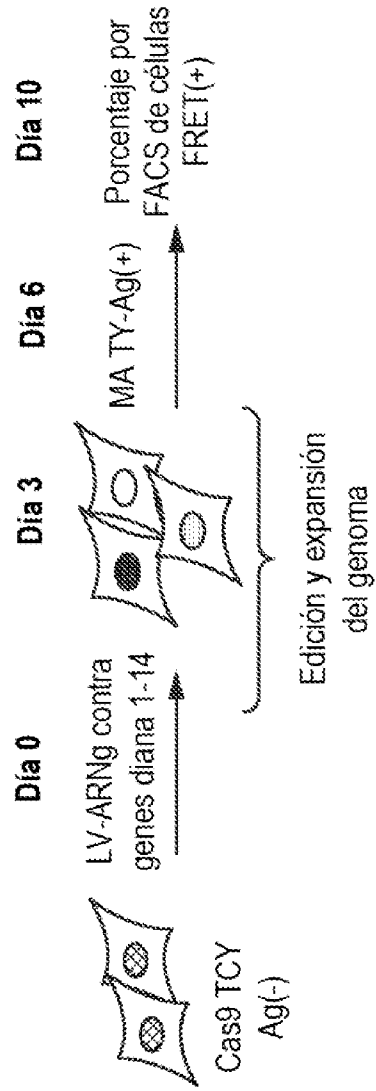


FIG. 9

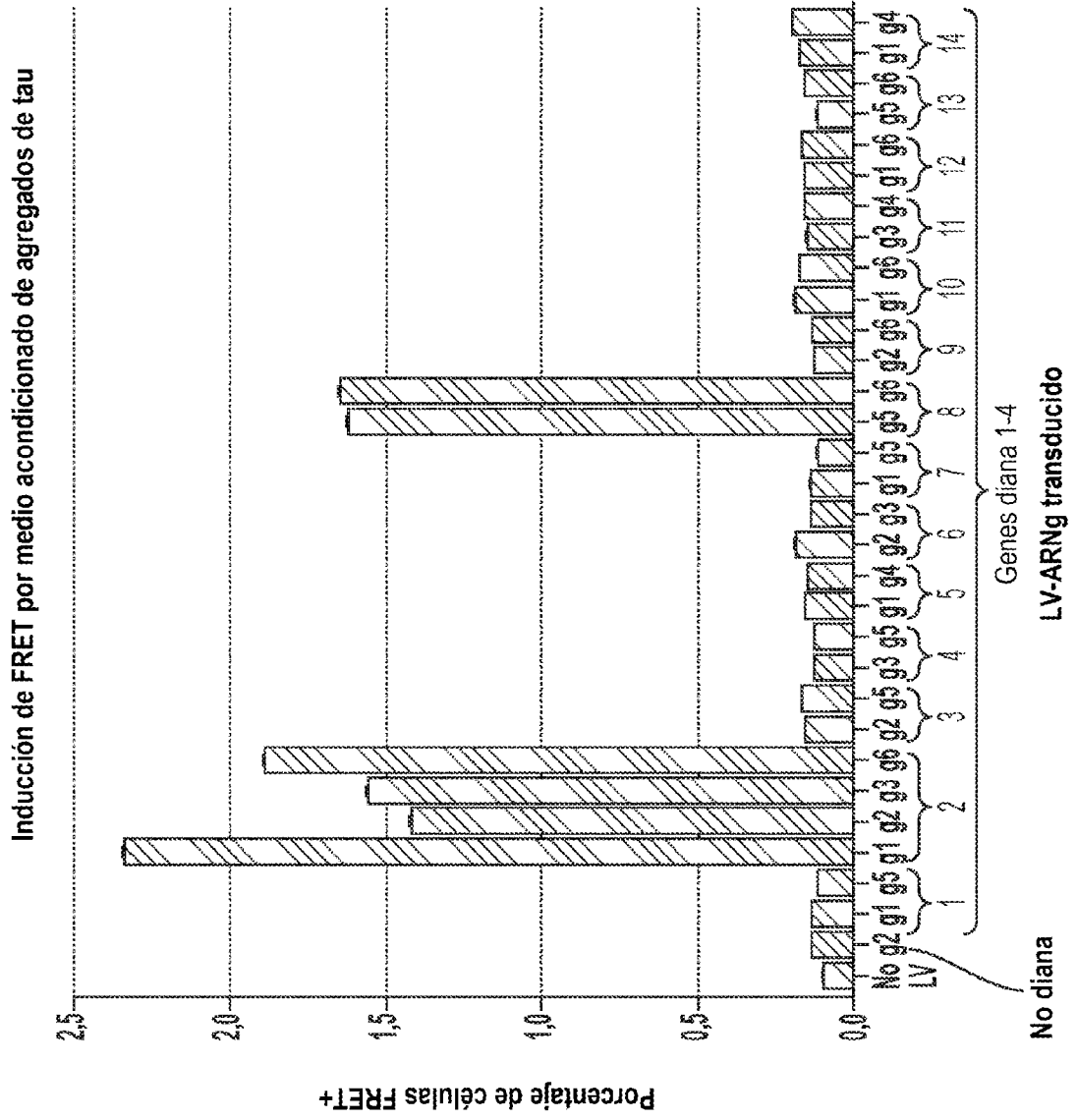


FIG. 10

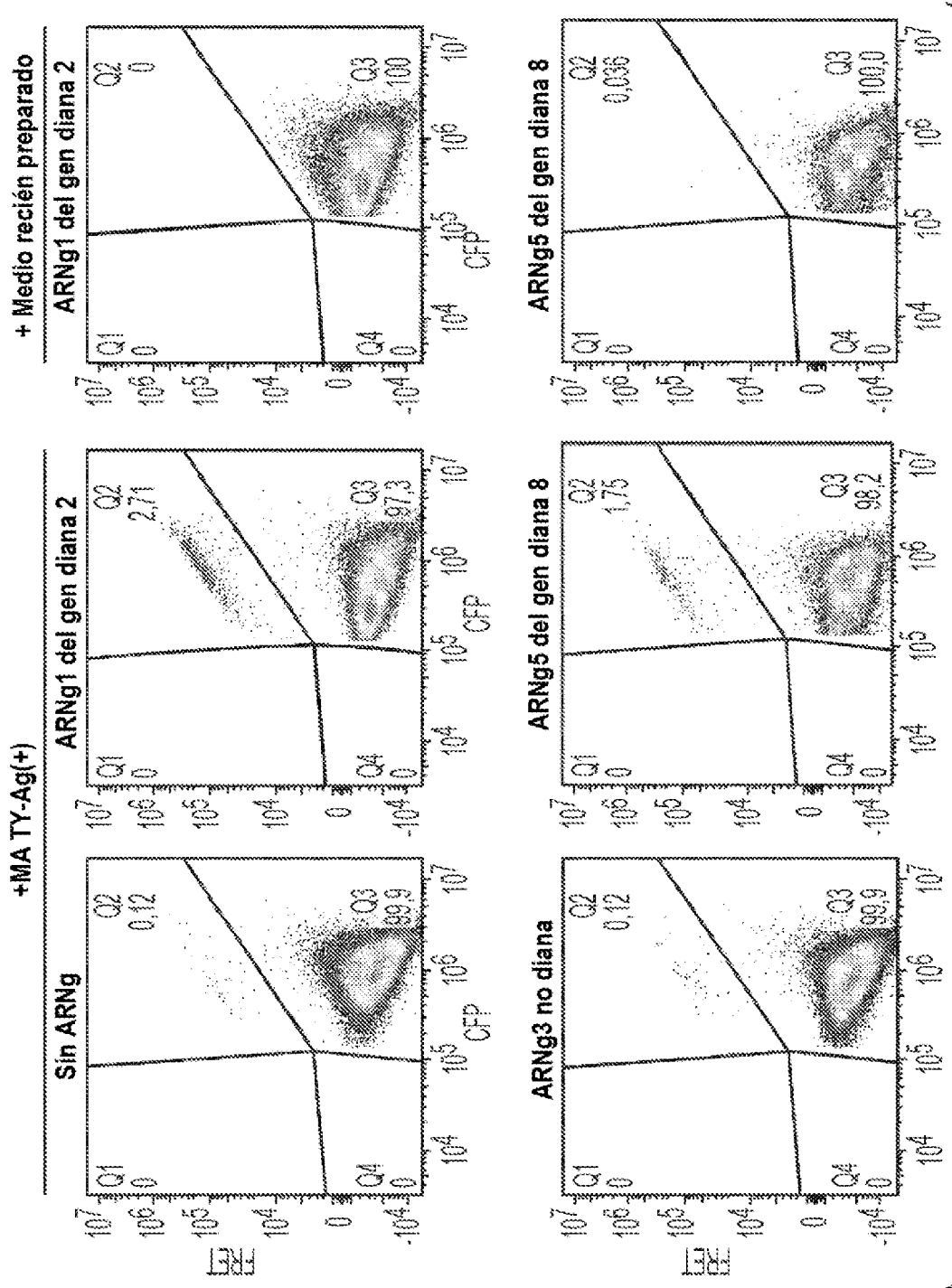


FIG. 11

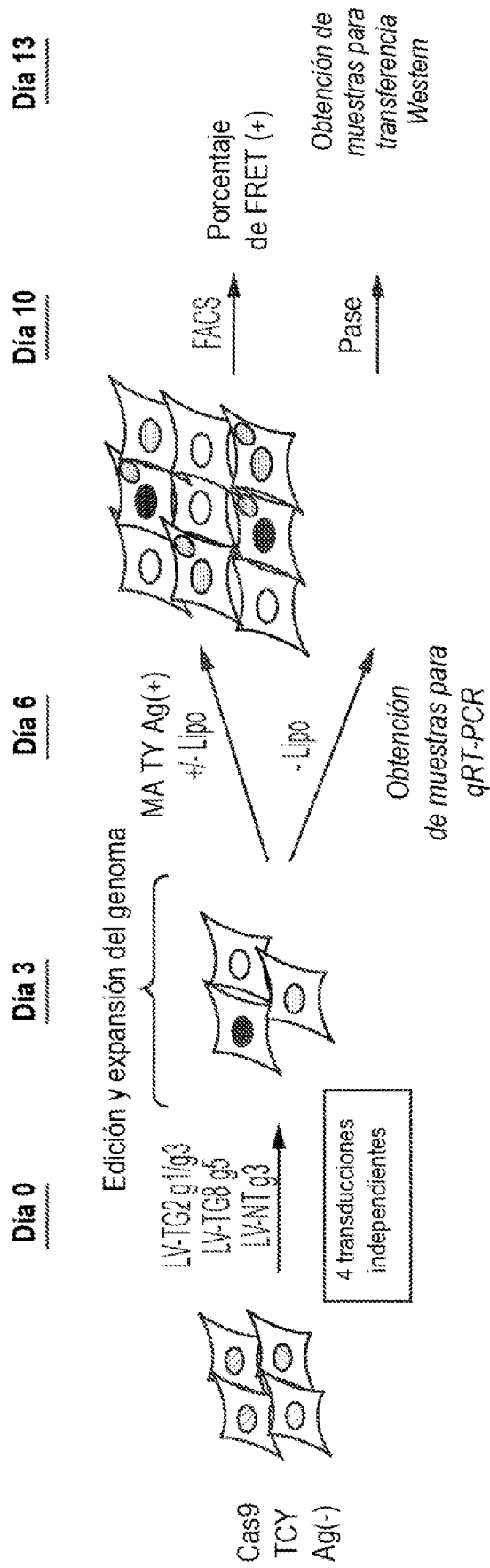
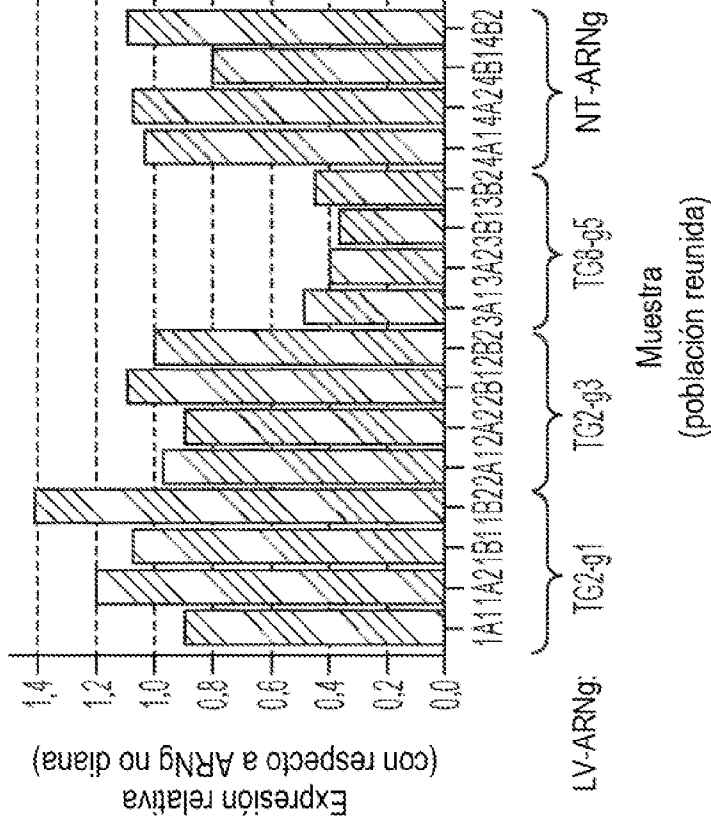


FIG. 12

Expresión del gen diana 8



Expresión del gen diana 2

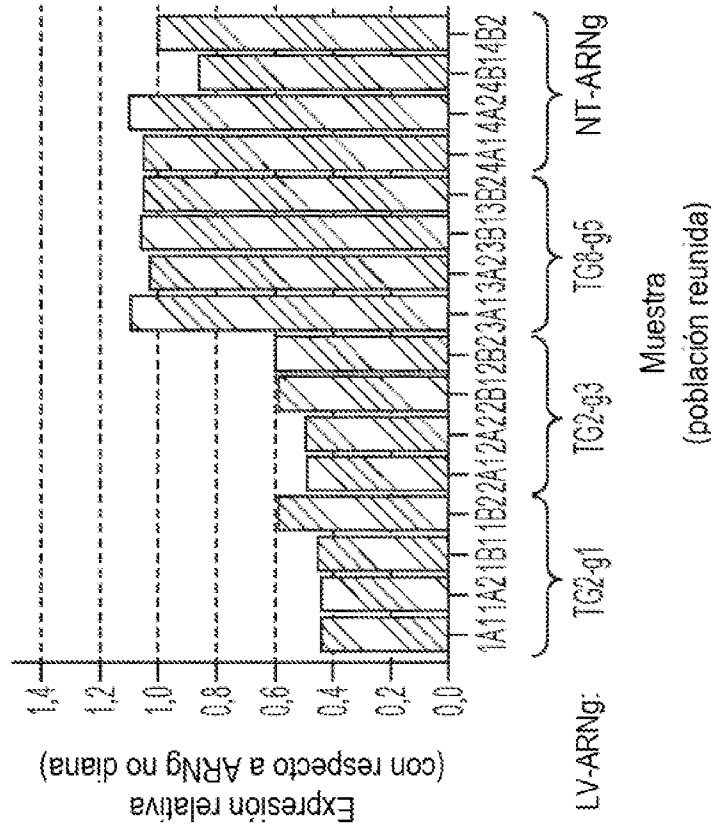


FIG. 13

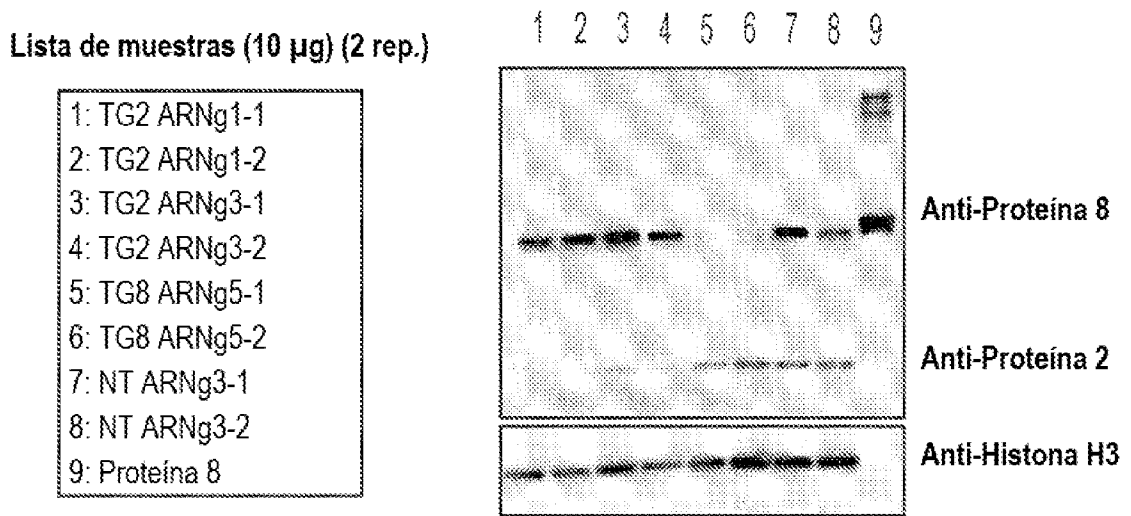


FIG. 14

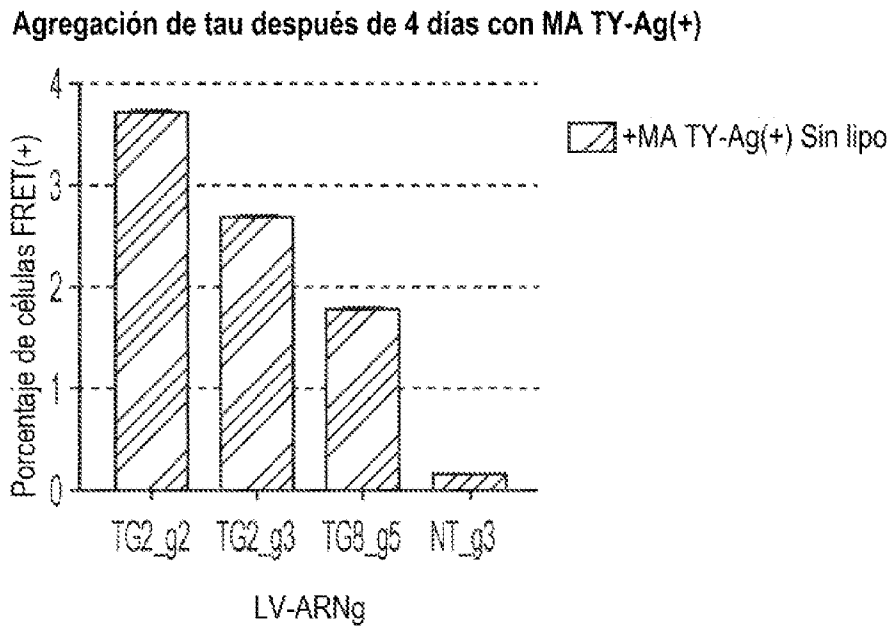


FIG. 15

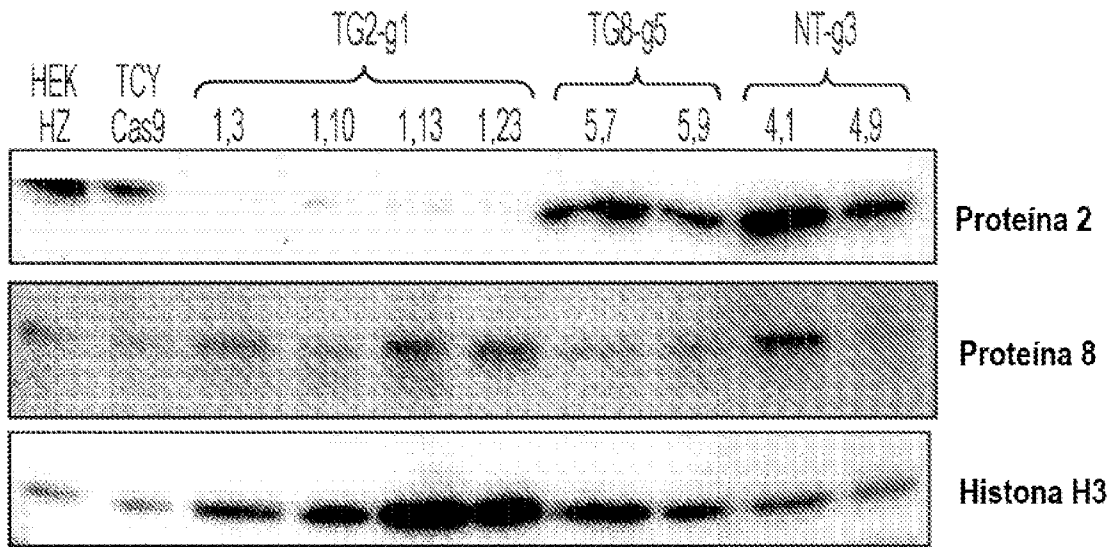


FIG. 16

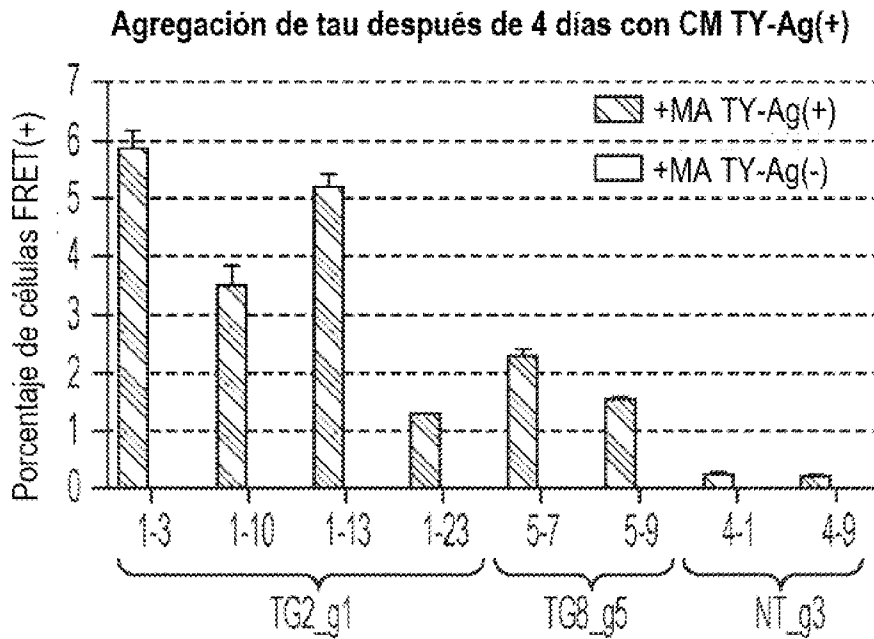


FIG. 17

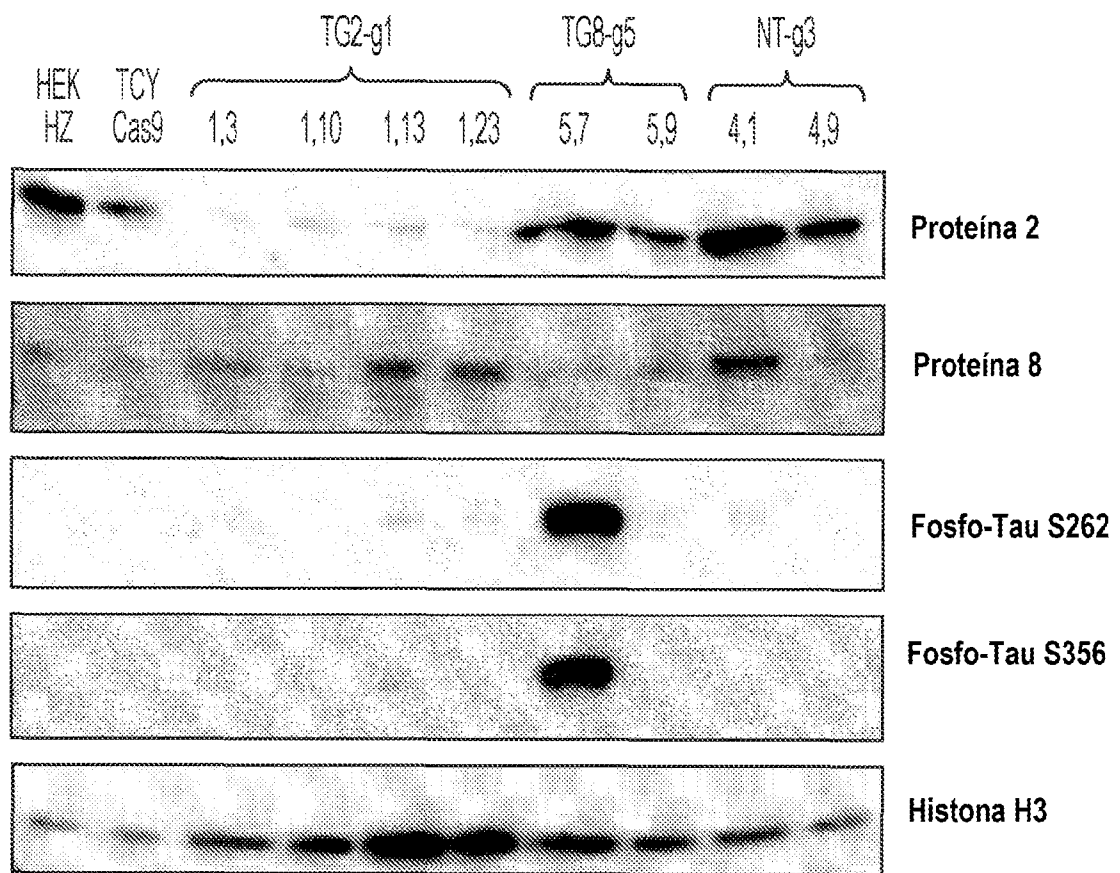


FIG. 18

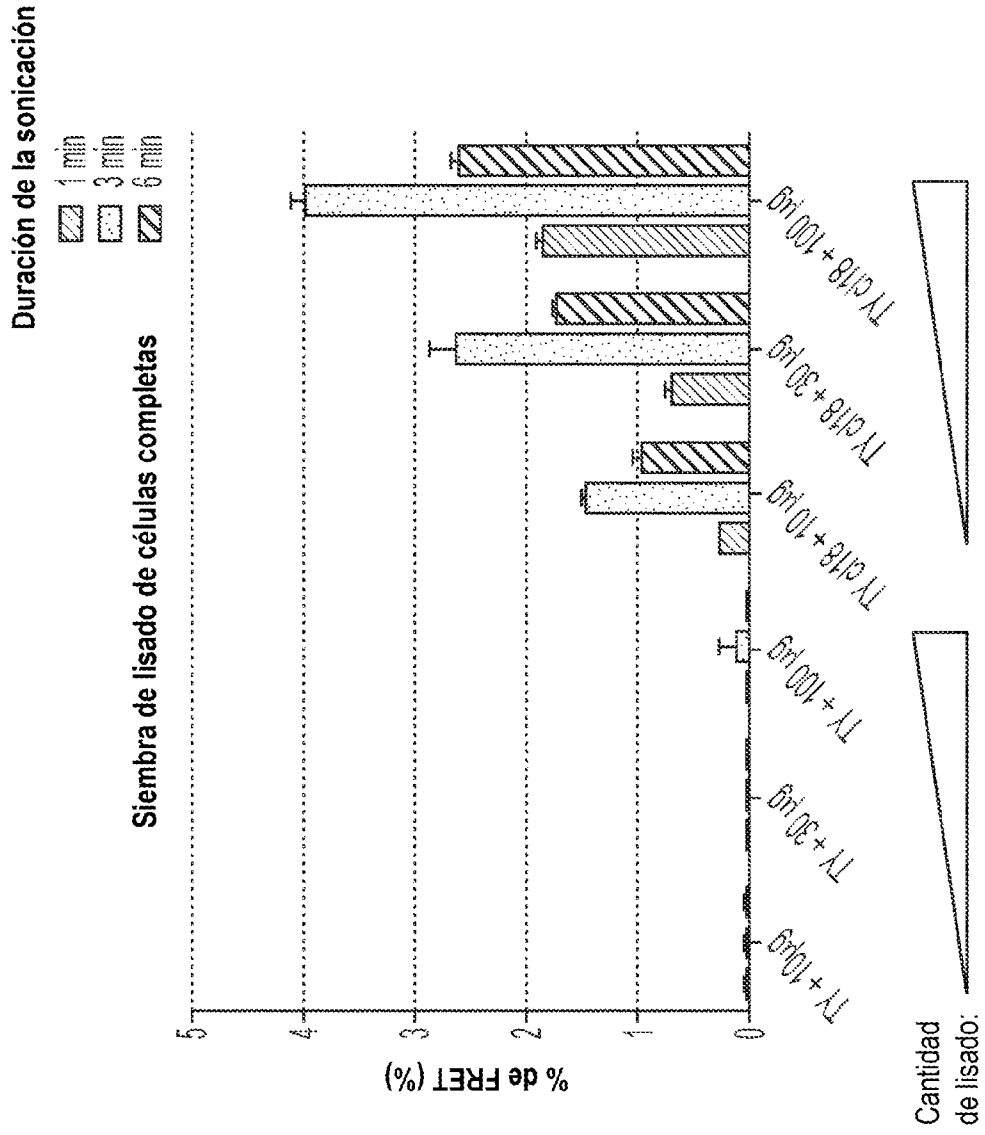


FIG. 19

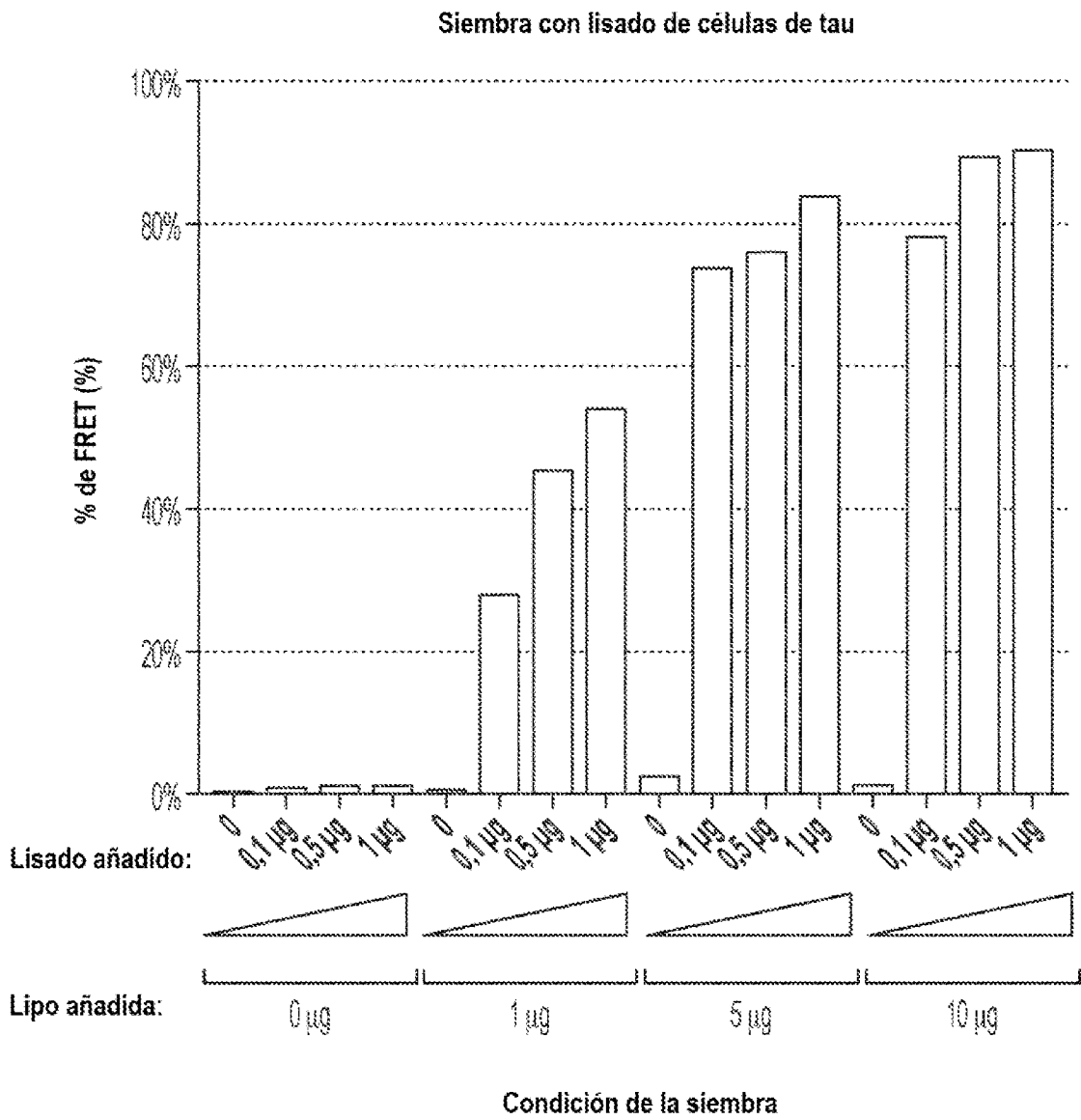
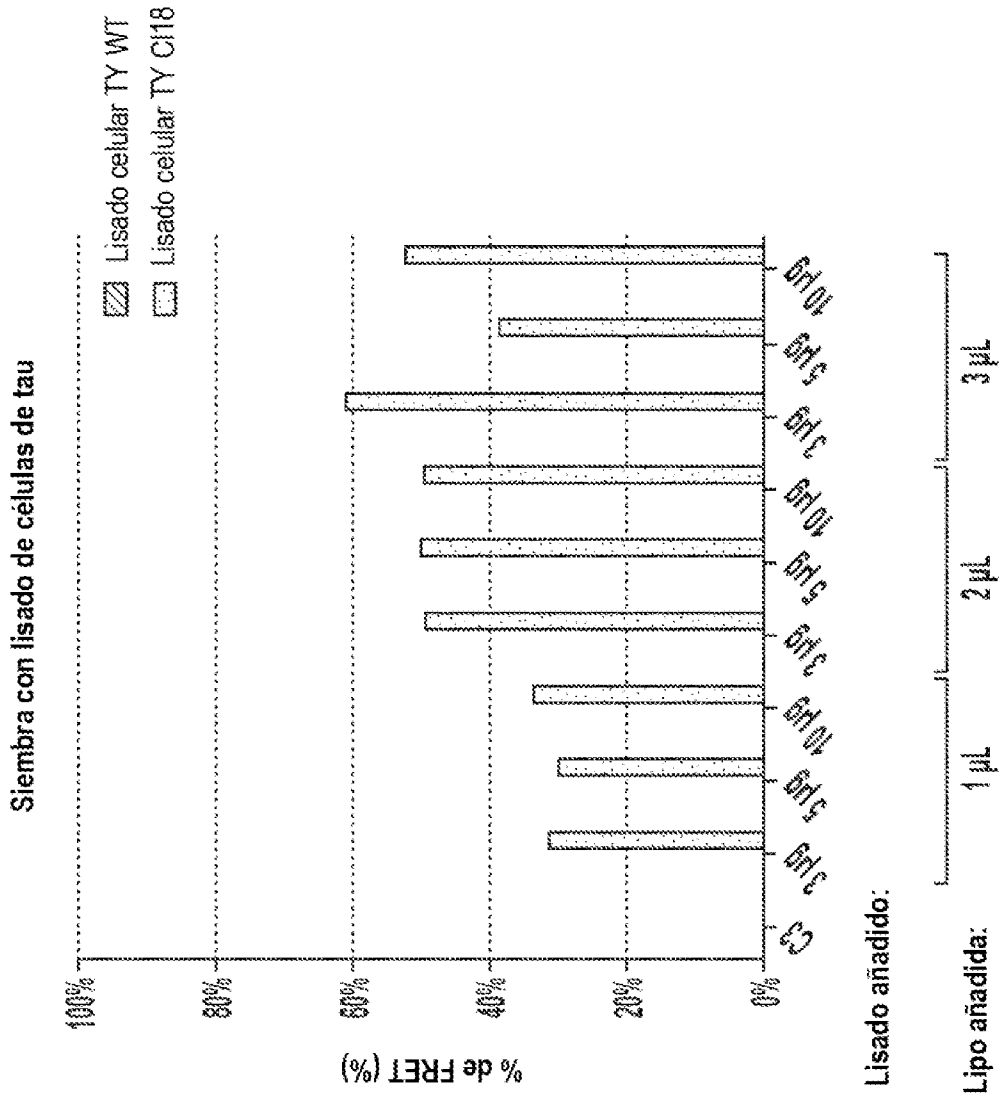


FIG. 20



Condición de la siembra
FIG. 21

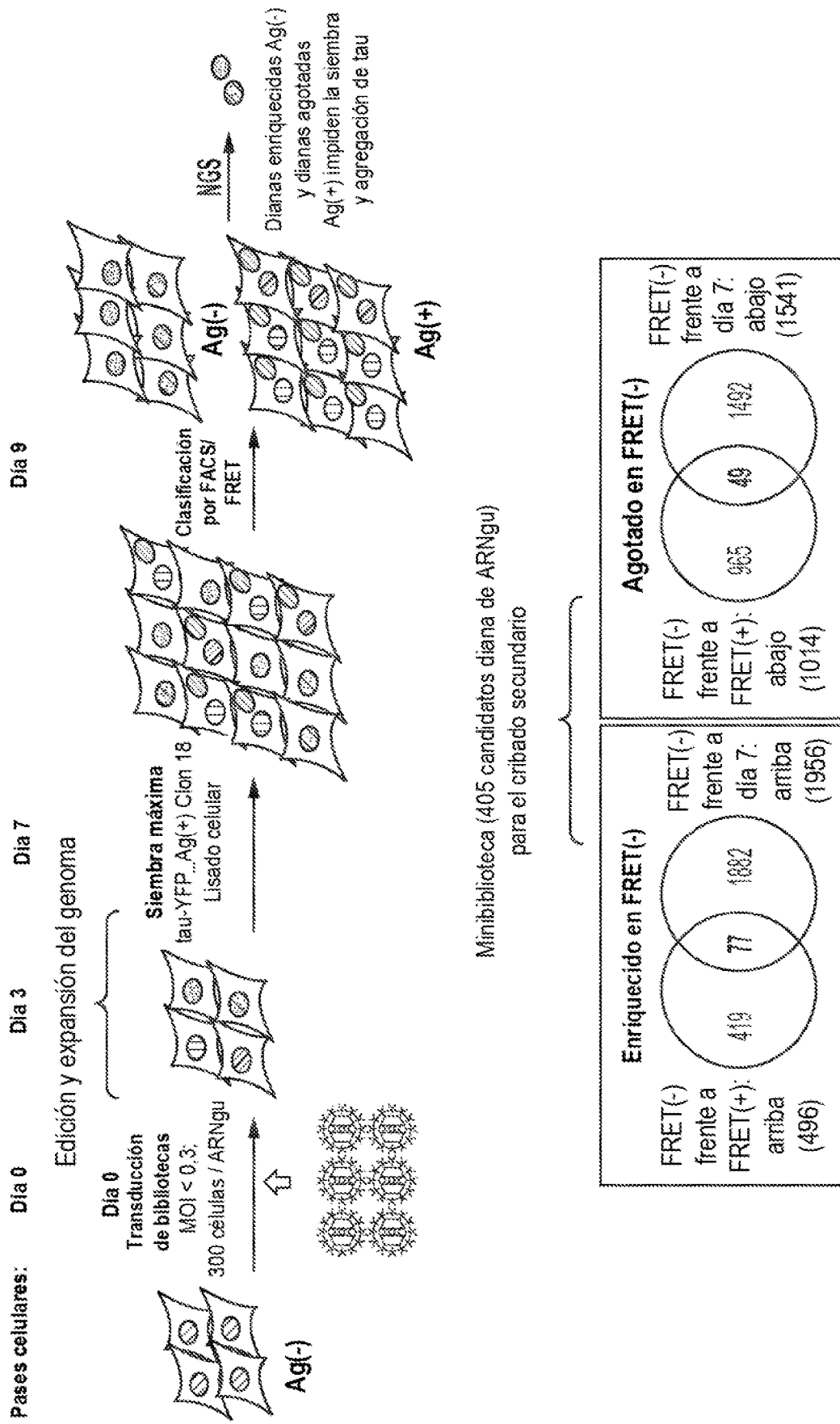
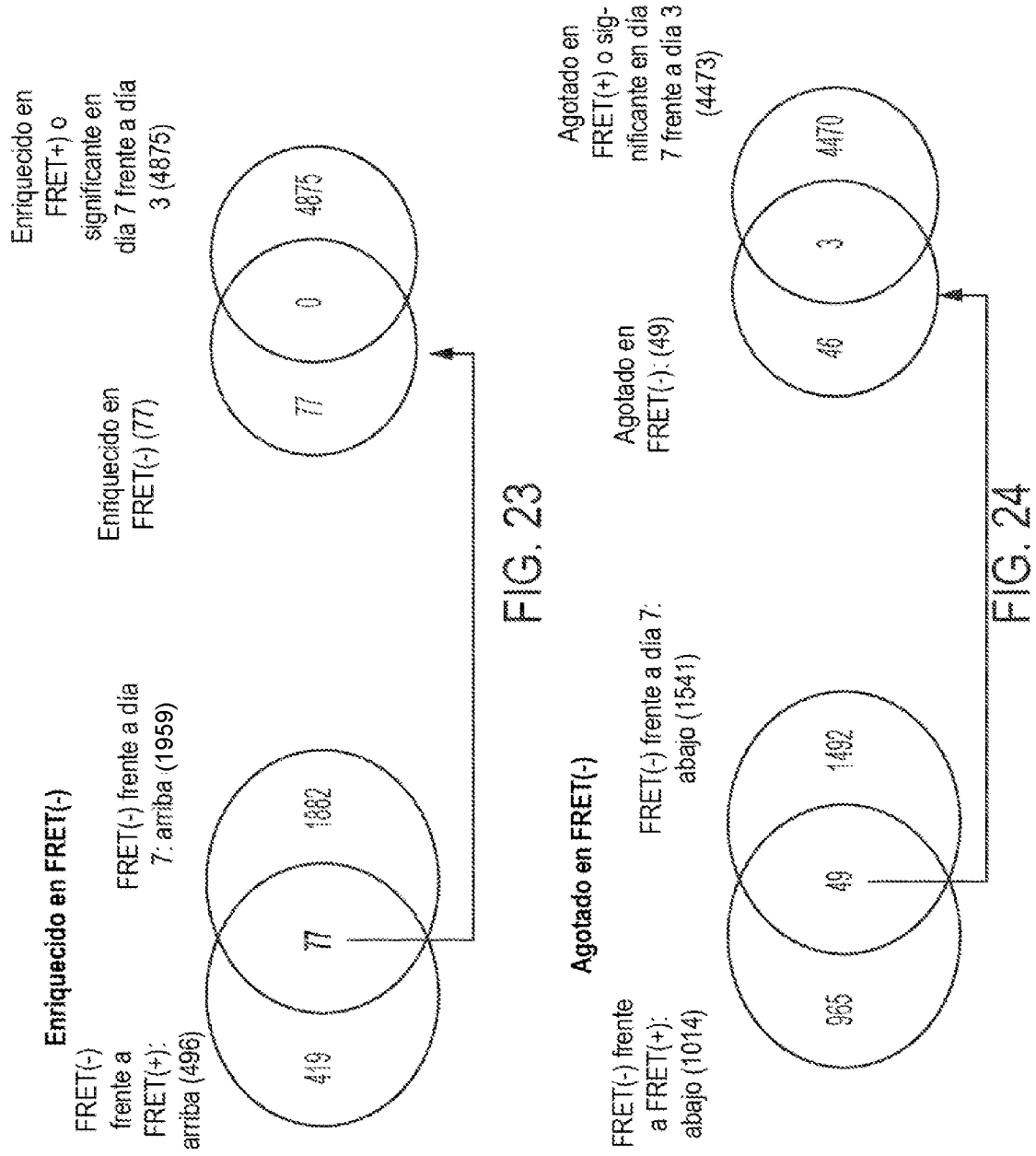


FIG. 22



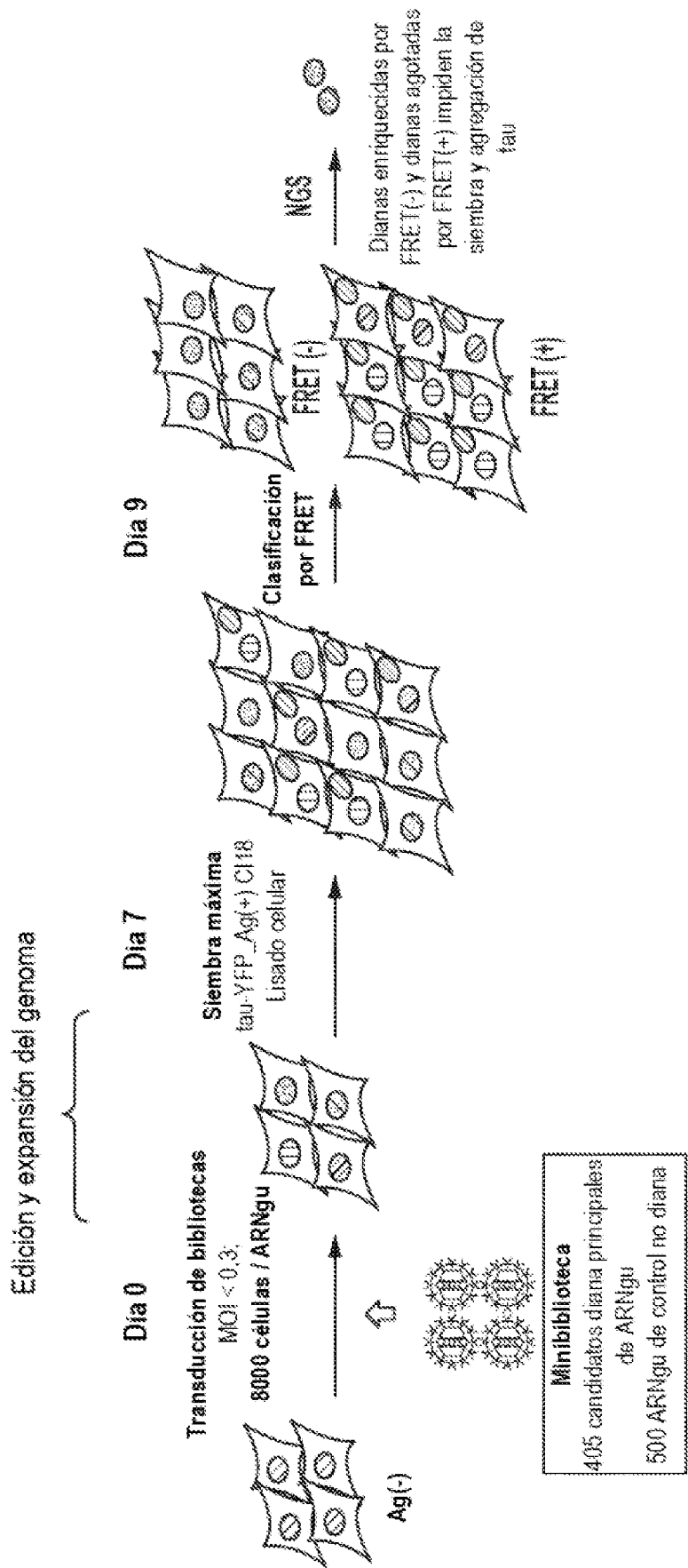


FIG. 25

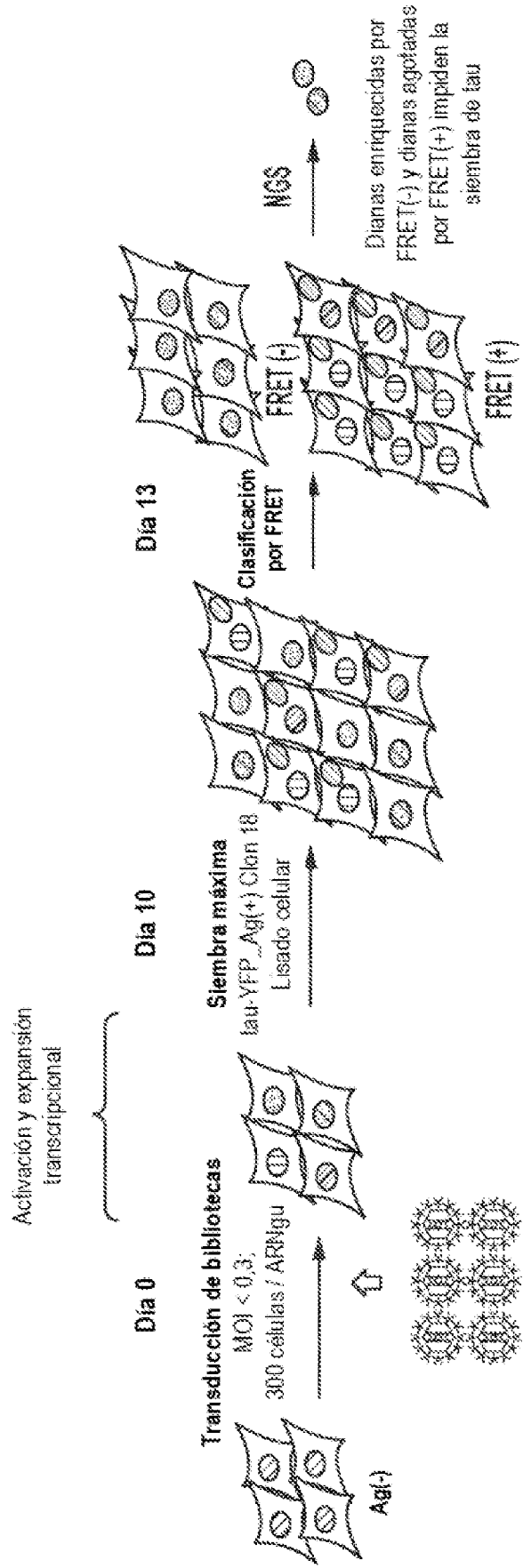


FIG. 26

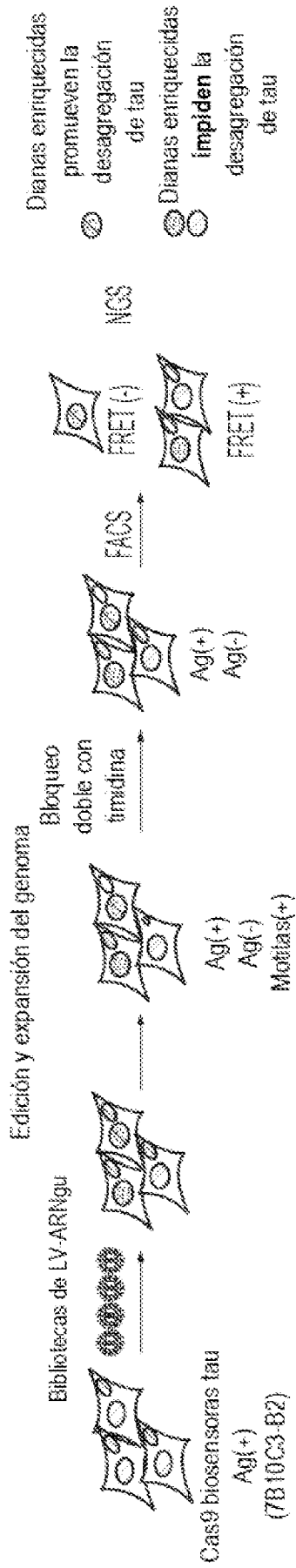


FIG. 27

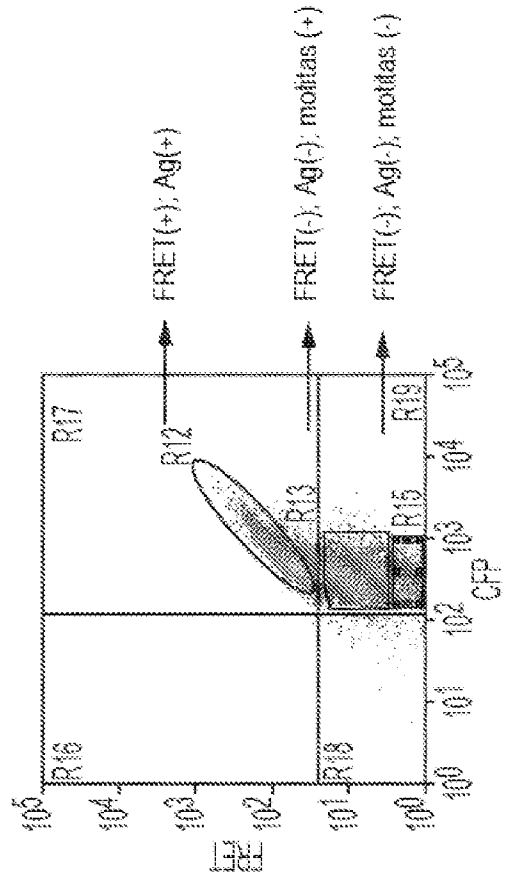


FIG. 28

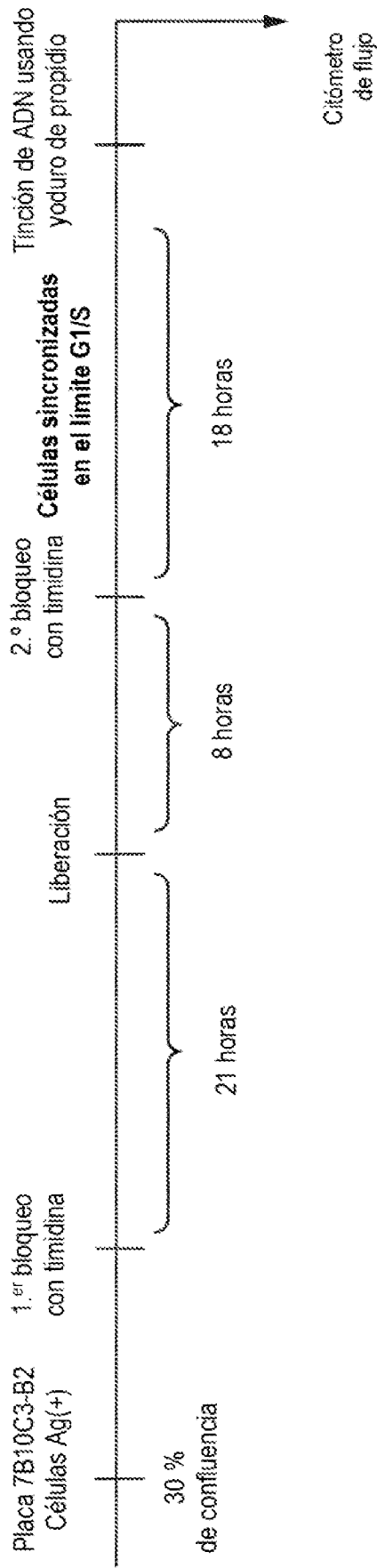


FIG. 29