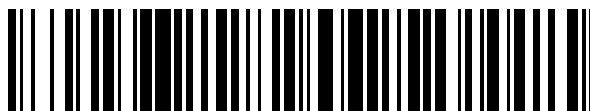


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 744 804**

51 Int. Cl.:

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.01.2014 PCT/EP2014/051204**

87 Fecha y número de publicación internacional: **31.07.2014 WO14114660**

96 Fecha de presentación y número de la solicitud europea: **22.01.2014 E 14701089 (6)**

97 Fecha y número de publicación de la concesión europea: **26.06.2019 EP 2948777**

54 Título: **Proteínas de repetición de dipéptidos como objetivo terapéutico en enfermedades neurodegenerativas con expansión de repetición de hexanucleótidos**

30 Prioridad:

**22.01.2013 WO PCT/EP2013/000190
20.09.2013 EP 13185419**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.02.2020

73 Titular/es:

**DEUTSCHES ZENTRUM FÜR
NEURODEGENERATIVE ERKRANKUNGEN E.V.
(100.0%)
Ludwig-Erhard Allee 2
53175 Bonn, DE**

72 Inventor/es:

**EDBAUER, DIETER;
HAASS, CHRISTIAN;
WENG, SHIH-MING;
MORI, KOHJI;
ARZTBERGER, THOMAS y
KREMMER, ELISABETH**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 744 804 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

Proteínas de repetición de dipéptidos como objetivo terapéutico en enfermedades neurodegenerativas con expansión de repetición de hexanucleótidos

5 La presente invención se refiere a anticuerpos y fragmentos de unión a antígeno de los mismos que se unen específicamente a un péptido que consiste en repeticiones de dipéptidos (Gly-Ala)_n, en las que n es un número entero de 16 o más en los que el anticuerpo o fragmento de unión a antígeno del mismo comprende las secuencias de aminoácidos HCDR1 a HCDR3 y LCDR1 a LCDR3 como se especifica en el presente documento.

Antecedentes de la invención.

10 Los trastornos neurodegenerativos se caracterizan generalmente por depósitos de proteínas que significan enfermedad. Además, en una serie de enfermedades neurodegenerativas, las mutaciones que causan variantes genéticamente heredadas de la enfermedad se asociaron con los genes que codifican los depósitos de proteínas, sus precursores o sus enzimas moduladoras. El análisis funcional de estas variantes genéticas ayudó fundamentalmente a comprender los mecanismos asociados a la enfermedad de la enfermedad de Alzheimer (EA) y la enfermedad de Parkinson (Gasser et al., 2011, Haass et al., 2007). La degeneración lobular frontotemporal (FTLD) y la esclerosis lateral amiotrófica (ALS) son los extremos de un espectro de trastornos neurodegenerativos superpuestos asociados de forma variable con demencia, cambios de personalidad, anormalidades del lenguaje y debilidad muscular progresiva (Josephs et al., 2011; Mackenzie et al., 2010; Rademakers et al., 2012). La investigación sobre ALS y FTLD se aceleró dramáticamente mediante la identificación de la proteína TDP-43 de unión a ARN/ADN (proteína de unión a ADN de Tar de 43 kDa) como una proteína depositada abundantemente (Arai et al., 2011; Neumann et al., 2006) y por el descubrimiento de que las mutaciones en TARDBP causan variantes familiares de ambas enfermedades (Benajiba et al., 2009; Sreedharan et al., 2008). La mayoría de los casos muestran inclusiones intracelulares que son fuertemente positivas para TDP-43 fosforilada. Estos hallazgos también ayudaron a desarrollar el concepto de que ALS y FTLD son trastornos multisistémicos con características clínicas y patológicas superpuestas y causas funcionales y genéticas similares (Rademakers et al., 2012; Sieben et al., 2012) y que, por lo tanto, se clasifican como FTLD-TDP, FTLD/ALS-TDP o ALS-TDP. Además del gen *TDP-43*, y el conocido *SOD1* (superóxido dismutasa 1), se descubrieron otros genes/factores de riesgo relacionados con ALS y/o FTLD, que incluyen los genes *FUS* (fusionado en sarcoma), *OPTN* (optineurina), *ataxina 2*, *Chmp2B*, *VCP* (proteína que contiene valosina), *TMEM106B*, *GRN* (Progranulina), *PFN* (Profilina) y *C9orf72*. Se han encontrado expansiones patológicas repetidas en *C9orf72* en aproximadamente el 40% de los pacientes con ALS familiar y en el 20% con FTLD familiar, lo que demuestra que *C9orf72* es la causa genética más común de estos trastornos incurables.

35 Recientemente, la expansión de una repetición de hexanucleótidos GGGGCC en el gen *C9orf72* se ha identificado como la mutación patógena más común en familias con FTLD, FTLD/ALS y ALS autosómico dominante (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Gijselink et al., 2012). Se reveló además que la expansión de repetición de hexanucleótidos dentro de la región reguladora del gen *C9orf72* es la causa más común de esclerosis lateral amiotrófica familiar y la segunda causa más común de degeneración lobular frontotemporal. La expansión de repetición de hexanucleótidos se encuentra secuencia arriba del marco de lectura abierto de *C9orf72*, ya sea en el primer intrón o en la región promotora, dependiendo de la isoforma de la transcripción (véase la Figura 8A). Aunque el contenido extremo de GC impide la secuenciación en pacientes, se cree que el número de unidades de repetición de GGGGCC es de al menos varios cientos en comparación con menos de 25 en controles sanos (van der Zee et al., 2012). Sin embargo, los patomecanismos que conducen a la enfermedad no fueron claros.

45 Los pacientes con una expansión de repetición de *C9orf72* tiene síntomas clínicos similares a otros pacientes con FTLD/ALS-TDP, pero muestran varias características patológicas únicas (Al-Sarraj et al., 2011; Boxer et al., 2011; Bigio et al., 2012; Whitwell et al., 2012). Los agregados de TDP-43 fosforilada están acompañados por abundantes inclusiones citoplasmáticas neuronales negativas para fosfo-TDP-43 similar a puntos y en forma de estrella, en particular en el cerebelo, hipocampo y neocorteza frontotemporal que solo pueden identificarse con anticuerpos para p62, ubiquitina o las ubiquilinas relacionadas. Estos agregados negativos para fosfo-TDP-43 son muy característicos de los portadores enfermos con la mutación *C9orf72* y están ausentes en otras variantes de FTLD/ALS-TDP. La identidad de la proteína o proteínas de la enfermedad en estas inclusiones y su relación con la expansión de repetición de hexanucleótidos de *C9orf72* se ha mantenido elusiva. A partir de la investigación sobre otras enfermedades neurodegenerativas con expansión de repetición fuera del marco de lectura abierto, se han propuesto dos patomecanismos principales:

55 Se cree que las expansiones repetidas en regiones reguladoras de genes no codificantes causan una enfermedad por dos mecanismos diferentes que no son mutuamente excluyentes. Primero, debido a la inmensa longitud de la transcripción y/o el empalme de la expansión de repetición, pueden verse afectados y producir haploinsuficiencia (van der Zee et al., 2012). En segundo lugar, la toxicidad por ARN causada por el secuestro de proteínas de unión a ARN también puede ser causante (Ranum et al., 2006). Actualmente, existe evidencia de ambas posibilidades. La observación de focos de ARN nuclear en pacientes con expansiones de repetición de hexanucleótidos GGGGCC, un hallazgo que, sin embargo, aún se discute de manera controvertida, sugiere que la captura de proteínas de unión a ARN esenciales puede estar involucrada en la enfermedad. Además, el hallazgo de una expresión disminuida del ARNm de *C9orf72* y una disminución de la actividad transcripcional del promotor de *C9orf72* en alelos intermedios (7-

24 repeticiones) (DeJesus-Hernandez et al., 2011; Gijssels et al., 2012; van der Zee et al., 2012) implica una pérdida de función como mecanismo causante de la enfermedad. Estos escenarios no son mutuamente excluyentes e incluso pueden ocurrir en paralelo.

5 No está claro cómo la expansión de repetición de *C9orf72* conduce a las inclusiones características positivas para p62/negativas para TDP-43 y la posterior neurodegeneración. Hay más de 150 artículos sobre genética y patología de *C9orf72*. Sin embargo, los datos funcionales siguen siendo muy limitados.

10 Se divulga en el presente documento por primera vez que la mayoría de estas inclusiones características contienen proteínas de repetición de dipéptidos (DPR) poli-(Gly-Ala) y, en menor medida, poli-(Gly-Pro) y poli-(Gly-Arg) que se generan por traducción no iniciada por ATG a partir de las repeticiones expandidas de GGGGCC en tres marcos de lectura. Estos hallazgos vinculan directamente la mutación genética asociada a FTL/ALS con la patología característica en pacientes con expansión de hexanucleótidos de *C9orf72*.

15 Hasta ahora solo hay evidencia de una traducción no iniciada por ATG de regiones de repetición exónicas en dos enfermedades (Zu et al., 2011). *ATXN8* codifica un tramo natural de poli-Q que puede causar inclusiones de poli-Q tras la expansión repetida en pacientes con ataxia espinocerebelosa tipo 8 (SCA8). La repetición de CAG expandida se traduce en los tres marcos de lectura (poli-Q, poli-A y poli-S) incluso después de la eliminación del codón de inicio endógeno. Poli-Q y poli-A se han encontrado en agregados de pacientes. Además, la distrofia miotónica tipo 1 (DM1) es causada por la expansión de CTG en las 3'UTR del gen *DMPK*. Se descubrió la traducción en agregados de poli-Q raros en pacientes con DM1 y modelos de ratón. El mecanismo subyacente se denominó traducción no iniciada por ATG asociada a repetición (RAN) y está patentado para trastornos de repetición de tri, tetra y pentanucleótidos (WO2010/115033 A9). La traducción RAN no se ha mostrado para repeticiones intrónicas o repeticiones de hexanucleótidos. En un artículo de comentario, Laura Ranum especula sobre la traducción RAN en pacientes con *C9orf72* (Ashizawa y Ranum 2012). También es posible que la traducción de las proteínas DPR se inicie desde el codón de inicio no canónico 5' de la región de repetición (Ivanov, I. P. et al., 2011; Peabody D. S. 1989; Touriol C. et al., 2003).

25 Se mostró por primera vez una traducción no iniciada por ATG de una expansión de repetición intrónica, que causa agregados positivos para p62/negativos para TDP43 de poli-GA, poli-GP y poli-GR. Este mecanismo de traducción inusual y el producto altamente anormal facilitarán enfoques terapéuticos más selectivos que los posibles para otras enfermedades neurodegenerativas. La inhibición de la generación o agregación anormal de proteínas de repetición de dipéptidos evitará o retrasará la progresión de la enfermedad en portadores de la mutación. La patología de DPR se puede utilizar para prevenir, retrasar o curar FTD/ALS en portadores de la mutación *C9orf72* desestabilizando específicamente el ARN de *C9ORF72* (intrónico) (por ejemplo, ARNpi, antisentido, que altera el empalme de *C9orf72*), inhibiendo la transcripción y/o traducción de la repetición en DPR (detección de inhibidores de traducción específicos de RAN) y prevención de la agregación o promoción de la degradación o eliminación de DPR (a través de compuestos químicos o inmunoterapia).

35 Es probable que se apliquen mecanismos similares para SCA36, causado por la expansión de hexanucleótidos GGCCCTG en *NOP56* (Kobayashi et al., 2011). Algunos de los marcos de lectura codifican las DPR como en *C9orf72*, poli-GP y poli-PR, pero con diferentes regiones flanqueantes. Por lo tanto, los fármacos terapéuticos descubiertos para pacientes con *C9orf72* también pueden ser útiles para pacientes con SCA36.

40 La agregación de repeticiones largas (Gly-Ala) no se conoce en otras proteínas humanas o virales. Sin embargo, el virus de Epstein-Barr codifica un antígeno rico en glicina y alanina. Dillner J. et al., (1984) PNAS 81: 4652 divulga anticuerpos contra un péptido sintético para unir el antígeno nuclear determinado por el virus de Epstein-Barr.

Sumario de la invención

45 En un primer aspecto, la presente invención proporciona un anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido que consiste en repeticiones del dipéptido (Gly-Ala)_a, en las que a es un número entero de 16 o más, en los que el anticuerpo o el fragmento de unión a antígeno del mismo comprende una secuencia de CDR3 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 53; una secuencia de CDR2 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 54; y una secuencia de CDR1 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 55.

50 y

(i) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende EILPGSGSTK de acuerdo con la SEQ ID NO: 64 y una secuencia de CDR3 que comprende GDFTNSHFAY de acuerdo con la SEQ ID NO: 60;

o

55 (ii) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGSTK de acuerdo con la SEQ ID NO: 65 y una

secuencia de CDR3 que comprende GDYSNSHFAY de acuerdo con la SEQ ID NO: 57;

o

(iii) una secuencia de CDR1 de cadena pesada que comprende GYKFIGYWIE de acuerdo con la SEQ ID NO: 69, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGTTK de acuerdo con la SEQ ID NO: 66 y una secuencia de CDR3 que comprende GDYSNSHFTY de acuerdo con la SEQ ID NO: 59.

En un segundo aspecto, la presente invención proporciona un producto farmacéutico que comprende el anticuerpo o fragmento de unión a antígeno del mismo según el primer aspecto de la invención y/o un ácido nucleico que codifica dicho anticuerpo o fragmento de unión a antígeno del mismo.

En un tercer aspecto, la presente invención proporciona el uso de un anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con el primer aspecto de la invención en el diagnóstico *in vitro* de una enfermedad caracterizada por una expansión de repeticiones genómicas de GGGGCC, en las que la enfermedad se selecciona del grupo que consiste en esclerosis lateral amiotrófica (ALS), demencia frontotemporal (FTD) y esclerosis lateral amiotrófica-demencia frontotemporal (ALS-FTD).

Listado de figuras

A continuación, se describirá el contenido de las figuras comprendidas en esta memoria descriptiva. En este contexto, consulte también la descripción detallada de la invención anterior y/o posterior.

Figura 1: Las repeticiones extendidas de GGGGCC se traducen en la agregación de proteínas de repetición de dipéptidos (DPR). (A) Validación de los anticuerpos específicos de DPR mediante inmunotransferencia con proteínas de fusión GST purificadas que contienen (GA)₁₅, (GP)₁₅ o (GR)₁₅. (B) Los constructos de repetición de GGGGCC con la longitud de repetición indicada que carecen de un ATG secuencia arriba se transfirieron en células HEK293. La digestión de restricción para estimar la longitud de la repetición de los constructos transfectados (panel superior). Las inmunotransferencias muestran la expresión dependiente de la longitud de proteínas poli-GA y poli-GP. Los productos de poli-GA fueron detectables a partir de ~38 repeticiones (flecha). El asterisco indica banda inespecífica. Los productos de Poli-GR no se detectaron (no se muestran). (C) Ensayo de trampa de filtro de cerebelo de paciente (véase Figura 2). Las fracciones insolubles en Triton-X100 se resuspendieron en SDS al 2%, se filtraron a través de membranas de acetato de celulosa y las proteínas retenidas se detectaron con los anticuerpos indicados. (D) La fracción insoluble en SDS de (C) se hirvió en regulador Lämmli 4x (que contenía 8% de SDS) y se analizó por inmunotransferencia. Las flechas marcan la parte superior del gel.

Figura 2: Sumario de hallazgos en casos de FTLD/ALS y controles. Sumario de los hallazgos inmunohistoquímicos para las inclusiones DPR, p62 y TDP-43 para todos los pacientes analizados. La patología de TDP-43 se examinó con anticuerpos fosfo-TDP-43 o pan-TDP-43. NE, no examinado. NA, no disponible.

Figura 3: Descripción general de la patología de DPR en el cerebelo. Inmunohistoquímica en la capa granular cerebelosa que compara la inclusión teñida con anticuerpos específicos p62, poli-GA, poli-GP y poli-GR en un paciente con FTLD/ALS con la mutación *C9orf72* (TJ-1). El número de agregados positivos para p62 (A) es similar al número de agregados de poli-GA (B), mientras que los agregados de poli-GP y poli-GR son menos frecuentes (C, D). En pacientes y controles, los anticuerpos poli-GR mostraron además una débil tinción nuclear y citoplasmática, probablemente debido a la reactividad cruzada con repeticiones cortas de Gly-Arg naturales presentes en varias proteínas de unión a ADN/ARN. Las barras de escala denotan 20 µm. Capa granular cerebelosa CBL-GL.

Figura 4: las proteínas de DPR forman las inclusiones negativas características de TDP-43 en pacientes con *C9orf72*. La inmunohistoquímica con anticuerpos específicos de DPR (GA, GP, GR) revela inclusiones de poli-GA, poli-GP y poli-GR que se asemejan a los agregados positivos para p62 en pacientes con FTLD/ALS con mutación *C9orf72* (compárese con la Figura 2). Inclusiones en forma de punto y en forma de hilo en la capa granular cerebelosa (CBL-GL) (A-E). Inclusiones citoplasmáticas en forma de estrella (F-I) e intranucleares en forma de punto (J) en las regiones 2 y 3 del cuerno de Amón del hipocampo (CA2 y CA3). Inclusión de la morfología mixta en la capa molecular cerebelosa (CBL-ML) y la capa granular de giro dentado (DG-GL) (K, L). Téngase en cuenta que en pacientes y controles, los anticuerpos poli-GR mostraron además una débil tinción nuclear y citoplasmática. Las barras de escala denotan 20 µm. La especificidad anti-GA y anti-GP se confirmó mediante experimentos de incubación previa con antígenos recombinantes (Figura 5 A-F). La validación de anti-GR solo fue posible mediante experimentos de inmunotransferencia (Figura 1A), porque el antígeno poli-GR se unía directamente al tejido (Figura 5 G-K). (M-O) La inmunofluorescencia doble revela la composición de los agregados de DPR en pacientes TJ1 con FTLD/ALS por *C9orf72*. No se observó colocalización de proteínas de DPR con fosfo-TDP-43 (véase la Figura 6). Las barras de escala denotan 10 µm.

Figura 5: La incubación previa con antígenos poli(Gly-Ala) y poli(Gly-Pro) confirma la especificidad de los anticuerpos anti-GA y anti-GP. Inmunohistoquímica en la región del hipocampo CA4 del caso TJ-1 de mutación *C9orf72* teñido con anticuerpos contra poli-GA, poli-GP y poli-GR incubados previamente con GST-(GA)₁₅, GST-(GP)₁₅, GST-(GR)₁₅ o GST (indicado en rosado). La incubación previa con GST-(GA)₁₅ y GST-(GP)₁₅ bloquea completamente las señales de tinción (comparar (A) y (B) para anti-GA y (D) y (E) para anti-GP). Sorprendentemente, GST-(GR)₁₅ mejoró

fuertemente las señales de poli-GR aparentes (comparar (G) y (H)). Esto se debió a la unión directa de la proteína GST-(GR)₁₅ cargada positivamente al tejido (presumiblemente a ADN, ARN y fosfolípidos), ya que un anticuerpo anti-GST mostró un patrón de tinción mejorado idéntico cuando se incubó previamente con GST-(GR)₁₅ (comparar el panel (J) y (K)). La incubación previa con GST no perjudicó la detección inmunohistoquímica de poli-GA, poli-GP y poli-GR (C, F, I). Las barras de escala denotan 20 μ m.

Figura 6: Los agregados de DPR se colocan con p62 pero no con TDP-43. La inmunofluorescencia del paciente TJ-1 con *C9orf72* con los anticuerpos indicados no revela una agregación conjunta de poli-GA y poli-GP con agregados positivos para fosfo-TDP-43 en el hipocampo (A-C). Sin embargo, similar a poli-GA (Figura 4M), poli-GP y poli-GR se localizaron conjuntamente con agregados positivos para p62 (D, E). El depósito de poli-GA se observó ocasionalmente dentro de agregados de fosfo-TDP-43, pero nunca al revés, lo que indica que el depósito de poli-GA precede a la agregación de TDP-43 (F). Colocalización de inclusiones positivas para p62 y DPR en el cerebelo y la región CA4 del hipocampo cuantificadas por inmunofluorescencia doble en tres pacientes con mutación *C9orf72* (TJ-1, TJ-2 y TJ-3). En el cerebelo, en total, se contabilizaron 370 a 822 (análisis de poli-GA), 123 a 566 (análisis de poli-GP) y 211 a 596 (análisis de poli-GR) por paciente. En CA4 en total se analizaron 48-73 inclusiones por paciente para cada especie de DPR (G). Las barras de escala denotan 10 μ m. CA3 y CA4, regiones 3 y 4 del cuerno de Amón del hipocampo; DG-GL, capa granular del giro dentado.

Figura 7: la patología de DPR es específica para pacientes con expansión de repetición de hexanucleótidos *C9orf72*. (A, B) La inmunohistoquímica con anticuerpos específicos de poli-GA (GA) no detecta agregados en un paciente con FTLT-DPR (TJ-13) sin expansión de repetición *C9orf72* y un caso con enfermedad de Huntington (TJ-11). Las inclusiones de fosfo-TDP-43 y Huntingtina son fácilmente detectables. Capa granular de giro dentado (DG-GL) y corteza frontal, respectivamente. (C) Inclusiones positivas para Poly-GA pero sin inclusiones de fosfo-TDP-43 en la corteza temporal del paciente TJ-10 con mutación *C9orf72* diagnosticada con FTLT-DPR. Las barras de escala denotan 20 μ m.

Figura 8: Regulación diferencial de transcripciones de *C9orf72* en portadores de la mutación. (A) Estructura genómica de *C9orf72* (dibujada no a escala). (B) El análisis por qPCR muestra una disminución de la expresión del ARNm de *C9orf72* maduro empalmado en el cerebelo de los casos de mutación *C9orf72*. Los cebadores del exón 2 a 3 detectan todas las isoformas *C9orf72*. Los cebadores del exón 1b a 2 detectan transcripciones alternativas que carecen de la expansión de repeticiones. (C) Expresión fuertemente aumentada de transcripciones sentido y antisentido que contienen el intrón 1 de *C9orf72*. Los cebadores utilizados para la transcripción inversa específica de cadena y qPCR se representan en (A). Los datos son la media \pm SEM. * denota $p < 0,05$, # denota $p < 0,058$ (prueba t de Student). Cerebelo de N = 3 pacientes con *C9orf72* y N = 6 controles incluyendo individuos sanos y pacientes con FTLT-DPR sin mutación *C9orf72* (como en la Figura 1C/D).

Figura 9: Identificación de proteínas de unión específica de la repetición de hexanucleótidos GGGGCC. (A) Geles representativos teñidos con plata que muestran proteínas extraídas por los respectivos ARN que contienen repeticiones. Los extractos nucleares de HEK293 se incubaron con las sondas de ARN indicadas con (+) o sin (-) un exceso de 50 veces de competidor de ARN no biotinilado. Las uniones de ARN/proteína se debilitaron con una concentración creciente de NaCl. En presencia del competidor GGGGCC (G+), se inhibió la unión de la proteína al ARN y aumentaron las proteínas en la fracción a través del flujo (FL). Los carriles encajonados en fracciones de elución de NaCl 500 mM se cortaron para la identificación de proteínas por LC-MS/MS. (B) Análisis de transferencia Western que confirma la unión específica de la repetición de GGGGCC de proteínas seleccionadas. Alícuotas de proteínas eluidas a diferentes concentraciones de sal fueron sometidas a electroforesis, y la transferencia Western se realizó usando los anticuerpos indicados. Todas las proteínas muestran la unión específica de repetición de GGGGCC a altas concentraciones de NaCl. Téngase en cuenta que la unión estaba completamente bloqueada por un exceso de 50 veces de GGGGCC no marcada (+). HnRNP F, que no es una de las 20 proteínas de unión específicas de repetición de GGGGCC, se usa como control negativo. A, repeticiones de AAAACC; FL, flujo a través; G, repeticiones de GGGGCC.

Figura 10: Listado de 20 proteínas seleccionadas que se unen específicamente a las repeticiones de GGGGCC. Se seleccionaron 20 proteínas como proteínas de unión a la repetición de GGGGCC específicas con base en los estrictos criterios descritos en el Ejemplo 9. El valor cuantitativo refleja la cantidad de proteína relativa estimada a partir de la intensidad de la señal de LC-MS/MS derivada de la elución 500mM usando el software Scaffold. En presencia de la repetición de GGGGCC no biotinilada el competidor (competición) de unión de todas las proteínas se suprime eficazmente. Además, estas proteínas muestran al menos 2 veces más unión a la repetición de ARN GGGGCC (GC) en comparación con la repetición de ARN AAAACC (AC). Los valores cuantitativos mostrados en el presente documento fueron los promedios de tres experimentos independientes.

Figura 11: Identificación de anticuerpos que se unen específicamente a la repetición de GGGGCC. (A) Un análisis representativo de transferencia Western que muestra la unión de anticuerpos específicos de GA seleccionados mediante inmunotransferencia con proteína de fusión de GST purificada que contiene la GA-DPR (GA)₁₅. Alícuotas de proteínas se sometieron a electroforesis, y la transferencia Western se realizó usando los anticuerpos indicados (GA 5E9, GA 5F2, GA 1A12). Las proteínas de fusión de GST que contienen GP-DPR, GR-DPR, AP-DPR, PR-DPR se usaron como control para la especificidad de unión a GA. (B) Experimentos de ELISA que demuestran una alta afinidad de los anticuerpos específicos de GA. Los anticuerpos se incubaron con cantidades crecientes de proteína de fusión

de GST (GA)₁₅ y se midió la absorbancia a una longitud de onda de 450 nm. El anticuerpo 2C8 de GST se usó como control. Para determinar la especificidad de unión de los anticuerpos GA seleccionados, se realizó un ELISA. Por lo tanto, las placas ELISA se recubrieron con anticuerpos anti-ratón de burro (adsorbidos contra Ig de rata, 100 ng/pozo) seguido por una etapa de bloqueo (1% de BSA en BBST). Las placas recubiertas previamente se incubaron con anticuerpos de captura anti-GA o anti-GST de ratón (100 µL de sobrenadante de hibridoma diluido 1/5) y luego se incubaron durante 1 hora a 37 °C con una serie de diluciones de GST-(GA)₁₅. El antígeno unido se detectó con anticuerpo anti-GST de rata y un anticuerpo secundario HRP anti-rata de burro (adsorbido contra 40 ng/pozo de Ig de ratón). Para el análisis cuantitativo, las placas se incubaron con sustrato TMB de HRP. La reacción de color se detuvo con HCl y se midió la absorbancia a 450 nm.

Figura 12: Secuencias de regiones variables de cadena ligera y regiones variables de cadena pesada de anticuerpos GA 1A12, GA 5E9 y GA 5F2. En este documento se muestran las secuencias de las regiones variables de cadenas ligeras (LCVR) y cadenas pesadas (HCVR). Las secuencias de las LCVR son idénticas, sin embargo, las secuencias de las HCVR difieren en los tres anticuerpos anti-GA diferentes (GA 1A12, GA 5E9, GA 5F2).

Las secuencias de las LCVR y HCVR pueden determinarse como comunes en la técnica. Por ejemplo, los datos se pueden obtener como se describe en <http://antibody.bath.ac.uk/align.html>. Las secuencias de las CDR dentro de la respectiva cadena pesada y ligera se indican con letras mayúsculas. Las regiones marco intermedias se indican en letras minúsculas.

Descripción detallada de la invención

Antes de que la presente invención se describa en detalle a continuación, debe entenderse que esta invención no se limita a la metodología, protocolos y reactivos particulares descritos en el presente documento, ya que pueden variar. También debe entenderse que la terminología utilizada en el presente documento tiene el propósito de describir realizaciones particulares solamente, y no pretende limitar el alcance de la presente invención, que estará limitado solo por las reivindicaciones adjuntas. A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en este documento tienen los mismos significados que los entendidos comúnmente por un experto en la materia.

Definiciones

A continuación, se proporcionan algunas definiciones de términos frecuentemente usados en esta memoria descriptiva. Estos términos tendrán, en cada caso de su uso, en el resto de la memoria descriptiva el significado definido respectivamente y los significados preferidos.

A lo largo de esta memoria descriptiva y las reivindicaciones que siguen, a menos que el contexto requiera lo contrario, la palabra "comprender", y variaciones tales como "comprende" y "que comprende", deben entenderse que implican la inclusión de un número entero o etapa establecido o grupo de números enteros o etapas, pero no la exclusión de ningún otro número entero o entero o grupo de números enteros o etapas.

Como se usa en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares "un", "uno, una" y "el, la" incluyen referencias al plural, a menos que el contenido indique claramente lo contrario.

Como se usa en el presente documento, un "individuo" significa cualquier mamífero, reptil o ave que pueda beneficiarse de la presente invención. Preferiblemente, se selecciona un individuo del grupo que consiste en animales de laboratorio (por ejemplo, ratón, rata o conejo), animales domésticos (incluyendo, por ejemplo, conejillo de indias, conejo, caballo, burro, vaca, oveja, cabra, cerdo, pollo, pato, camello, gato, perro, tortuga terrestre, tortuga de mar, serpiente o lagarto), o primates, incluidos chimpancés, bonobos, gorilas y seres humanos. Se prefiere particularmente que el "individuo" sea un ser humano.

Un "paciente" es cualquier receptor de servicios de atención médica. Por lo general, el paciente está enfermo o lesionado, o es susceptible a enfermedades o lesiones o está en riesgo de desarrollar una enfermedad y, por lo tanto, necesita tratamiento por parte de un médico, asistente médico, enfermera registrada de práctica avanzada, veterinario u otro proveedor de atención médica. Como se usa en este documento, un paciente significa cualquier mamífero, reptil o pájaro que pueda beneficiarse de la invención descrita en este documento. Preferiblemente, se selecciona un "paciente" del grupo que consiste en animales de laboratorio (por ejemplo, ratón o rata), animales domésticos (por ejemplo, conejillo de indias, conejo, caballo, burro, vaca, oveja, cabra, cerdo, pollo, camello, gato, perro, tortuga terrestre, tortuga de mar, serpiente o lagarto) o primates, incluidos chimpancés, bonobos, gorilas y seres humanos. Se prefiere particularmente que el "paciente" sea un ser humano.

El término "tejido", como se usa en el presente documento, se refiere a un conjunto de células del mismo origen que cumplen una función específica de manera concertada. Los ejemplos de un tejido incluyen pero no se limitan a tejido nervioso, tejido muscular, hueso, cartílago, tejido conectivo y tejido epitelial. Múltiples tejidos juntos forman un "órgano" para llevar a cabo una función específica. Los ejemplos de un órgano incluyen, pero no se limitan a, cerebro, músculo, corazón, sangre, esqueleto, articulación, hígado, riñón, estómago y piel.

El término "célula" como se usa en el presente documento puede referirse a una célula procariota (por ejemplo, una

célula bacteriana) o una célula eucariota (por ejemplo, una célula de hongo, vegetal o animal). Los organismos multicelulares comprenden varios tipos de células diferenciadas para cumplir diferentes funciones en dicho organismo. Estas incluyen, entre otras, células madre, células madre embrionarias, células del sistema nervioso, células sanguíneas, células del sistema inmunitario, células mesenquimales, células epiteliales, células intersticiales, células de metabolismo y almacenamiento, células de glándulas, células de matriz extracelular, células contráctiles, células pigmentarias, células germinales y células tumorales. El término "célula", como se usa en el presente documento, también se refiere a aquellas células que se eliminan de su entorno natural, tales como células primarias aisladas o líneas celulares de cualquiera de los tipos de células nombrados anteriormente. Típicamente, se usan células tales como células bacterianas, células de levadura, células primarias aisladas o líneas celulares en ensayos biotecnológicos. En el contexto de la presente invención, la célula primaria aislada o las líneas celulares son preferiblemente de origen mamífero.

Los términos "polinucleótido" y "ácido nucleico" se usan indistintamente en el presente documento. Las moléculas de ácido nucleico se entienden como una macromolécula polimérica u oligomérica hecha de monómeros de nucleótidos. Los monómeros de nucleótidos están compuestos de una nucleobase, un azúcar de cinco carbonos (como, entre otros, ribosa o 2'-desoxirribosa) y uno a tres grupos fosfato. Típicamente, un polinucleótido se forma a través de enlaces fosfodiéster entre los monómeros de nucleótidos individuales. En el contexto de la presente invención, las moléculas de ácido nucleico referidas incluyen, pero no se limitan a, ácido ribonucleico (ARN), ácido desoxirribonucleico (ADN) y mezclas de los mismos, como por ejemplo, híbridos de ARN-ADN, así como ADNc, ADN genómico, ADN recombinante, ARNc y ARNm. Un ácido nucleico puede consistir en un gen completo, o una porción del mismo, el ácido nucleico también puede ser un microARN (miARN) o un ARN pequeño de interferencia (ARNpi). Los miARN son moléculas cortas de ácido ribonucleico (ARN), en promedio solo 22 nucleótidos de longitud, que se encuentran en todas las células eucariotas. Los microARN (miARN) son reguladores postranscripcionales que se unen a secuencias complementarias en transcripciones de ARN mensajero (ARNm) objetivo, lo que generalmente produce represión traduccional y silenciamiento génico. Los ARN pequeños de interferencia (ARNpi), a veces conocidos como ARN interferente corto o ARN silenciador, son ácidos ribonucleicos cortos (moléculas de ARN), con una longitud de entre 20 y 25 nucleótidos. Están involucrados en la vía de ARN de interferencia (ARNi), en la que interfieren con la expresión de genes específicos. El ácido nucleico también puede ser un ácido nucleico artificial. Los ácidos nucleicos artificiales incluyen poliamida o ácido nucleico peptídico (PNA), morfolino y ácido nucleico bloqueado (LNA), así como ácido nucleico glicólico (GNA) y ácido nucleico de treosa (TNA). Cada uno de estos se distingue del ADN o ARN natural por cambios en la cadena principal de la molécula.

Los métodos adecuados para detectar ácidos nucleicos tales como ADN y/o ARN incluyen, por ejemplo, análisis de transferencia Northern, análisis de transferencia Southern, ensayos de protección de nucleasa (NPA), hibridación *in situ* (ISH), reacción en cadena de la polimerasa (PCR para detección de ADN) y reacción en cadena de la polimerasa de transcripción inversa (RT-PCR para detección de ARN).

El término "marco de lectura abierto" (ORF) se refiere a una secuencia de nucleótidos, que puede traducirse en aminoácidos. Típicamente, dicho ORF contiene un codón de inicio, una región posterior que generalmente tiene una longitud que es un múltiplo de 3 nucleótidos, pero no contiene un codón de parada (TAG, TAA, TGA, UAG, UAA o UGA) en el marco de lectura dado. Por lo general, los ORF se producen de forma natural o se construyen artificialmente, es decir, por medio de tecnología genética. Un ORF codifica un péptido, polipéptido o proteína donde los aminoácidos en los que puede traducirse forman una cadena unida a péptidos.

Los términos "expresión génica" o "expresión" se usan indistintamente en el presente documento y se refieren al proceso mediante el cual se usa la información genética para sintetizar un producto génico funcional. Típicamente, dicho producto génico es un péptido, polipéptido o proteína, o un ácido nucleico tal como un ARN ribosómico (ARNr), ARN de transferencia (ARNt) o ARN nuclear pequeño (ARNnp). La expresión génica incluye los pasos de transcripción, empalme de ARN, traducción y modificación postraduccional. Preferiblemente, el término se usa para referirse a la síntesis de un péptido, polipéptido o proteína. Por lo tanto, el término "detección de expresión" se usa preferiblemente para referirse a la detección de expresión de un péptido, polipéptido o proteína. Dicha detección puede llevarse a cabo mediante métodos conocidos en la técnica, en particular mediante el uso de ligandos que se unen específicamente al péptido, polipéptido o proteína.

Normalmente, en la síntesis de proteínas, una secuencia de ADN que codifica un gen se transcribe primero en ARNm del que se eliminan los intrones mediante empalme de ARN y los exones se unen y posteriormente se traducen para producir una cadena de aminoácidos, que luego se pliega en una proteína.

El término "transcripción" se refiere al proceso en el que un segmento particular de ADN, típicamente un gen, se transcribe en ARN mediante la enzima ARN polimerasa. Durante la transcripción, una ARN polimerasa lee una secuencia de ADN, que produce una cadena de ARN antiparalela complementaria. A diferencia de la replicación de ADN, la transcripción da como resultado un complemento de ARN que incluye uracilo (U) en todos los casos en los que la timina (T) estaría presente en un complemento de ADN. Si el gen transcrito codifica una proteína, el resultado de la transcripción es un ARN mensajero previo (ARNm) o una molécula de ARNm, que luego se traducirá en un péptido, polipéptido o proteína. Alternativamente, el gen transcrito puede codificar genes de ARN no codificantes (tal como microARN, ARNncil, etc.) o ARN ribosómico (ARNr) o ARN de transferencia (ARNt), otros componentes del proceso de ensamblaje de proteínas u otras ribozimas. El "empalme de ARN" ocurre simultáneamente o después del

proceso de transcripción y se refiere al proceso en el que los "intrones" comprendidos en el pre-ARNm se eliminan y los "exones" se unen covalentemente.

El término "intrón" se refiere a cualquier secuencia de nucleótidos dentro de un gen que se elimina mediante empalme de ARN. En la técnica, el término intrón se usa típicamente para referirse tanto a la secuencia de ADN dentro de un gen como a la secuencia correspondiente en el transcrito de ARN que se elimina mediante empalme de ARN. Las secuencias que se unen en el ARN maduro final después del empalme de ARN se denominan "exones". Nuevamente, el término exón se usa típicamente en la técnica para referirse tanto a las secuencias de ADN dentro de un gen como a las secuencias correspondientes en el transcrito de ARN que se unen durante el empalme de ARN después de la eliminación del intrón.

Los ribosomas facilitan el proceso de "traducción" de ARNm en una cadena de aminoácidos al inducir la unión de los ARNt con secuencias anticodónicas complementarias a la del ARNm. Los ARNt transportan aminoácidos específicos que se encadenan en un polipéptido a medida que el ARNm pasa y el ribosoma lo "lee". Típicamente, la traducción es un proceso dependiente de AUG, en el que un codón AUG del ARNm (correspondiente a un codón ATG del ADN) se reconoce como sitio de inicio de la traducción que da como resultado que la metionina sea el primer aminoácido en la cadena de aminoácidos producida. Sin embargo, también existe un mecanismo de traducción independiente de AUG, por ejemplo, en el que el ARNt de metionina interactúa con un codón complementario a solo dos nucleótidos. En los virus, el mecanismo de traducción independiente de AUG incluye el uso de sitios internos de entrada de ribosomas (IRES) que imitan estructuralmente el ARN iniciador y manipulan los ribosomas para iniciar la traducción en un sitio que no es AUG.

El término "traducción RAN" se refiere específicamente a un mecanismo de traducción independiente de AUG de repeticiones de nucleótidos. En la técnica anterior se ha demostrado que la traducción RAN se produce para ARN exónico que comprende repeticiones de trinucleótidos, tetranucleótidos y pentanucleótidos. En el trabajo que condujo a la presente invención, se demostró sorprendentemente que también el ARN intrónico transcrito a partir de repeticiones de hexanucleótidos, presentes en el genoma, podría traducirse mediante traducción RAN. Los polipéptidos que fueron traducidos por RAN de tales repeticiones de nucleótidos pueden (pero no necesariamente tienen que) diferir de los péptidos o polipéptidos que se tradujeron de una manera dependiente de AUG en que carecen de la metionina inicial. Por ejemplo, la traducción RAN puede facilitarse porque la presencia de dichas repeticiones de nucleótidos promueve la formación de una estructura de horquilla que posteriormente desencadena la traducción RAN.

Los términos "cadena de aminoácidos" y "cadena de polipéptidos" se usan como sinónimos en el contexto de la presente invención.

En el contexto de la presente invención, el término "péptido" se refiere a un polímero corto de aminoácidos unidos por enlaces peptídicos. Tiene los mismos enlaces químicos (péptidos) que las proteínas, pero generalmente es de longitud más corta. El péptido más corto es un "dipéptido", que consta de dos aminoácidos unidos por un único enlace peptídico. También puede haber un tripéptido, tetrapéptido, pentapéptido, etc. El péptido también puede tener una longitud de hasta 8, 10, 12, 15, 18 o 19 aminoácidos. Un péptido tiene un extremo amino y un extremo carboxilo, a menos que sea un péptido cíclico. En el contexto de la presente invención, el término "repetición de dipéptido (DPR)" se refiere a un dipéptido de dos aminoácidos unidos por un enlace peptídico único que se duplica varias veces para formar un péptido o polipéptido más largo, es decir, una repetición de dipéptido puede comprender 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000 o más dipéptidos unidos por enlaces peptídicos. Preferiblemente, las repeticiones de dipéptidos están en el intervalo de 500 a 10.000 más preferiblemente en el intervalo de 700 a 4500. Las "repeticiones de dipéptidos" son los productos de traducción de las repeticiones de hexanucleótidos. El significado de este término se describe adicionalmente más adelante. Requiere, entre otras cosas, que la unidad de repetición más pequeña tenga una longitud de seis nucleótidos, por ejemplo, en la secuencia de nucleótidos "CGCGCGCGCG" (SEQ ID NO: 83) la unidad de repetición más pequeña es "CG" y, por lo tanto, esta secuencia es una repetición de dinucleótidos. Por otro lado, en la secuencia de nucleótidos GGGCCCGGCC (SEQ ID N: 84) la unidad de repetición más pequeña es "GGGCC" y, por lo tanto, esta es una repetición de hexanucleótidos. Un hexanucleótido que no comprende un codón de PARADA codificará así un dipéptido y una serie de repeticiones de dichos hexanucleótidos codificará una repetición de dipéptido. Debido a la degeneración del código genético, es posible que una secuencia de nucleótidos que cumpla los criterios de ser una repetición de hexanucleótidos codifique un dipéptido de aminoácidos idénticos, por ejemplo, la unidad de repetición más pequeña de la secuencia de nucleótidos "GGTGGCGGTGGC" (SEQ ID NO: 85) es "GGTGGC", que codifica Gly-Gly. Sin embargo, se prefiere que el dipéptido comprenda dos aminoácidos diferentes, por ejemplo, (Gly-Ala), (Gly-Pro), (Gly-Arg), (Ala-Pro), (Pro-Arg), (Gly-Leu), (Ala-Trp), (Pro-Gly), (Ala-Gln). De lo contrario, es decir, si los aminoácidos de un dipéptido pudieran ser idénticos, no podría determinarse para un péptido de seis aminoácidos de longitud que hay tres repeticiones de dipéptidos, pero el experto consideraría que dicha secuencia es un hexapéptido o una repetición de seis veces de un monómero. En consecuencia, el término "repetición de dipéptido (DPR)" se refiere a un dipéptido de dos aminoácidos diferentes unidos por un enlace peptídico único que se duplica varias veces para formar un péptido o un polipéptido más largo, es decir, una repetición de dipéptido puede comprender 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000 o más dipéptidos unidos por enlaces peptídicos.

El término "polipéptido" se refiere a una cadena lineal única de aminoácidos unidos por enlaces peptídicos y preferiblemente comprende al menos aproximadamente 20 aminoácidos. Un polipéptido puede ser una cadena de una proteína que está compuesta de más de una cadena, o puede ser la proteína en sí misma si la proteína está compuesta de una sola cadena.

5 El término "proteína" se refiere a una molécula que comprende uno o más polipéptidos que reanudan una estructura secundaria y terciaria y además se refiere a una proteína que está compuesta por varios polipéptidos, es decir, varias subunidades, formando estructuras cuaternarias. La proteína a veces tiene grupos no peptídicos unidos, que pueden llamarse grupos prostéticos o cofactores.

10 Un "péptido aislado", "polipéptido aislado" o "proteína aislada" se refiere a un péptido, polipéptido o proteína que se ha removido de su entorno natural en una célula de manera que otro material celular normalmente cercano ya no está presente. En el contexto de la presente invención, también péptidos, polipéptidos o proteínas producidos fuera de su entorno celular natural, por ejemplo, por medios químicos o por medios recombinantes en un entorno no natural, se consideran péptidos, polipéptidos o proteínas aislados.

15 Los polipéptidos o proteínas (que incluyen derivados de proteínas, variantes de proteínas, fragmentos de proteínas, segmentos de proteínas, epítomos de proteínas y dominios de proteínas) pueden modificarse adicionalmente mediante modificación química. Por lo tanto, un polipéptido modificado químicamente puede comprender grupos químicos distintos de los residuos encontrados en los 20 aminoácidos naturales. Los ejemplos de tales otros grupos químicos incluyen, sin limitación, aminoácidos glicosilados y aminoácidos fosforilados. Las modificaciones químicas de un polipéptido pueden proporcionar propiedades ventajosas en comparación con el polipéptido parental, por ejemplo, uno
20 o más de mayor estabilidad, mayor vida media biológica o mayor solubilidad en agua. Las modificaciones químicas incluyen sin limitación: PEGilación, glicosilación de polipéptidos parentales no glicosilados. Dichas modificaciones químicas aplicables a las variantes utilizables en la presente invención pueden ocurrir con la traducción o posterior a la traducción.

25 Una "proteína antigénica" como se menciona en la presente solicitud es un polipéptido como se definió anteriormente que contiene al menos un epítomo. Un "fragmento antigénico" de una proteína antigénica es una secuencia parcial de dicha proteína antigénica que comprende al menos un epítomo. Para fines de inmunización, solo aquellas partes de una proteína que son relevantes provocan una respuesta inmune. Por lo tanto, el constructo de ácido nucleico no necesita codificar la proteína antigénica de longitud completa como se encuentra en, por ejemplo, una célula enferma, una célula cancerosa o un patógeno. Un fragmento acortado de dicha proteína es suficiente siempre que su secuencia
30 de aminoácidos comprenda el epítomo o epítomos responsables del reconocimiento de la proteína antigénica por el sistema inmunitario. El término "antígeno", como se usa en el presente documento, se refiere a cualquier molécula o parte de una molécula, que incluye pero no se limita a ácido nucleico, aminoácido, péptido, polipéptido, proteína, carbohidrato y lípido, al que se une un ligando de la invención.

35 El término "epítomo" como se usa en el presente documento se refiere a un determinante antigénico, que es parte de un antígeno que está unido específicamente por un ligando de la invención, preferiblemente un anticuerpo o fragmento de unión a antígeno del mismo. Los epítomos generalmente consisten en agrupaciones de moléculas químicamente activas en la superficie, tales como aminoácidos o cadenas laterales de azúcar, y generalmente tienen características estructurales tridimensionales específicas, así como características de carga específicas. Los epítomos de un antígeno pueden ser un epítomo conformacional o un epítomo no conformacional, es decir, un epítomo lineal. Los epítomos conformacionales y no conformacionales se distinguen porque la unión al primero pero no al último se pierde en presencia de disolventes desnaturalizantes. Un epítomo conformacional se compone de secciones discontinuas de la secuencia de aminoácidos del antígeno. Estos epítomos interactúan con el ligando en función de las características de la superficie tridimensional y la forma o estructura terciaria del antígeno. La mayoría de los epítomos son conformacionales. Por el contrario, los epítomos lineales interactúan con el ligando en función de su estructura primaria.
40 Un epítomo lineal está formado por una secuencia continua de aminoácidos del antígeno. Los epítomos conformacionales comprenden preferiblemente entre 8 y 20 aminoácidos discontinuos, preferiblemente entre 8 y 15 aminoácidos. Los epítomos lineales tienen una longitud de entre 6 y 20 aminoácidos, más preferiblemente entre 8 y 15 aminoácidos.

45 Los péptidos, polipéptidos o proteínas pueden detectarse a través de diversos métodos que incluyen, pero no se limitan a, un ensayo de trampa de filtro, transferencia Western, ensayo por inmunoabsorción ligado a enzimas (ELISA), inmunohistoquímica (IHC), inmunocitoquímica (ICC) y cromatografía de exclusión por tamaño. (SEC).

50 La "inmunotinción" que incluye pero no se limita a inmunohistoquímica (IHC) o inmunocitoquímica (ICC), es un método basado en anticuerpos para detectar una proteína específica en una muestra. El término inmunotinción se usó originalmente para referirse a la tinción inmunohistoquímica de secciones de tejido. Ahora, sin embargo, la inmunotinción abarca una amplia gama de técnicas utilizadas en histología, biología celular y biología molecular que utilizan métodos de tinción basados en anticuerpos. Mientras que los primeros casos de tinción con IHC usaban tintes fluorescentes, otros métodos no fluorescentes que usan enzimas como la peroxidasa y la fosfatasa alcalina ahora se usan con más frecuencia. Estas enzimas son capaces de catalizar reacciones que producen un producto coloreado que es fácilmente detectable por microscopía óptica. Alternativamente, los elementos radiactivos pueden usarse como marcadores, y las inmunorreacciones pueden visualizarse mediante autorradiografía. La preparación o fijación de
60

tejidos es esencial para la preservación de la morfología celular y la arquitectura de los tejidos. La fijación inadecuada o prolongada puede disminuir significativamente la capacidad de unión del anticuerpo. Muchos antígenos pueden demostrarse con éxito en secciones de tejido embebidas en parafina fijadas con formalina. La optimización de los métodos y tiempos de fijación, el pretratamiento con agentes bloqueantes, la incubación de anticuerpos con alto contenido de sal y la optimización de los reguladores de lavado y los tiempos de lavado posteriores al anticuerpo pueden ser importantes para obtener inmunotinción de alta calidad.

La "cromatografía de exclusión por tamaño (SEC)" es un método cromático que permite la separación de moléculas en solución, por ejemplo, proteínas, por su tamaño y, en algunos casos, por peso molecular. Típicamente, SEC se aplica a moléculas grandes o complejos macromoleculares como polipéptidos o proteínas. La "cromatografía de filtración en gel" se refiere típicamente a una SEC en la que se usa una solución acuosa para transportar la muestra a través de la columna, mientras que la "cromatografía de permeación en gel" se refiere a la SEC en la que se usa un disolvente orgánico como fase móvil.

El término "ensayo de retención de filtro" y "ensayo de trampa de filtro" se usan simultáneamente en el presente documento para referirse al análisis de agregados de proteínas. Las proteínas purificadas se filtran a través de filtros de acetato de celulosa que dan como resultado grandes agregados retenidos en el filtro. Posteriormente, estos agregados unidos a la membrana pueden caracterizarse por cualquier medio o método conocido en la técnica.

La "transferencia Western" permite la detección de proteínas específicas (nativas o desnaturalizadas) a partir de extractos hechos de células o tejidos, antes o después de cualquier etapa de purificación. Las proteínas generalmente se separan por tamaño usando electroforesis en gel antes de ser transferidas a una membrana sintética (típicamente nitrocelulosa o PVDF) a través de métodos de transferencia seca, semiseca o húmeda. La membrana se puede sondear usando anticuerpos usando métodos similares a la inmunohistoquímica, pero sin necesidad de fijación. La detección se realiza típicamente usando anticuerpos unidos a peroxidasa para catalizar una reacción quimioluminiscente. La transferencia Western es un método de rutina de biología molecular que se puede utilizar para comparar semicuantitativamente o cuantitativamente los niveles de proteínas entre extractos. La separación del tamaño antes de la transferencia permite medir el peso molecular de la proteína en comparación con los marcadores de peso molecular conocidos. La transferencia Western es una técnica analítica utilizada para detectar proteínas específicas en una muestra dada de homogeneizado o extracto de tejido. Utiliza electroforesis en gel para separar proteínas por la longitud del polipéptido (condiciones desnaturalizantes) o por la estructura tridimensional de la proteína (condiciones nativas/no desnaturalizantes).

El "ensayo por inmunoabsorción ligado a enzimas (ELISA)" es un método de diagnóstico para determinar cuantitativa o semicuantitativamente las concentraciones de proteínas a partir de plasma sanguíneo, suero o extractos de células/tejidos en un formato de placa de múltiples pozos (generalmente 96 pozos por placa). En términos generales, las proteínas en solución se adsorben en placas de ELISA. Los anticuerpos que se unen específicamente a la proteína de interés se usan para sondear la placa.

La "microscopía electrónica (EM)" puede usarse para estudiar la microarquitectura detallada de tejidos o células. "Inmuno-EM" permite la detección de proteínas específicas en secciones de tejido ultradelgadas. Los anticuerpos marcados con partículas de metales pesados (por ejemplo, oro) se pueden visualizar directamente mediante microscopía electrónica de transmisión.

Un "marcador", "etiqueta" o "rótulo" es cualquier tipo de sustancia que puede indicar la presencia de otra sustancia o complejo de sustancias. El marcador puede ser una sustancia que está vinculada o introducida en la sustancia a detectar. Los marcadores detectables se utilizan en biología molecular y biotecnología para detectar, por ejemplo, una proteína, un producto de una reacción enzimática, un segundo mensajero, ADN, interacciones de moléculas, etc. Los ejemplos de marcadores o etiquetas adecuados incluyen un fluoróforo, un cromóforo, un radiomarcador, un coloide metálico, una enzima o una molécula quimioluminiscente o bioluminiscente. Los ejemplos de fluoróforos incluyen varias formas de proteína fluorescente verde (GFP) tales como EnGFP, RFP, CYP, BFP, YFP, dsRed, etc., ficobiliproteínas (aloficocianina, ficocianina, ficoeritrina y ficoeritrocianina), fluoresceína (isotiocianato de fluoresceína, FITC), rodamina (isotiocianato de tetrametil rodamina, TRITC) y colorantes de cianina (tales como C2, Cy3, Cy5, Cy7). Los ejemplos de radiomarcadores incluyen ^3H , ^{14}C , ^{32}P , ^{33}P , ^{35}S , $^{99\text{m}}\text{Tc}$ o ^{125}I . Los ejemplos de enzimas incluyen luciferasa, beta-galactosidasa, peroxidasa de rábano picante, fosfatasa alcalina, glucosa oxidasa y ureasa.

Como se usa en el presente documento, el término "marcador de selección" se refiere a un gen indicador que transmite un cierto rasgo a una célula u organismo que permite que dicha célula u organismo sobreviva en un entorno en el que no sobreviviría sin dicho gen indicador. En consecuencia, la expresión de un marcador de selección permite la selección artificial de células u organismos cuyo genoma comprende dicho gen informador. Los marcadores de selección incluyen, pero no se limitan a, genes marcadores auxotróficos y genes marcadores de resistencia que confieren resistencia a una toxina o un antibiótico. Los términos "marcador de selección", "gen indicador" y "gen marcador" se usan indistintamente en el presente documento. Los marcadores de selección pueden ser genes marcadores auxotróficos o genes marcadores de resistencia que confieren resistencia a una toxina o un antibiótico. El "gen marcador de auxotrofia" se entiende como un gen que permite el crecimiento de tales células u organismos auxotróficos en condiciones de cultivo selectivas mediante la codificación de una molécula requerida para la síntesis de un producto necesario para la supervivencia de un procarionte auxotrófico en un medio selectivo utilizado para cultivo

celular. Cualquier gen marcador que sea capaz de complementar el gen correspondiente que no es funcional en una célula u organismo auxotrófico puede ser útil en el contexto de esta invención. Como se usa en el presente documento, los términos "genes marcadores de resistencia que confieren resistencia a una toxina y/o un antibiótico", o variaciones de los mismos como "genes marcadores de resistencia a la toxina" o "genes de resistencia a antibióticos" se refieren a genes informadores que codifican una enzima que inactiva funcionalmente una toxina o un antibiótico. La inactivación funcional de una toxina o antibiótico se puede lograr mediante la expresión de un gen marcador que porta una mutación o mutaciones que hacen que el producto génico respectivo sea insensible a una toxina o antibiótico. Alternativamente, la inactivación funcional de una toxina o antibiótico se puede lograr mediante la expresión de un gen marcador que inhibe la toxina o antibiótico, por ejemplo, al interactuar o unirse a él. La inactivación funcional de una toxina o antibiótico también puede lograrse mediante la expresión de un gen marcador que contrarresta los efectos de la toxina o antibiótico. Los compuestos antibióticos incluyen, pero no se limitan a, tetraciclinas, sulfonamidas, penicilinas, cefalosporinas, ansamicinas, carbapenémicos, macrólidos, quinolonas, aminonucleósidos, aminoglicósidos, péptidos, glicopéptidos y lipopéptidos. A modo de ejemplo, la higromicina B, la neomicina, la kanamicina, la gentamicina y el G418 (también conocido como Geneticina) son antibióticos aminoglicósidos de estructura similar. En general, la neomicina y la kanamicina se usan para los procariotas, mientras que G418 es necesaria para los eucariotas.

Se pueden conjugar diferentes tipos de rótulos o etiquetas químicas con anticuerpos secundarios o primarios y otras moléculas para facilitar su visualización (es decir, detección y medición) por diversos métodos. Los radioisótopos se usaron ampliamente en el pasado, pero son costosos, tienen una vida útil corta, no ofrecen mejoras en la relación señal/ruido y requieren un manejo y eliminación especiales. Las enzimas y los fluoróforos han reemplazado en gran medida a los isótopos radiactivos como etiquetas detectables para los ensayos. Una serie de avances en reactivos e instrumentación hacen que estas nuevas tecnologías sean más versátiles y potentes. Las etiquetas enzimáticas como la peroxidasa de rábano picante (HRP) se utilizan con mayor frecuencia para los métodos de transferencia, inmunoensayos e inmunohistoquímica. Las etiquetas fluorescentes se usan predominantemente para obtención de imágenes celulares, amplificación y secuenciación de ácidos nucleicos y microarreglos; sin embargo, la tecnología de fluorescencia se está desarrollando rápidamente para su aplicación en todo tipo de ensayos.

El término "nivel de expresión" se refiere a la cantidad de producto génico (por ejemplo, DPR) presente en el cuerpo o una muestra en un determinado punto de tiempo. El nivel de expresión puede, por ejemplo, medirse/cuantificarse/detectarse por medio de las cantidades de la proteína o del ARNm que codifica la proteína. Por ejemplo, el nivel de expresión puede cuantificarse normalizando la cantidad de producto génico de interés (por ejemplo, DPR) presente en una muestra con la cantidad total de producto génico de la misma categoría (proteína total o ARNm) en la misma muestra o en una muestra de referencia (por ejemplo, una muestra tomada al mismo tiempo del mismo individuo o una parte del mismo tamaño (peso, volumen) de la misma muestra) o identificando la cantidad de producto génico de interés por tamaño de muestra definido (peso, volumen, etc.). El nivel de expresión puede medirse/cuantificarse/detectarse mediante cualquier método conocido en la técnica, por ejemplo, métodos para la detección y cuantificación directa del producto génico de interés (como la espectrometría de masas) o métodos para la detección y medición indirecta del producto génico de interés que generalmente funcionan mediante la unión del producto génico de interés con una o más moléculas diferentes o medios de detección (por ejemplo, cebador o cebadores, sondas, anticuerpos, proteínas de andamio) específicos para el producto génico de interés (por ejemplo, DPR). Preferiblemente, el nivel de expresión se determina con base en la proteína más que con base en el ARNm.

El término "toxicidad" como se usa en el presente documento se refiere al grado en que un compuesto/sustancia puede dañar un organismo o una subestructura del organismo, tal como una célula (citotoxicidad), tejido o un órgano. En consecuencia, el término "efecto tóxico" se refiere al efecto perjudicial que un compuesto/sustancia tiene sobre un organismo, órgano, tejido o célula. Un compuesto puede exhibir un efecto tóxico porque daña la función y/o estructura de un organismo, órgano, tejido o célula, lo que puede provocar una función alterada o una pérdida de la función de ciertos elementos o partes del organismo, órgano, tejido o célula, o incluso puede provocar la muerte de dicho organismo, órgano, tejido o célula. El término "compuesto tóxico", por lo tanto, se refiere a una sustancia, por ejemplo, un ácido nucleico, un péptido, polipéptido o proteína, o una sustancia o compuesto químico, que exhibe un efecto tóxico en el organismo, órgano, tejido o célula.

La toxicidad o el efecto tóxico de un compuesto se puede medir usando uno de varios ensayos de viabilidad conocidos en la técnica que incluyen, pero no se limitan a, ensayos basados en formazán (MTT/XTT), ensayo de lactato deshidrogenasa (LDH), prueba de ATP, Calceína AM, ensayo clonogénico, ensayo de homodímero de etidio, azul Evans, hidrólisis de diacetato de fluoresceína/tinción con yoduro de propidio (tinción FDA/PI), citometría de flujo, ensayo TUNEL, con proteína fluorescente verde (GFP), violeta de metilo, yoduro de propidio, azul de tripano, o resazurina. Además, las tinciones de ADN pueden usarse para diferenciar entre células necróticas, apoptóticas y normales.

El término "enfermedad" y "trastorno" se usan indistintamente en el presente documento, haciendo referencia a una afección anormal, especialmente una afección médica anormal tal como una enfermedad o lesión, en la que una célula, un tejido, un órgano o un individuo no es capaz de seguir cumpliendo eficientemente su función. Típicamente, pero no necesariamente, una enfermedad está asociada con síntomas o signos específicos que indican la presencia de dicha enfermedad. La presencia de tales síntomas o signos puede, por lo tanto, ser indicativos de una célula, un tejido, un órgano o un individuo que padece una enfermedad. Una alteración de estos síntomas o signos puede ser indicativa de la progresión de dicha enfermedad. La progresión de una enfermedad se caracteriza típicamente por un

aumento o disminución de tales síntomas o signos que pueden indicar un "empeoramiento" o "mejoría" de la enfermedad. El "empeoramiento" de una enfermedad se caracteriza por una capacidad decreciente de una célula, tejido, órgano o individuo/paciente para cumplir su función de manera eficiente, mientras que el "mejoramiento" de una enfermedad se caracteriza típicamente por un aumento en la capacidad de una célula, tejido, órgano o individuo/paciente para cumplir su función de manera eficiente. Una célula, un tejido, un órgano o un individuo que es "susceptible" a una enfermedad se encuentra en un estado saludable pero especialmente vulnerable a la aparición de una enfermedad, por ejemplo, debido a predisposición genética, falta de vacunación, inmunidad poco desarrollada o inmadura, mal estado nutricional o similares. El brote de la enfermedad aún puede prevenirse mediante profilaxis o tratamiento preventivo. Se puede sospechar que una célula, un tejido, un órgano o un individuo tiene una enfermedad en la que dicha célula, tejido, órgano o individuo típicamente muestra signos o síntomas tempranos o débiles de dicha enfermedad. En tal caso, el inicio de la enfermedad aún puede prevenirse o su progresión puede reducirse o prevenirse mediante el tratamiento.

Como se usa en el presente documento, "detectar", "que detecta" o "detección" de una enfermedad o trastorno se refiere al establecimiento de la presencia o ausencia de una enfermedad en un paciente. Por ejemplo, una fracción utilizada en la detección de una enfermedad puede identificar la presencia o ausencia de un indicador de una enfermedad en una muestra o en un individuo o paciente. Por ejemplo, una enfermedad puede detectarse por medio de un ligando o un ligando marcado que interactúa, es decir, se une o forma un complejo con un ácido nucleico o péptido, polipéptido o proteína específico de la enfermedad. Una enfermedad también puede detectarse por medio de un inhibidor que bloquea el mecanismo de acción subyacente a la enfermedad y, por lo tanto, altera los síntomas de modo que se pueda identificar la enfermedad subyacente.

Como se usa en el presente documento, "tratar", "que tratar", "tratamiento" o "terapia" de una enfermedad o trastorno significa lograr uno o más de los siguientes: (a) reducir la gravedad del trastorno; (b) limitar o prevenir el desarrollo de síntomas característicos de los trastornos que se tratan; (c) inhibir el empeoramiento de los síntomas característicos de los trastornos que se tratan; (d) limitar o prevenir la recurrencia del trastorno o trastornos en un individuo que previamente ha tenido el trastorno o trastornos; y (e) limitar o prevenir la recurrencia de síntomas en individuos que anteriormente eran sintomáticos para el trastorno o trastornos. En consecuencia, una terapia trata una enfermedad o trastorno, o los síntomas de una enfermedad o trastorno mediante el logro de uno o más de los efectos mencionados anteriormente (a) - (e). Por ejemplo, una enfermedad puede tratarse mediante la inhibición/bloqueo del mecanismo de acción subyacente a la enfermedad, por ejemplo, a través de compuestos que inhiben la expresión de polipéptidos/proteínas específicos de la enfermedad o mediante la inhibición de procesos adicionales en los que está involucrado dicho polipéptido/proteína. Una enfermedad también puede tratarse activando el sistema inmunitario del paciente, por ejemplo, mediante inmunización activa, o apoyando el sistema inmunitario del paciente, por ejemplo, por inmunización pasiva.

Dos o más proteínas antigénicas o fragmentos antigénicos de las mismas son "inmunológicamente idénticos" si son reconocidos por el mismo anticuerpo, célula T o célula B. El reconocimiento de dos o más polipéptidos inmunogénicos por el mismo anticuerpo, célula T o célula B también se conoce como "reactividad cruzada" de dicho anticuerpo, célula T o célula B. Preferiblemente, el reconocimiento de dos o más polipéptidos inmunológicamente idénticos por el mismo anticuerpo, célula T o célula B se debe a la presencia de epítopos idénticos o similares en todos los polipéptidos. Epítopos similares comparten suficientes características estructurales y/o de carga para unirse por la región Fab del mismo anticuerpo o receptor de células B o por la región V del mismo receptor de células T. Las características de unión de un anticuerpo, receptor de células T o receptor de células B se definen, preferiblemente, por la afinidad de unión del receptor al epítipo en cuestión. Dos polipéptidos inmunogénicos son "inmunológicamente idénticos" como se entiende por la presente solicitud si la constante de afinidad del polipéptido con la constante de menor afinidad es al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95% o al menos 98% de la constante de afinidad del polipéptido con la constante de mayor afinidad. Los métodos para determinar la afinidad de unión de un polipéptido a un receptor, como la diálisis de equilibrio o el ensayo inmunoabsorbente ligado a enzimas (ELISA) son bien conocidos en la técnica. Preferiblemente, dos o más polipéptidos "inmunológicamente idénticos" comprenden al menos un epítipo idéntico. Los efectos de vacunación más fuertes generalmente se pueden obtener, si los polipéptidos inmunogénicos comprenden epítopos idénticos o si tienen una secuencia de aminoácidos idéntica. El término "vacuna" se refiere a una preparación biológica, típicamente una farmacéutica, que mejora la inmunidad a una enfermedad específica. Dicha preparación puede comprender uno o más inmunógenos específicos de enfermedad adecuados para provocar una respuesta inmune. En el contexto de la presente invención, dicho compuesto puede ser un polipéptido que es sustancialmente idéntico o inmunológicamente idéntico a un polipéptido como se especifica a continuación que comprende la repetición de dipéptido (DPR) especificada. Alternativamente, la vacuna puede comprender una construcción de ácido nucleico que codifica un polipéptido inmunogénico que es sustancialmente idéntico o inmunológicamente idéntico a un polipéptido que comprende la repetición de dipéptido especificada. En el último caso, se prefiere que el polipéptido se exprese en el individuo tratado con la vacuna. En el contexto de la presente invención, dicha construcción de ácido nucleico puede ser un vector.

Como se usa en el presente documento, el término "vector" se refiere a una proteína o un polinucleótido o una mezcla de los mismos que es capaz de introducirse o de introducir las proteínas y/o el ácido nucleico contenidos en la célula. Además, el término "vector" se refiere a al menos un polinucleótido formulado con una preparación de liposomas o nanopartículas lipídicas que es capaz de transfectar una célula con al menos un polinucleótido como se describe, por

ejemplo, por Geall et al., 2012. Además del polinucleótido que codifica el gen de interés, se pueden introducir polinucleótidos y/o polipéptidos adicionales en la célula. La adición de polinucleótidos y/o polipéptidos adicionales se prefiere especialmente si dichos polinucleótidos y/o polipéptidos adicionales son necesarios para introducir el constructo de ácido nucleico en la célula o si la introducción de polinucleótidos y/o polipéptidos adicionales aumenta la expresión del polipéptido inmunogénico codificado por el constructo de ácido nucleico de la presente invención.

En el contexto de la presente invención, se prefiere que los genes de interés codificados por el polinucleótido introducido se expresen dentro de la célula tras la introducción del vector o vectores. Los ejemplos de vectores adecuados incluyen, pero sin limitación, plásmidos, cósmidos, fagos, virus o cromosomas artificiales.

Los ejemplos de una enfermedad incluyen, pero no se limitan a, trastornos neurológicos, enfermedades inflamatorias, enfermedades infecciosas, afecciones cutáneas, enfermedades endocrinas, enfermedades intestinales, trastornos genéticos, enfermedades autoinmunes, enfermedades traumáticas, enfermedades articulares y varios tipos de cáncer.

"Trastornos neurológicos" se refiere a cualquier trastorno del sistema nervioso en el que se producen anomalías estructurales, bioquímicas o eléctricas en el cerebro, la médula espinal u otros nervios que afectan una gama de síntomas que incluyen, pero no se limitan a, parálisis, debilidad muscular, problemas de coordinación, pérdida de sensibilidad, convulsiones, confusión, dolor y niveles alterados de conciencia. Los ejemplos de trastornos neurológicos incluyen, pero no se limitan a, daño del cerebro o partes individuales del cerebro (por ejemplo, daño de la corteza prefrontal, lóbulo frontal, lóbulo parietal, lóbulo temporal, lóbulo occipital, cerebelo, hipocampo, tronco encefálico, sistema límbico), disfunción del cerebro o partes individuales del cerebro (por ejemplo, afasia, disartria, apraxia, agnosia, amnesia, ataxia) o inflamación del cerebro (por ejemplo, encefalitis, encefalitis viral, trombosis del seno cavernoso, absceso cerebral, amebiana). Las "enfermedades neurodegenerativas" como el Alzheimer, la enfermedad de Parkinson, la enfermedad de Huntington, la esclerosis lateral amiotrófica (ALS), la demencia frontotemporal (FTD), que también se llama degeneración lobular frontotemporal (FTLD), esclerosis lateral amiotrófica-demencia frontotemporal (ALS-FTD, también llamada ALS-FTLD), y la ataxia espinocerebelosa son otros ejemplos de trastornos neurológicos. Los términos clínicos y patológicos FTD y FTLD (y ALS-FTD y ALS-FTLD) se usan como sinónimos en esta solicitud. Los trastornos neurológicos también incluyen trastornos de la médula espinal (por ejemplo, siringomielia, siringobulbia, síndrome de Morvan, mielopatía vascular, síndrome de Foix-Alajouanine, compresión de la médula espinal) e inflamación de la médula espinal (por ejemplo, mielitis, poliomielitis, enfermedad desmielinizante, mielitis transversa, paraparesia espástica tropical, absceso epidural), neuropatía central y/o periférica, trastornos del nervio craneal (por ejemplo, neuralgia del trigémino), trastornos del movimiento del sistema nervioso central y/o periférico (por ejemplo, enfermedad de Parkinson, ALS, síndrome de Tourette, esclerosis múltiple), trastornos del sueño (por ejemplo, insomnio, hipersomnia), apnea del sueño, narcolepsia, cataplejía, Kleine-Levin, trastorno del sueño del ritmo circadiano, trastorno avanzado de la fase del sueño, trastorno de la fase del sueño retardado), dolor de cabeza (por ejemplo, migraña, cefalea en racimos, tensión), enfermedades neuropsiquiátricas, delirio, demencia (por ejemplo, enfermedad de Alzheimer, demencia vascular, FTD, demencia semántica y demencia con cuerpos de Lewy), accidente cerebrovascular (por ejemplo, MCA, ACA, PCA, de Foville, Millard-Gubler, medular lateral, de Weber, accidente cerebrovascular lacunar), tumores (por ejemplo, gliomas, meningiomas, adenomas hipofisarios, tumores de la vaina nerviosa), síndrome de dolor regional complejo y enfermedades de las neuronas motoras (MND) (por ejemplo, ALS, esclerosis lateral primaria (PLS), atrofia muscular progresiva (PMA), parálisis bulbar progresiva (PBP), parálisis pseudobulbar).

Los "síntomas" de una enfermedad son implicaciones de la enfermedad notable por una célula, tejido, órgano o individuo que tiene dicha enfermedad e incluyen, pero no se limitan a, dolor, debilidad, sensibilidad, tensión, rigidez y espasmo de la célula, tejido, un órgano o un individuo. Los "signos" o "señales" de una enfermedad incluyen, pero no se limitan a, el cambio o la alteración, tal como la presencia, ausencia, aumento o elevación, disminución o declive, de indicadores específicos, tal como biomarcadores o marcadores moleculares, o el desarrollo, la presencia o empeoramiento de los síntomas.

Los términos "indicador" o "biomarcador" se usan indistintamente en el presente documento. En el contexto de la presente invención, un "indicador" puede definirse como una sustancia dentro de un sistema biológico que se usa como un indicador de un estado biológico de dicho sistema. En la técnica, el término "biomarcador" a veces también se aplica a medios para la detección de dichas sustancias endógenas (por ejemplo, anticuerpos, sondas de ácido nucleico, sistemas de formación de imágenes). Sin embargo, en el contexto de la presente invención, el término "biomarcador" solo se aplicará a la sustancia, no a los medios de detección. Por lo tanto, los biomarcadores pueden ser cualquier tipo de molécula presente en un organismo vivo, tal como un ácido nucleico (ADN, ARNm, miARN, ARNr, etc.), una proteína (receptor de la superficie celular, proteína citosólica, etc.), un metabolito u hormona (azúcar en la sangre, insulina, estrógenos, etc.), una molécula característica de una cierta modificación de otra molécula (por ejemplo, fracciones de azúcar o fracciones de fosforilo en proteínas, fracciones de metilo en el ADN genómico, expansión de repeticiones de nucleótidos) o una sustancia que se ha internalizado por el organismo o un metabolito de dicha sustancia. Por consiguiente, una enfermedad o trastorno puede caracterizarse por la presencia o ausencia, aumento o disminución de dicho indicador. Dicho indicador de una enfermedad puede o no causar la enfermedad.

Los indicadores de la presencia y/o progresión de una enfermedad incluyen "marcadores genéticos" tales como repetición de nucleótidos, VNTR (repetición en tándem de número variable; por ejemplo, STR (repetición en tándem corta), AFLP (polimorfismo de longitud de fragmento amplificado), SSR (repetición de secuencia simple), MLVA, SSLP

(polimorfismo de longitud de secuencia simple), RFLP (polimorfismo de longitud de fragmento de restricción), RAPD (amplificación aleatoria de ADN polimórfico), SNP (polimorfismo de un solo nucleótido), SFP (polimorfismo de una sola característica), DArT (tecnología de matrices de diversidad), marcadores de RAD (marcadores de ADN asociados al sitio de restricción).

- 5 El término "repetición de nucleótidos" se refiere a una ubicación en el genoma en la que una secuencia de nucleótidos corta forma secuencias repetitivas de 2, 3, 4, 5 o 6 nucleótidos, es decir, dinucleótidos, trinucleótidos, tetranucleótidos, pentanucleótidos y hexanucleótidos. En un estado saludable, el número de repeticiones de la repetición puede variar entre 1 y 50, es decir, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50. Sin embargo, en un estado de enfermedad, estas repeticiones pueden expandirse para repetirse con mucha más frecuencia que en un estado
- 10 saludable, por ejemplo, pueden expandirse de 30 a 40 veces, lo que lleva a varios cientos de repeticiones de la repetición, es decir, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000, 5.500, 6.000, 6.500, 7.000, 7.500, 8.000, 8.500, 9.000, 9.500, o 10.000, o incluso más repeticiones. En consecuencia, el nivel de expansión de dinucleótidos, trinucleótidos, tetranucleótidos, pentanucleótidos o hexanucleótidos puede ser un
- 15 indicador (y/o la causa) de una enfermedad, tal como por ejemplo, una enfermedad neurológica, en particular una enfermedad neurodegenerativa. Por ejemplo, la expansión de las repeticiones de trinucleótidos (CAG)_n y del complemento inverso (CTG)_n en la cadena complementaria está asociada con la enfermedad de Huntington, la expansión de repeticiones de pentanucleótidos (ATTCT)_n y del complemento inverso (AGAAT)_n en la cadena complementaria está asociada con la ataxia espinocerebelosa SCA10, mientras que la expansión de los hexanucleótidos (GGGGCC)_n y el complemento inverso (CCCCGG)_n en la cadena complementaria está asociada con ALS, FTD y FTD-ALS y la expansión de los hexanucleótidos (GGCCTG)_n del complemento inverso (CAGGCC)_n en la cadena complementaria se asocia con ataxia espinocerebelosa SCA36. En principio, las repeticiones de nucleótidos pueden estar presentes en el exón o el intrón de un gen y pueden o no transcribirse y pueden o no traducirse en un péptido o polipéptido.
- 20 Las enfermedades neurodegenerativas ALS, FTD y ALS-FTD se caracterizan porque las repeticiones de hexanucleótidos (GGGGCC)_n localizadas dentro del intrón I del gen C9orf72 se expanden. En un estado saludable, el número "n" de las repeticiones de GGGGCC puede variar entre 1 y 19, es decir, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 y 19. En un estado de enfermedad el número "n" de estas repeticiones puede expandirse hasta repetirse con mucha más frecuencia que en un estado saludable, por ejemplo, típicamente pueden expandirse mucho más de 30 a 40 veces.
- 25 Sin embargo, algunos pacientes ya pueden mostrar síntomas a partir de 20 repeticiones. Por lo tanto, en un estado de enfermedad, la expansión de las repeticiones de hexanucleótidos puede conducir a varios cientos de repeticiones de la repetición de hexanucleótidos, es decir, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000 y 10.000 repeticiones, o más. El número exacto de repeticiones no puede determinarse con las técnicas actuales de secuenciación. La transcripción de las repeticiones de GGGGCC y de las repeticiones de GGCCCC genómicas en la cadena de ADN complementaria (también denominada cadena antisentido) conduce a productos de transcripción de ARN que comprenden o bien repeticiones de GGGGCC o de GGCCCC con los números indicados anteriormente. Sorprendentemente, la traducción de estos transcritos se inicia a partir de codones no ATG. El transcrito sentido se traduce en los tres marcos de lectura (poly-(Gly-Ala), poly-(Gly-Pro) y poly-(Gly-Arg)). Usando transcripción inversa específica de la cadena, se detectaron tanto transcritos sentido como antisentido (Figura 8C). Por lo tanto, existen 6 marcos de lectura, pero dos dan como resultado DPR idénticas, a saber, DPR poli-(Gly-Pro). Por lo tanto, las cinco DPR posibles diferentes son las siguientes: poli-(Gly-Ala), poli-(Gly-Pro), poli-(Gly-Arg), poli-(Ala-Pro) y poli-(Pro-Arg). En particular, los polipéptidos poli-GA, poli-GP y poli-AP son altamente hidrófobos y forman agregados en un entorno intracelular.
- 30
- 35
- 40
- 45 La secuencia genómica de la región que rodea la repetición del hexanucleótido del gen C9orf72 se lee como sigue en la secuencia de referencia de la NCBI (SEQ ID NO: 1)

acgtaacctacgggtgtcccgttaggaaagagaggtgctcaaacagcgacaagtccgccacgtaaaagatgacgcttg
 ggtgtcagccgtccctgctgcccgggtgcttctcttttggggggcggggcttagcaagagcaggtgtgggttaggaggtgtgtttttg
 ttttccaccctctctcccactacttctctcacagtactcgtgaggggtgaacaagaaaagacctgataaagattaaccagaagaaa
 caaggagggaacaaccgcagcctgtagcaagctctggaactcaggagtcgcgcgcta[GGGGCC]_nggggcgtggtcgggg
 cggggccccggggggcggggccggggcggggctgcggttgcggtgctgccccgcgggcggggaggcgcagggcgggtggcgagt
 ggggtgagtgaggaggcggcatcctggcgggtggctgtttgggggtcggctgccgggaagaggcggggtagaagcgggggctctc
 ctcagagctcgacgcatttttactttccctctcatttctgaccgaagctgggtgtcgggctttcgcctctagcgactggtggaattgctg
 catccggggccccgggcttccggcgggcggcgggcgggcgaggacaagggatggggatctggccttctcttctgcttt
 cccgcctcagtaccgagctgtctcttccggggaccgctggggagcgtgccgctgcgggctcgagaaaaggagcctcgggt
 actgagaggcctcgcctgggggaaggccggagggtggcgggcgcgggcttctgcgaccaagtcgggggtcgttaggaacccg
 agacgggtccctgccggcgaggagatcatgcgggg,

en la que "n" tiene los significados preferidos y más preferibles expuestos anteriormente.

El término "polipéptidos que comprenden o consisten en repeticiones de dipéptidos codificadas por repeticiones genómicas de hexanucleótidos" se refiere a polipéptidos que comprenden repeticiones de dipéptidos, preferiblemente de 9 o más de dipéptidos idénticos consecutivos. Se prefiere más que los polipéptidos comprendan más de 10, más de 15, más de 20, más de 25, más de 30, más de 35, más de 40, más de 45, más de 50 repeticiones de dipéptidos. En realizaciones preferidas, el polipéptido comprende o consiste en repeticiones de dipéptidos con dos aminoácidos diferentes, preferiblemente con una secuencia seleccionada del grupo que consiste en (Gly-Ala)_a, (Gly-Pro)_b, (Gly-Arg)_c, (Ala-Pro)_d, (Pro-Arg)_e, (Gly-Leu)_f, (Ala-Trp)_g, (Pro-Gly)_h, (Ala-Gln)_i, (Gly-Pro)_k, y (Pro-Arg)_l en las que a es un número entero de 16 o más, preferiblemente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; b es un número entero de 28 o más, preferiblemente 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; c es un número entero de 21 o más, preferiblemente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; d es un número entero de 17 o más, preferiblemente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; e es un número entero de 24 o más, preferiblemente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; f es un número entero de 15 o más, preferiblemente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; g es un número entero de 9 o más, preferiblemente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; h es un número entero de 28 o más, preferiblemente 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; j es un número entero de 16 o más, preferiblemente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; k es un número entero de 28 o más, preferiblemente 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más; y l es un número entero de 24 o más, preferiblemente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.500, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más.

El término "agregados" como se usa en el presente documento se refiere a la acumulación intra o extracelular de péptidos, polipéptidos o proteínas.

Los términos "muestra" o "muestra de interés" se usan indistintamente en el presente documento, haciendo referencia a una parte o pieza de un tejido, órgano o individuo, que típicamente es más pequeño que dicho tejido, órgano o individuo, destinado a representar el total del tejido, órgano o individuo. Tras el análisis, una muestra proporciona información sobre el estado del tejido o el estado de salud o enfermedad de un órgano o individuo. Los ejemplos de muestras incluyen, pero no se limitan a, muestras de líquidos como sangre, suero, plasma, líquido cefalorraquídeo y otros líquidos corporales (orina, saliva), así como muestras de biopsia o muestras sólidas, como extractos de tejido del cerebro o la médula espinal. Otros ejemplos de muestras incluyen cultivos celulares (por ejemplo, linfoblastos derivados del paciente) o cultivos de tejidos.

El análisis de una muestra se puede realizar de forma visual o química. El análisis visual incluye, entre otros, imágenes microscópicas o exploración radiográfica de un tejido, órgano o individuo que permite la evaluación de una muestra. El análisis químico incluye, pero no se limita a la detección de la presencia o ausencia de indicadores específicos o la detección de las alteraciones en su cantidad o nivel.

- 5 El término "ligando" como se usa en el presente documento se refiere a cualquier sustancia o compuesto que puede interactuar específicamente con, por ejemplo, para unirse específicamente o formar un complejo con la molécula especificada, por ejemplo, un polipéptido de la invención que comprende o que consiste en una repetición de dipéptido.

El término "inhibidor" se refiere a una sustancia, por ejemplo, un ligando, que bloquea la acción de otro compuesto, es decir, una molécula receptora. Típicamente, los inhibidores actúan uniéndose al sitio activo de la molécula del receptor, o interactuando con sitios de unión únicos que normalmente no están involucrados en la regulación de la actividad de la molécula del receptor. La actividad del inhibidor puede ser reversible o irreversible dependiendo de la longevidad de la interacción del complejo molécula de inhibidor-receptor. Los ejemplos de inhibidores incluyen, pero no se limitan a, moléculas de ácido nucleico, como ARNpi o miARN, o proteínas como factores de transcripción, moléculas de inmunoglobulina, anticuerpos, proteínas similares a anticuerpos, peptidomiméticos, hormonas, citoquinas, factores de crecimiento o neurotransmisores.

El término "inmunoglobulina (Ig)" como se usa en el presente documento se refiere a las glucoproteínas que confieren inmunidad de la superfamilia de inmunoglobulinas. Las "inmunoglobulinas de superficie" están unidas a la membrana de las células efectoras por su región transmembrana y abarcan moléculas tales como, pero sin limitarse a, receptores de células B, receptores de células T, proteínas del complejo mayor de histocompatibilidad (MHC) clase I y II, microglobulina beta-2 (B2M), CD3, CD4 y CD8. Típicamente, el término "anticuerpo" como se usa en el presente documento se refiere a inmunoglobulinas secretadas que carecen de la región transmembrana y, por lo tanto, pueden liberarse en el torrente sanguíneo y las cavidades corporales. Los anticuerpos se agrupan en diferentes isotipos en función de la cadena pesada que poseen. Hay cinco tipos de cadenas pesadas de Ig humana designadas por las letras griegas: α , δ , ϵ , γ y μ . El tipo de cadena pesada presente define la clase de anticuerpo, es decir, estas cadenas se encuentran en los anticuerpos IgA, IgD, IgE, IgG e IgM, respectivamente, cada uno de los cuales desempeña diferentes funciones y dirige la respuesta inmune apropiada contra diferentes tipos de antígenos. Las cadenas pesadas distintas difieren en tamaño y composición; α y γ comprenden aproximadamente 450 aminoácidos, mientras que μ y ϵ tienen aproximadamente 550 aminoácidos (Janeway et al., (2001) Immunobiology, Garland Science). Los anticuerpos comprenden cuatro cadenas de polipéptidos, a saber, dos cadenas pesadas (H) y dos cadenas ligeras (L) interconectadas por enlaces disulfuro. Cada cadena pesada está compuesta por una región variable de la cadena pesada (abreviada aquí como HCVR o VH) y una región constante de la cadena pesada. La región constante de la cadena pesada está compuesta por tres dominios, CH1, CH2 y CH3. Cada cadena ligera está compuesta por una región variable de la cadena ligera (abreviada aquí como LCVR o VL) y una región constante de la cadena ligera. La región constante de la cadena ligera se compone de un dominio, CL. Las regiones VH y VL se pueden subdividir en regiones de hipervariabilidad, denominadas regiones determinantes de complementariedad (CDR), intercaladas con regiones que están más conservadas, denominadas regiones marco (FR). Cada VH y VL está compuesto por tres CDR y cuatro FR, dispuestas desde el terminal amino hasta el terminal carboxilo en el siguiente orden: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. Las CDR para las cadenas pesadas y ligeras se pueden determinar como se conoce en la técnica. Por ejemplo, el siguiente conjunto de reglas puede usarse para encontrar las CDR dentro de una secuencia de cadena ligera y pesada de anticuerpo, respectivamente:

CDR-1 de cadena ligera: Inicio: aproximadamente 24 residuos, residuo antes del CDR-1 siempre un Cys, residuo después del CDR1 siempre un Trp. Típicamente Trp-Tyr-Gln, pero también, Trp-Leu-Gln, Trp-Phe-Gln, Trp-Tyr-Leu; Longitud: 10 a 17 residuos.

CDR-2: de cadena ligera: Inicio: siempre 16 residuos después del final de L1, residuos antes generalmente Ile-Tyr, pero también, Val-Tyr, Ile-Lys, Ile-Phe, longitud siempre 7 residuos.

CDR-3 de cadena ligera: Inicio: siempre 33 residuos después del final de CDR-2; residuo antes siempre Cys, residuos después siempre Phe-Gly-XXX-Gly, longitud: 7 a 11 residuos.

CDR-1 de cadena pesada: Inicio: aproximadamente 26 residuos siempre 4 después de un Cys (con base en la definición de AbM de Chothia, la definición de Kabat comienza con 5 residuos más tarde); residuos antes siempre Cys-XXX-XXX-XXX; residuos después siempre un Trp. Típicamente Trp-Val, pero también, Trp-Ile, Trp-Ala, longitud: 10 a 12 residuos [definición de AbM, la definición Chothia excluye los últimos 4 residuos).

CDR-2 de cadena pesada: Inicio: siempre 15 residuos después del final de la definición de Kabat/AbM, de CDR-1 de cadena pesada; residuos antes: típicamente Leu-Glu-Trp-Ile-Gly, pero una serie de variaciones, residuos después de Lys/Arg-Leu/Ile/Val/Phe/Thr/Ala-Thr/Ser/Ile/Ala, definición de Kabat de longitud de 16 a 19 residuos (definición de acuerdo con AbM; la definición de Chothia termina 7 residuos antes)

CDR-3 de cadena pesada: Inicio: siempre 33 residuos después del final de CDR-2 de cadena pesada (siempre 2 residuos de aminoácidos después de un Cys); residuos antes siempre Cys-XXX-XXX (típicamente Cys-Ala-Arg); residuos después siempre Trp-Gly-XXX-Gly; Longitud: de 3 a 25 residuos. Este conjunto de reglas es conocido por la

persona experta y también se puede encontrar en <http://www.bioinf.org.uk/abs/#cdrid>.

El término "anticuerpo humano", como se usa en el presente documento, pretende incluir anticuerpos que tienen regiones variables y constantes derivadas de secuencias de inmunoglobulina de línea germinal humana. Los mAb humana (por ejemplo, mutaciones introducidas por mutagénesis aleatoria o específica del sitio *in vitro* o por mutación somática *in vivo*), por ejemplo en las CDR. Sin embargo, el término "anticuerpo humano", como se usa en el presente documento, no pretende incluir "anticuerpos humanizados" en los que las secuencias de CDR derivadas de la línea germinal de otra especie de mamífero (por ejemplo, ratón), se han injertado en secuencias de FR humano. Los anticuerpos humanos también incluyen anticuerpos aislados de bibliotecas de inmunoglobulinas humanas o de animales transgénicos para una o más inmunoglobulinas humanas y que no expresan inmunoglobulinas endógenas.

El término "anticuerpo monoclonal" como se usa en el presente documento se refiere a una preparación de moléculas de anticuerpo de composición molecular única. Un anticuerpo monoclonal muestra una especificidad y afinidad de unión única por un epítipo particular. En una realización, los anticuerpos monoclonales son producidos por un hibridoma que incluye una célula B obtenida de un animal no humano, por ejemplo, ratón, fusionado a una célula inmortalizada. El término "anticuerpo recombinante", como se usa en el presente documento, incluye todos los anticuerpos que se preparan, expresan, crean o aíslan por medios recombinantes, tales como (a) anticuerpos aislados de un animal (por ejemplo, un ratón) que es transgénico o transcrómico con respecto a los genes de inmunoglobulina o un hibridoma preparado a partir de ellos, (b) anticuerpos aislados de una célula huésped transformada para expresar el anticuerpo, por ejemplo, de un transfectoma, (c) anticuerpos aislados de una biblioteca recombinante de anticuerpos combinatorios, y (d) anticuerpos preparados, expresados, creados o aislados por cualquier otro medio que implique el empalme de secuencias de genes de inmunoglobulina a otras secuencias de ADN. Como se usa en el presente documento, un "anticuerpo heterólogo" se define en relación con un organismo transgénico que produce dicho anticuerpo. Este término se refiere a un anticuerpo que tiene una secuencia de aminoácidos o una secuencia de ácido nucleico codificante correspondiente a la encontrada en un organismo que no consiste en el organismo transgénico, y que generalmente se deriva de una especie distinta del organismo transgénico. Como se usa en el presente documento, un "anticuerpo heterohíbrido" se refiere a un anticuerpo que tiene cadenas ligeras y pesadas que se originan de diferentes organismos. Por ejemplo, un anticuerpo que tiene una cadena pesada humana asociada con una cadena ligera murina es un anticuerpo heterohíbrido.

El término "fragmentos de unión a antígeno" se refiere a fragmentos de un anticuerpo que retienen la función de unirse específicamente a un antígeno o proteína antigénica pero que carecen de algunas o de todas las demás características estructurales de un anticuerpo o constructos artificiales que comprenden partes de anticuerpos. Los ejemplos preferidos de fragmentos de unión a antígeno incluyen, pero no se limitan a los siguientes fragmentos Fab, fragmento Fc, fragmento Fab', F(ab')₂, Fv de cadena sencilla, fragmentos variables de cadena sencilla divalente (di-scFv), scFv en tándem, diacuerpos, diacuerpos de cadena sencilla (scDB), triacuerpos, captadores de células T bi-específicas (BiTEs®), o moléculas de redireccionamiento de doble afinidad (moléculas DART®).

Los "fragmentos Fab" (también denominados "porción Fab" o "región Fab") cada uno con un único sitio de unión al antígeno, y un "fragmento Fc" residual (también denominado "porción Fc" o "región Fc") cuyo nombre refleja su capacidad de cristalizar fácilmente. El "fragmento Fab", que se refiere a un fragmento Fab que comprende adicionalmente la región bisagra de una molécula de Ig, mientras que los "fragmentos F(ab')₂" se entienden que comprenden dos fragmentos Fab' unidos químicamente o conectados mediante un enlace disulfuro. Al mismo tiempo que sdAb (Desmyter et al., 1996), di-scFv se pueden modificar genéticamente enlazando dos scFv (scFvA-scFvB). Esto se puede hacer produciendo una sola cadena peptídica con dos regiones VH y dos regiones VL, produciendo "scFv en tándem" (V_HA-V_LA-V_HB-V_LB). Otra posibilidad es la creación de scFv con enlazadores que son demasiado cortos para que las dos regiones variables se plieguen entre sí, lo que obliga a los scFv a dimerizarse. Usualmente, los enlazadores con una longitud de 5 residuos se usan para generar estos dímeros. Este tipo se conoce como "diacuerpos". Los enlazadores aún más cortos (uno o dos aminoácidos) entre un dominio de V_H y V_L conducen a la formación de trímeros mono-específicos, denominados "triacuerpos" o "triacuerpos". Los diacuerpos biespecíficos se forman expresando en cadenas con la disposición V_HA-V_LB y V_HB-V_LA o V_LA-V_HB y V_LB-V_HA, respectivamente. Los diacuerpos de cadena simple (scDb) comprenden un fragmento V_HA-V_LB y un fragmento V_HB-V_LA que están enlazados por un péptido enlazador (P) de 12-20 aminoácidos, preferiblemente 14 aminoácidos, (V_HA-V_LB-P-V_HB-V_LA). Los "captadores de células T bi-específicos (BiTE)" son proteínas de fusión que consisten en dos scFv de diferentes anticuerpos en los que uno de los scFv se une a las células T a través del receptor de CD3, y el otro a una célula tumoral a través de una molécula específica del tumor (Kufer et al., 2004). Las moléculas de redireccionamiento de doble afinidad (moléculas "DART") son diacuerpos adicionalmente estabilizados a través de un puente disulfuro en el terminal C.

Los términos "unión específica" o "uniendo específicamente" a un antígeno se refieren a la capacidad de un ligando para unirse a un determinante antigénico de un antígeno con alta afinidad. En ese contexto, "alta afinidad" significa que la K_d para la interacción está por debajo de 1 x 10⁻⁵ M, preferiblemente por debajo de 1 x 10⁻⁶ M, más preferiblemente por debajo de 1 x 10⁻⁷, incluso más preferiblemente debajo de 1 x 10⁻⁸ M y más preferiblemente debajo de 1 x 10⁻⁹ M.

Los términos "farmacéutico", "composición farmacéutica", "medicamento", "agente medicamentoso", "agente" y

"fármaco" se usan indistintamente en el presente documento en referencia a una sustancia y/o una combinación de sustancias que se usan para la identificación, prevención o tratamiento del estado o enfermedad de un tejido.

"Farmacéuticamente aceptable" significa aprobado por una agencia reguladora del gobierno federal o estatal o enumerado en la Farmacopea de los Estados Unidos u otra farmacopea generalmente reconocida para su uso en animales, y más particularmente en humanos.

El término "ingrediente activo" se refiere a la sustancia en una composición o formulación farmacéutica que es biológicamente activa, es decir, que proporciona valor farmacéutico. Una composición farmacéutica puede comprender uno o más ingredientes activos que pueden actuar conjuntamente o independientemente uno del otro. El ingrediente activo puede formularse como formas neutras o de sal. Las sales farmacéuticamente aceptables incluyen las formadas con grupos amino libres, tales como aquellas derivadas de ácidos clorhídrico, fosfórico, acético, oxálico, tartárico, etc., y aquellas formadas con grupos carboxilo libres, tales como pero sin limitarse a, las derivadas de sodio, potasio, amonio, calcio, hidróxidos férricos, isopropilamina, trietilamina, 2-etilaminoetanol, histidina, procaína y similares.

El término "portador", como se usa en el presente documento, se refiere a una sustancia farmacológicamente inactiva tal como, pero sin limitarse a, un diluyente, excipiente o vehículo con el que se administra el ingrediente terapéuticamente activo. Tales portadores farmacéuticos pueden ser líquidos o sólidos. Los portadores líquidos incluyen, pero no se limitan a, líquidos estériles, tales como soluciones salinas en agua y aceites, incluidos aquellos de petróleo, de origen animal, vegetal o sintético, tales como aceite de maní, aceite de soja, aceite mineral, aceite de sésamo y similares. Las soluciones salinas y las soluciones acuosas de dextrosa y glicerol también pueden emplearse como portadores líquidos, particularmente para soluciones inyectables. Una solución salina es un portador preferido cuando la composición farmacéutica se administra por vía intravenosa. Ejemplos de portadores farmacéuticos adecuados se describen en "Remington's Pharmaceutical Sciences" por E. W. Martin.

Los "excipientes" farmacéuticos adecuados incluyen almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, tiza, gel de sílice, estearato de sodio, monoestearato de glicerol, talco, cloruro de sodio, leche descremada en polvo, glicerol, propileno, glicol, agua, etanol y similares.

El término "adyuvante" se refiere a agentes que aumentan, estimulan, activan, potencian o modulan la respuesta inmune al ingrediente activo de la composición a nivel celular o humoral, por ejemplo, los adyuvantes inmunológicos estimulan la respuesta del sistema inmunitario al antígeno real, pero no tienen efecto inmunológico por sí mismos. Los ejemplos de tales adyuvantes incluyen, pero no se limitan a, adyuvantes inorgánicos (por ejemplo, sales metálicas inorgánicas tales como fosfato de aluminio o hidróxido de aluminio), adyuvantes orgánicos (por ejemplo, saponinas o escualeno), adyuvantes a base de aceite (por ejemplo, adyuvante completo de Freund y adyuvante incompleto de Freund), citoquinas (por ejemplo, IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF e INF- γ), adyuvantes en partículas (por ejemplo, complejos inmunoestimuladores (ISCOMS), liposomas o microesferas biodegradables), virosomas, adyuvantes bacterianos (por ejemplo, lípido A monofosforilado o péptidos de muramilo), adyuvantes sintéticos (por ejemplo, copolímeros de bloque no iónicos, análogos de péptidos de muramilo o lípido sintético A), o adyuvantes de polinucleótidos sintéticos (por ejemplo, poliarginina o polilisina).

Los productos farmacéuticos se administran por una vía de administración adecuada para el caso. Las rutas de administración incluyen, pero no se limitan a, administración intranasal, administración intramuscular, administración subcutánea, administración oral y administración tópica.

Una "administración intranasal" es la administración de un producto farmacéutico a la mucosa del tracto respiratorio completo, incluido el pulmón. Típicamente, el producto farmacéutico se administra a la mucosa de la nariz. La administración intranasal se logra mediante instilación, atomización o aerosol. Preferiblemente, dicha administración no implica la perforación de la mucosa por medios mecánicos tales como una aguja.

El término "administración intramuscular" se refiere a la inyección de un producto farmacéutico en cualquier músculo de un individuo o paciente. Las inyecciones intramusculares preferidas se administran en el deltoides, el vasto lateral o las áreas ventrogluteal y dorsogluteal.

El término "administración subcutánea" se refiere a la inyección de un producto farmacéutico en la hipodermis.

El término "administración oral" se refiere a la administración de un producto farmacéutico a través de la boca al sistema gástrico.

Una "administración tópica" es la administración de un producto farmacéutico a cualquier parte de la piel sin penetrar en la piel con una aguja o un dispositivo comparable. El producto farmacéutico también puede administrarse tópicamente a la mucosa de la boca, nariz, región genital y recto.

El término "soporte de datos" se refiere a cualquier medio de almacenamiento utilizable para proporcionar información al usuario. Estos pueden incluir, pero sin limitarse a, manuales, libros, folletos, volantes, papel, así como un medio de almacenamiento legible por ordenador. El término "medio de almacenamiento legible por ordenador", como se usa en el presente documento, se refiere a cualquier medio capaz de almacenar datos en un formato legible por un dispositivo mecánico tal como por ejemplo, un ordenador. Los ejemplos de medios incluyen, pero no se limitan a, un disco

magnético, un disco óptico, un disco duro, una RAM, una ROM, una EEPROM (por ejemplo, memoria flash), una EPROM, un DVD, un disco Blu-Ray, un disquete, un disco DVD-RW, un código de barras y caracteres de tinta magnética.

Realizaciones

5 En los siguientes pasajes se definen diferentes aspectos de la invención con más detalle. Cada aspecto así definido puede combinarse con cualquier otro aspecto o aspectos a menos que se indique claramente lo contrario. En particular, cualquier característica indicada como preferida o ventajosa puede combinarse con cualquier otra característica o características indicadas como preferidas o ventajosas. En el trabajo que condujo a la presente invención, se demostró sorprendentemente que también el ARN intrínico transcrito a partir de repeticiones de hexanucleótidos, presentes en
10 el genoma, podría traducirse mediante traducción RAN.

En un primer aspecto, la presente invención proporciona un anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido que consiste en repeticiones de dipéptidos (Gly-Ala)_a, en las que a es un número entero de 16 o más, en los que el anticuerpo o el fragmento de unión a antígeno del mismo comprende una
15 secuencia CDR3 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 53; una secuencia de CDR2 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 54; y una secuencia de CDR1 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 55,

y

20 (i) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende EILPGSGSTK de acuerdo con la SEQ ID NO: 64 y una secuencia de CDR3 que comprende GDFTNSHFAY de acuerdo con la SEQ ID NO: 60;

o

25 (ii) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGSTK de acuerdo con la SEQ ID NO: 65 y una secuencia de CDR3 que comprende GDYSNSHFAY de acuerdo con la SEQ ID NO: 57;

o

(iii) una secuencia de CDR1 de cadena pesada que comprende GYKFIGYWIE de acuerdo con la SEQ ID NO: 69, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGTTK de acuerdo con la SEQ ID NO: 66 y una secuencia de CDR3 que comprende GDYSNSHFTY de acuerdo con la SEQ ID NO: 59.

30 La enfermedad caracterizada por una expansión de las repeticiones genómicas de hexanucleótidos es una enfermedad neurológica, más preferiblemente una enfermedad neurodegenerativa. Las repeticiones genómicas de hexanucleótidos en tales enfermedades se seleccionan del grupo que consiste en GGGGCC y GGCCTG en una cadena de ADN y las repeticiones de hexanucleótidos GGCCCC y CAGGCC en la cadena de ADN complementaria respectiva, es decir, la enfermedad se caracteriza por una expansión de las repeticiones de hexanucleótidos GGGGCC
35 y/o GGCCTG en una cadena de ADN y las repeticiones de hexanucleótidos GGCCCC y CAGGCC en la cadena de ADN complementaria respectiva. Preferiblemente, la expansión de repeticiones genómicas de hexanucleótidos GGGGCC (y/o las repeticiones de hexanucleótido complementarios GGCCCC) caracteriza una enfermedad seleccionada del grupo que consiste en ALS, FTD, ALS-FTD, y la expansión de repeticiones genómicas de hexanucleótidos GGCCTG (y/ o las repeticiones de hexanucleótidos complementarios CAGGCC) caracteriza la
40 enfermedad SCA36. Por consiguiente, la expansión de repeticiones de hexanucleótido genómicos GGGGCC y/o GGCCTG en una cadena de ADN y la repetición de hexanucleótido GGCCCC y CAGGCC en la cadena de ADN complementaria respectiva es un indicador de las enfermedades respectivas.

Se divulga un polipéptido que comprende o consiste en repeticiones de dipéptidos con una secuencia seleccionada del grupo que consiste en (Gly-Ala)_a, en la que a es un número entero de 16 o más, preferiblemente 20, 25, 30, 35,
45 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 o más.

En realizaciones preferidas, una etiqueta detectable se une, preferiblemente se enlaza covalentemente, a dicho polipéptido. Las etiquetas adecuadas se seleccionan preferiblemente del grupo que consiste en un fluoróforo, un cromóforo, un radiomarcador, un coloide metálico, una enzima, una molécula quimioluminiscente o bioluminiscente.
50 Se prefiere particularmente que la etiqueta detectable sea un fluoróforo, preferiblemente GFP, o una enzima, preferiblemente una luciferasa.

El anticuerpo o fragmento de unión a antígeno del mismo del primer aspecto de la invención se une o forma específicamente un complejo que consiste en repeticiones de dipéptidos con una secuencia (Gly-Ala)_a, en la que a es un número entero de 16 o más, preferiblemente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150,
55 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500,

4000, 4500, 5000 o más. El anticuerpo o fragmento de unión al mismo se une a un epítipo formado por el dipéptido que se repite.

Preferentemente, el anticuerpo se selecciona del grupo que consiste en anticuerpos monoclonal, monovalente, biespecífico, heteroconjugado, multiespecífico, recombinante, heterólogo, heterohíbrido, quimérico, humanizado (en particular injertados con CDR), desimmunizado o humano. Más preferiblemente, el anticuerpo es un anticuerpo humanizado. En realizaciones preferidas adicionales, el anticuerpo es un anticuerpo IgG. En realizaciones preferidas adicionales, el fragmento de unión a antígeno se selecciona del grupo que consiste en fragmento Fc, fragmentos Fab, fragmento Fab', fragmentos F(ab')₂, Fv de cadena sencilla (scFv), scFv en tándem, diacuerpos, diacuerpos biespecíficos, diacuerpos de cadena simple (scDb), triacuerpos, captadores de células T bi-específicos (BiTEs®) y moléculas de redireccionamiento de doble afinidad (DART®).

Los anticuerpos o fragmentos de unión a antígeno de los mismos de acuerdo con el primer aspecto de la presente invención comprenden una cadena ligera que comprende una secuencia de CDR3 que comprende, que consiste esencialmente o que consiste en la secuencia de aminoácidos QQLVEYPLT (SEQ ID NO: 53); una secuencia de CDR2 de cadena ligera que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos LMSTRAS (SEQ ID NO: 54) y una secuencia de CDR1 de cadena ligera que comprende, que consiste esencialmente en, o que consiste en una secuencia de aminoácidos RSSKSLLYKDGKTYLN (SEQ ID NO: 55) y

(i) una cadena pesada que comprende una secuencia de CDR3 que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos GDFTNSHFAY (SEQ ID NO: 60); una secuencia de CDR2 de cadena pesada que consiste esencialmente en, o que consiste en la secuencia de aminoácidos EILPGSGSTK (SEQ ID NO: 64) y una secuencia de CDR1 de cadena pesada que comprende, que consiste esencialmente en, o que consiste en una secuencia de aminoácidos GYTFTGYWIE (SEQ ID NO: 68), preferiblemente este anticuerpo comprende una o más, lo más preferiblemente todas las regiones marco del anticuerpo GA 1A12;

o

(ii) una cadena pesada que comprende una secuencia de CDR3 que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos GDYSNSHFAY (SEQ ID NO: 57); una secuencia de CDR2 de cadena pesada que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos ENLPGSGSTK (SEQ ID NO: 65) y una secuencia de CDR1 de cadena pesada que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos GYTFTGYWIE (SEQ ID NO 68), preferiblemente este anticuerpo comprende una o más, lo más preferiblemente todas las regiones marco del anticuerpo GA 5E9;

o

(iii) una cadena pesada que comprende una secuencia de CDR3 que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos GDYSNSHFTY (SEQ ID NO: 59); una secuencia de CDR2 de cadena pesada que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos ENLPGSGTTK (SEQ ID NO: 66) y una secuencia de CDR1 de cadena pesada que comprende, que consiste esencialmente en, o que consiste en la secuencia de aminoácidos GYKFIGYWIE (SEQ ID NO: 69), preferiblemente este anticuerpo comprende una o más, lo más preferiblemente todas las regiones marco del anticuerpo GA 5F2.

En otra realización más, la invención proporciona un anticuerpo, o un fragmento de unión a antígeno del mismo, con una región variable de cadena ligera (LCVR) que comprende, que consiste esencialmente en o que consiste en la secuencia de aminoácidos de la SEQ ID NO: 80 y una región variable de cadena pesada (HCVR) que comprende, que consiste esencialmente en o que consiste en la secuencia de aminoácidos seleccionada del grupo de la SEQ ID NO: 80, 81 y 82.

En realizaciones preferidas del cuarto aspecto, dicho anticuerpo o fragmento de unión a antígeno se une específicamente al polipéptido (Gly-Ala)_n de acuerdo con el segundo y tercer aspecto con un valor de K_d de 10 nM o menos, preferiblemente 9 nM o menos, más preferiblemente 8 nM o menos, más preferiblemente 7 nM o menos, más preferiblemente 6 nM o menos, más preferiblemente 5 nM o menos, más preferiblemente 4 nM o menos, más preferiblemente 3 nM o menos, más preferiblemente 2 nM o menos, e incluso más preferiblemente 1 nM o menos.

En realizaciones preferidas, una etiqueta detectable está unida, preferiblemente se enlaza covalentemente, al anticuerpo o fragmento de unión a antígeno del mismo. Las etiquetas adecuadas se seleccionan preferiblemente del grupo que consiste en un fluoróforo, un cromóforo, un radiomarcador, un coloide metálico, una enzima, una molécula quimioluminiscente o bioluminiscente. Se prefiere particularmente que la etiqueta detectable sea un fluoróforo, preferiblemente GFP, o una enzima, preferiblemente una luciferasa.

En un tercer aspecto, la presente invención proporciona el anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con el primer aspecto de la invención para el diagnóstico *in vitro* de una enfermedad caracterizada por una expansión de repeticiones genómicas de GGGGCC, en la que la enfermedad se selecciona de el grupo que consiste en esclerosis lateral amiotrófica (ALS), demencia frontotemporal (FTD) y esclerosis lateral amiotrófica-demencia frontotemporal (ALS-FTD).

- En realizaciones preferidas, el anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con el primer aspecto de la invención se usa para diagnosticar *in vitro* una enfermedad caracterizada por una expansión de repeticiones de hexanucleótidos genómicos, preferiblemente detectando dicho polipéptido en un individuo que tiene, se sospecha que tiene o es susceptible a dicha enfermedad. Preferiblemente, dicho polipéptido se detecta por cualquiera de los métodos descritos anteriormente, preferiblemente por medio de transferencia Western, ELISA, FACS e inmunocitoquímica (por ejemplo, mediante microscopía de fluorescencia o microscopía electrónica). En realizaciones preferidas, la detección del polipéptido que consiste en repeticiones de dipéptidos (Gly-Ala) es un indicador de la presencia de la enfermedad.
- En un segundo aspecto, la presente invención proporciona un producto farmacéutico que comprende el anticuerpo o fragmento de unión al mismo de acuerdo con el primer aspecto de la invención y/o un ácido nucleico que codifica dicho anticuerpo o fragmento de unión a antígeno del mismo.
- En realizaciones preferidas, el producto farmacéutico comprende además un portador, excipiente y/o adyuvante farmacéuticamente aceptable y opcionalmente una o más sustancias activas adicionales. Las sustancias activas adicionales pueden ser efectivas contra la misma o contra una enfermedad o trastorno diferente. En consecuencia, las sustancias activas adicionales comprenden aquellas que tienen un efecto beneficioso en el tratamiento de una enfermedad caracterizada por una expansión de repeticiones genómicas de hexanucleótidos. Además, las sustancias activas adicionales pueden dirigirse al tratamiento de diferentes enfermedades o trastornos o afecciones adecuadas para administrarse en combinación con el ingrediente activo para el tratamiento de una enfermedad caracterizada por una expansión de las repeticiones genómicas de hexanucleótidos.
- Preferiblemente, el producto farmacéutico del segundo aspecto contiene una cantidad terapéuticamente eficaz de dicho anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con el primer aspecto de la invención y/o un ácido nucleico que codifica dicho anticuerpo o fragmento de unión a antígeno del mismo, preferiblemente en forma purificada, junto con una cantidad adecuada de portador y/o excipiente para proporcionar la forma de administración adecuada al paciente. La formulación debe adaptarse al modo de administración.
- El producto farmacéutico puede tomar la forma de soluciones, suspensiones, emulsiones, tabletas, píldoras, cápsulas, polvos, formulaciones de liberación sostenida y similares. El producto farmacéutico puede formularse como un supositorio, con aglutinantes y portadores tradicionales tales como los triglicéridos.
- Para preparar los productos farmacéuticos de la presente invención, los vehículos farmacéuticamente aceptables pueden ser sólidos o líquidos. Las composiciones en forma sólida incluyen polvos, tabletas, píldoras, cápsulas, pastillas, cápsulas lisas, supositorios y gránulos dispersables. Un excipiente sólido puede ser una o más sustancias, que también pueden actuar como diluyentes, agentes saborizantes, aglutinantes, conservantes, agentes desintegrantes de tabletas o un material encapsulante. En polvos, el excipiente es preferiblemente un sólido finamente dividido, que está en una mezcla con el inhibidor finamente dividido de la presente invención. En tabletas, el ingrediente activo se mezcla con el portador que tiene las propiedades de unión necesarias en proporciones adecuadas y se compacta en la forma y tamaño deseados. Los excipientes adecuados son carbonato de magnesio, estearato de magnesio, talco, azúcar, lactosa, pectina, dextrina, almidón, gelatina, tragacanto, metilcelulosa, carboximetilcelulosa de sodio, una cera de bajo punto de fusión, manteca de cacao y similares. Para preparar supositorios, primero se derrite una cera de bajo punto de fusión, tal como una mezcla de glicéridos de ácidos grasos o manteca de cacao, y el componente activo se dispersa homogéneamente, tal como por agitación. La mezcla homogénea fundida se vierte en moldes de tamaño conveniente, se deja enfriar y, por lo tanto, se solidifica. Las tabletas, polvos, cápsulas, píldoras, cápsulas lisas y pastillas se pueden usar como formas de dosificación sólidas adecuadas para administración oral.
- Las composiciones en forma líquida incluyen soluciones, suspensiones y emulsiones, por ejemplo, agua, soluciones salinas, dextrosa acuosa, soluciones de glicerol o soluciones de agua/propilenglicol. Para inyecciones parenterales (por ejemplo, inyecciones intravenosa, intraarterial, infusión intraósea, intramuscular, subcutánea, intraperitoneal, intradérmica e intratecal), las preparaciones líquidas se pueden formular en solución, por ejemplo, en solución acuosa de polietilenglicol. Una solución salina es un portador preferido cuando la composición farmacéutica se administra por vía intravenosa.
- Preferiblemente, la composición farmacéutica está en forma de dosificación unitaria. En tal forma, la composición puede subdividirse en dosis unitarias que contienen cantidades apropiadas del componente activo. La forma de dosificación unitaria puede ser una composición empaquetada, conteniendo el empaque cantidades discretas de la composición, tales como tabletas, cápsulas y polvos empacados en viales o ampollas. Además, la forma de dosificación unitaria puede ser una cápsula, un vial de inyección, una tableta, una cápsula lisa o una pastilla misma, o puede ser el número apropiado de cualquiera de estas en forma empacada. La composición, si se desea, también puede contener cantidades menores de agentes humectantes o emulsionantes, o agentes reguladores de pH.
- En un tercer aspecto, la presente invención proporciona el uso de un anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con la reivindicación 1 en el diagnóstico *in vitro* de una enfermedad caracterizada por una expansión de repeticiones genómicas de GGGGCC, en la que la enfermedad se selecciona de grupo consiste en esclerosis lateral amiotrófica (ALS), demencia frontotemporal (FTD) y esclerosis lateral amiotrófica-demencia frontotemporal (ALS-FTD).

Las repeticiones de hexanucleótidos tienen una secuencia de GGGGCC. Preferiblemente, el hexanucleótido (GGGGCC)_n es responsable de la formación de agregados responsables de ALS, FTD o ALS-FTD, y el número "n" es al menos 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450 y 500 o más.

- 5 Los siguientes ejemplos y figuras son meramente ilustrativos de la presente invención y no deben constituirse en limitantes del alcance de la invención como se indica por las reivindicaciones adjuntas de ninguna manera.

Ejemplos

Ejemplo 1 - Anticuerpos

- 10 Para detectar directamente la expresión de las proteínas de repetición postuladas, se generaron anticuerpos (anti-GA y anti-GP) contra péptidos (GA)₁₅ y (GP)₁₅ fusionados con la proteína de unión a maltosa y se probó un anticuerpo monoclonal (anti GR) que originalmente se generó contra un epítipo EBNA2A que incluye una repetición (GR)₆ (Kremmer et al., 1995).

- 15 Se inmunizaron conejos con proteínas de fusión MBP que contenían repeticiones (GA)₁₅ o (GP)₁₅ en el terminal C (Eurogentec). El suero resultante se purificó con proteínas de fusión GST que contenían (GA)₁₅ o (GP)₁₅ en el terminal C. Brevemente, la proteína expresada en *E. coli* se entrelazó con Glutatión Sefarosa (Pierce) y los anticuerpos unidos se eluyeron mediante cambio de pH (NaCl 0,15 M, glicina 0,1 M pH = 2,5) y se neutralizaron inmediatamente (0,1 volúmenes de Tris 1 M pH = 9,5). El anticuerpo 5A2 monoclonal específico de la repetición de poli-GR de rata se generó originalmente contra el antígeno peptídico (GQSRGQSRGRGRGRGRGRGKGGK; SEQ ID NO: 47) de EBNA2 y detecta específicamente GST-(GR)₁₅ (Figura 1A) (Kremmer et al., 1995). Anticuerpo p62 (ratón, BD transducción),
20 TDP de conejo (Protein Tech), fosfo-TDP-43 (Ser409/Ser410, clon 1D3) (Neumann et al., 2009) y GST de rata (6G9, Sigma) están disponibles comercialmente.

Los tres anticuerpos DPR detectaron el antígeno de repetición respectivo mediante inmunotransferencia sin reacción cruzada con las otras dos proteínas DPR (Figura 1A).

Ejemplo 2 - Constructos de repetición de hexanucleótidos

- 25 Para probar si tales proteínas de repetición pueden traducirse en ausencia de un codón de inicio, se clonó la región de repetición de pacientes con *C9orf72* en un vector de expresión mamífero usando PCR cebada de repetición.

- 30 La repetición del hexanucleótido se amplificó por PCR cebada de repetición (6) de linfoblastos del paciente (AGTACTCGCTGAGGGTGAAC; SEQ ID NO: 48; y CGTACGC ATCCCAGTTTGAGAGCCCCGGCCCCGGCCCCGG CCCC; SEQ ID NO: 49) y clonación TOPO en pEF6-V5/His (Invitrogen). Ni el sitio de clonación múltiple ni la secuencia genómica *C9orf72* contienen codones de inicio ATG. Adicionalmente, múltiples codones de parada están presentes en todos los marcos de lectura secuencia arriba de la repetición de GGGGCC. Los plásmidos se transformaron en bacterias Stbl2 (Invitrogen) que se cultivaron a 32 °C para evitar la recombinación. Debido al contenido extremo de GC, solo se pueden secuenciar ~20 repeticiones. Por lo tanto, se estimaron números de repetición más altos por digestión de restricción (Figura 1B).

- 35 Para los constructos más largos solo se pudo usar digestión de restricción para estimar el número de repetición que oscila entre ~28 y ~145, debido a que el contenido extremo de GC impide la secuenciación (Figura 1B panel superior). La región secuencia arriba de la repetición de GGGGCC carece de codones de inicio ATG y está construida en células HEK293 (cultivadas con DMEM suplementado con FCS al 10% y penicilina/estreptomina), los anticuerpos anti-GA detectaron proteínas de tamaño creciente comenzando con un producto débil de ~38 repeticiones, lo que sugiere que
40 los mecanismos de traducción requieren una longitud mínima de repetición (Figura 1B, flecha). Sin embargo, no se puede excluir la traducción de polipéptidos más cortos que no se detectaron debido a la degradación rápida o un menor número de epítipos repetidos. No se detectaron productos de poli-GR (datos no mostrados) y solo el constructo más largo con ~145 repeticiones expresaron adicionalmente cantidades detectables de poli-GP (Figura 1B). Juntos, estos datos indican que las repeticiones de GGGGCC de longitud media se traducen predominantemente en poli-GA
45 y cantidades menores de proteínas poli-GP en células HEK293 (Figura 1B).

Ejemplo 3 - Ensayo de trampa de filtro

Para analizar la agregación de poli-GA, se realizaron ensayos de trampa de filtro de cerebelo post mortem humano de controles sanos y pacientes con FTL/ALS-TDP con y sin expansión repetida de hexanucleótidos de *C9orf72*. Este método está bien establecido para el estudio de la agregación de poli-Q (Li et al., 2003).

- 50 El material del paciente fue proporcionado por Neurobiobank Munich, Ludwig-Maximilians-University Munich y el Antwerp Bio Bank en el Instituto Born-Bunge, Universidad de Antwerp y fue recopilado y distribuido de acuerdo con las directrices de los comités éticos locales.

El cerebelo se homogeneizó en PBS que contenía Triton-X 100 al 1%, MgCl₂ 15 mM y 0,2 mg/mL de DNasa I. Tras la centrifugación (180.000 g durante 30 min a 12 °C) el sedimento insoluble se resuspendió en SDS al 2% 100 mM Tris

pH = 7 y se incubó a temperatura ambiente durante 1 hora y luego se filtró a través de una membrana de acetato de celulosa con un tamaño de poro de 0,2 μm (Li et al., 2003).

Se observó una fuerte señal de poli-GA solo en pacientes con FTLD/ALS con expansión repetida de hexanucleótidos (Figura 1C) que indica que poli-GA forma agregados insolubles en SDS en el cerebelo. Además, también se detectó poli-GP y poli-GR insolubles en pacientes con *C9orf72* (Figura 1C). El material insoluble en SDS al 2% se solubilizó parcialmente al hervir en SDS al 8% y se pudo analizar mediante SDS-PAGE. Los tres anticuerpos específicos de DPR detectaron agregados de alto peso molecular en la parte superior del gel en portadores de mutación *C9orf72*, pero no en los controles (Figura 1D). Estos hallazgos indican que se genera agregación de proteínas DPR en los tres marcos de lectura específicamente en pacientes con mutación *C9orf72*.

10 Ejemplo 4 - Inmunotinción

Las secciones embebidas en parafina se desparafinizaron primero con xileno y etanol. Después de un breve lavado con agua desionizada, la recuperación de antígeno se realizó sometiendo a microondas las secciones seis veces 3 min en regulador de citrato 100 mM (pH 6,0). Las secciones se dejaron reposar durante 20 minutos seguidas de un breve lavado con agua desionizada. La peroxidasa endógena se bloqueó con H_2O_2 al 5% en metanol durante 15 minutos. Después de enjuagar con agua desionizada, las secciones se asentaron en PBS con Brij35 al 0,05%. El bloqueo se realizó con FCS al 2% en PBS durante 5 min. Los anticuerpos primarios se diluyeron con solución de bloqueo y se incubaron durante la noche a 4 °C. Para los anticuerpos monoclonales de rata 5A2 y 1D3, se usó adicionalmente anticuerpo secundario anti-rata de conejo (PARIS Anticorps) a una dilución 1:2000 durante 1 h. Las secciones se lavaron dos veces con PBS con Brij35 al 0,05% durante 5 minutos y se desarrollaron con el kit de IHC 20 DCS SuperVision 2 HRP (DCS) o el sistema de detección de polímeros NovoLink para anticuerpos de rata (Leica) usando DAB como cromógeno. Se obtuvieron imágenes microscópicas con un microscopio BX50 y el software Cell-D (Olympus). Para las especies de inmunofluorescencia, se usaron anticuerpos secundarios específicos conjugados con Alexa-488 o Alexa-544 (Invitrogen) para doble tinción. Se usó TOPRO-3 (Invitrogen) como tinción nuclear. La autofluorescencia se redujo con Negro Sudán al 0,3%. Se obtuvieron imágenes confocales con LSM510 (Zeiss) con 25 una lente de inmersión en aceite 63x/1.4, utilizando un diámetro de agujero de alfiler de 1 unidad Airy. El brillo y el contraste de cada imagen se mejoraron linealmente utilizando el software LSM (Zeiss).

Ejemplo 5 - Patrones de distribución celular de las proteínas DPR

Para determinar los patrones de distribución celular de estas proteínas DPR en pacientes con mutaciones *C9orf72*, nos enfocamos en el cerebelo y el hipocampo en el análisis inmunohistoquímico ya que estas regiones del cerebro 30 contienen abundantes patologías de inclusiones (Al-Sarraj et al., 2011; Mahoney et al., 2012; Bigio et al., 2012) (Figura 3A y 4A/B). Sorprendentemente, en todos los casos con expansión de repetición de *C9orf72*, los anticuerpos específicos de poli-GA detectaron inclusiones citoplasmáticas neuronales en forma de punto (Figura 3B y 4C). Su forma y abundancia fue similar a las inclusiones positivas para p62/negativas para TDP-43 consideradas 35 patognomónicas para los casos de mutación *C9orf72*. Además, estos tipos de inclusiones también fueron detectados por anticuerpos contra poli-GP y poli-GR, sin embargo, en un grado mucho menor (Figura 3C/D y 4D/E). Esto sugiere que, en comparación con las proteínas poli-GA, las proteínas poli-GP y poli-GR se traducen de manera menos eficiente o tienen una menor propensión de agregación (compárese con la Figura 1B). En el hipocampo, la mayoría de las inclusiones teñidas por los tres anticuerpos DPR se parecían a las inclusiones en forma de estrella positivas para p62 40 típicas de los pacientes con *C9orf72* (Figura 4F-I y 5). Como se informó para las tinciones de p62 (Al-Sarraj et al., 2011; Bigio et al., 2012), también se observaron algunas inclusiones intranucleares neuronales positivas para DPR (Figura 4J). Una patología DPR similar es visible en otras regiones del cerebro, incluida la capa granular del giro dentado, la capa molecular del cerebelo y la neocorteza (Figura 4K/L y Figura 2).

Ejemplo 6 - Co-localización de agregados positivos para DPR y positivos para p62

A continuación se analizó si los agregados positivos para DPR son idénticos a los agregados positivos para p62. La 45 inmunofluorescencia dual en secciones del hipocampo de un caso con *C9orf72* mostró co-localización de poli-GA, poli-GP y poli-GR con p62 en las inclusiones características similares a estrella (Figura 4M-O y 6D/E). Sin embargo, no hubo coagregación de las proteínas fosfo-TDP-43 y DPR (Figura 6A-C), aunque ocasionalmente los pequeños agregados esféricos de poli-GA estaban rodeados de fosfo-TDP-43 agregados formando un núcleo dentro de las inclusiones de fosfo-TDP-43 (Figura 6F). De acuerdo con el ensayo de trampa de filtro (véase la Figura 1C), tales 50 agregados de poli-GA no fueron detectables en casos de FTLD-TDP sin expansiones repetidas de *C9orf72* o en casos con enfermedad de Huntington, que presenta estiramientos poli-Q expandidos (Figura 7A-B). En total, se identificaron agregados intraneuronales que contenían poli-GA, poli-GP y poli-GR en los siete pacientes analizados con expansión de repetición de *C9orf72* genéticamente confirmada, pero no en otros nueve casos con una longitud de repetición normal (véase la Figura 2). Curiosamente, algunos pacientes con mutaciones *C9orf72* mostraron notablemente pocas 55 inclusiones de fosfo-TDP-43 en todo el cerebro. Hasta ahora, solo se ha informado un caso excepcional (TJ-9) clasificado como FTLD-UPS con mutación *C9orf72* y prominente patología de ubiquitina pero sin patología detectable de TDP-43 (3, 6). Se descubrieron abundantes agregados de poli-GA y algunos agregados de poli-GP y poli-GR en la corteza temporal de este paciente (Figura 7C y Figura 2), lo que sugiere que las proteínas DPR son cruciales para la patogénesis de FTLD en este caso. Por lo tanto, se propone que poli-GA es la principal especie de agregación en 60 pacientes con FTLD-UPS con expansión de repetición de *C9orf72* y, por lo tanto, puede ser la principal causa de

neurodegeneración.

Ejemplo 7 - Análisis de RT-PCR del nivel de ARNm de *C9orf72* y la repetición de GGGGCC

Hasta ahora, los datos más sólidos para el patomecanismo en pacientes con *C9orf72* apoyan la haploinsuficiencia a través de una expresión fuertemente reducida del alelo mutante (DeJesus-Hernandez et al., 2011, y Gijssels et al., 2012, van der Zee et al., Human Mutation 2012).

Para confirmar aún más el patomecanismo de DPR se evaluó el nivel de ARNm de *C9orf72* en pacientes a través de análisis de RT-PCR (la posición de los cebadores utilizados se indica en la Figura 8A).

El ARN del paciente se extrajo usando el kit RNeasy (Qiagen) que incluye una etapa de ADNasa I para eliminar el ADN genómico. Se confirmó una calidad de ARN suficiente utilizando el Nano kit Agilent RNA 6000 en el Bioanizador Agilent 2100. Los valores de RIN resultantes (4,7 - 7,2) no fueron estadísticamente diferentes entre los casos y controles de *C9orf72* ($p = 0,25$, prueba t de Student). El ADNc se generó usando el kit de transcripción inversa MicroARN Taqman (Life Technologies) con hexámeros aleatorios o cebadores específicos de cadena para *C9orf72* (CAATCCACCAGTCGCTAGA; SEQ ID NO: 37; y CTGCGGTTGCGGTGCCTGC; SEQ ID NO: 38). Se realizó PCR cuantitativa utilizando SsoFast EvaGreen (Biorad) usando cebadores intrónicas específicos de *C9orf72* (AAGAGGCGCGGTTAGAAG; SEQ ID NO: 39; y AGTCGCTAGAGGCGAAAGC; SEQ ID NO: 40) y cebadores exónicos que abarcan intrones para *C9orf72* (todas las isoformas: ACTGGAAT GGGGATCGCAGCA; SEQ ID NO: 41; y ACCCTGATCTTCCATTCTCTGTGCC; SEQ ID NO: 42; isoformas que contienen el exón alternativo 1b: CTGCGGTTGCGGTGCCTGC; SEQ ID NO: 43; y AGCTGGAGATGGCGGTGGGC; SEQ ID NO: 44) y el gen de mantenimiento *YHAZ* (TGAACAAAAGACGGAAGGTGCTG; SEQ ID NO: 45; y TCTGATAGGATGTGTTGTTGCA; SEQ ID NO: 46).

Se confirmó que los niveles de ARNm de *C9orf72* se redujeron ~ 50% en el cerebro del paciente (Figura 8B). Sin embargo, la expresión del intrón 1 (donde se encuentra la repetición de GGGGCC) aumenta fuertemente en los pacientes con *C9orf72* que argumenta una estabilización selectiva de la repetición que contiene pre-ARNm (o el intrón 1 extirpado solo). Se estableció que, además de los transcritos sentido, también se aumentaron los transcritos antisentido en pacientes con *C9orf72* (véase la Figura 8B). Este hallazgo tiene implicaciones importantes para los esfuerzos terapéuticos que utilizan oligonucleótidos antisentido como tratamiento para pacientes con *C9orf72* y enlaza los tres mecanismos propuestos de la enfermedad: haploinsuficiencia, captura de proteínas de unión a ARN y traducción de RAN.

Ejemplo 8 - Análisis cuantitativo de la colocalización de inclusiones de p62 y poli-GA/GP/GR

Para analizar qué fracción de tipos de células individuales en cualquier región cerebral afectada muestra acumulaciones de la repetición del dipéptido y evaluar el grado preciso de colocalización de inclusiones inmunorreactivas de poli-GA/GP/GR con aquellas inmunomarcadas por el anticuerpo p62, se analizó cuantitativamente la colocalización de poli-GA con p62 en las inclusiones características en tres pacientes (véase la Figura 6G). Colocalización de inclusiones positivas para p62 y DPR en el cerebelo y la región CA4 del hipocampo cuantificadas por inmunofluorescencia doble en tres pacientes con mutación *C9orf72* (TJ-1, TJ-2 y TJ-3). En el cerebelo, en total, se contabilizaron 370 a 822 (análisis de poli-GA), 123 a 566 (análisis de poli-GP) y 211 a 596 (análisis de poli-GR) inclusiones por paciente. En CA4 en total se analizaron 48-73 inclusiones por paciente para cada especie de DPR. En el cerebelo, más del 90% de las inclusiones positivas para p62 también fueron positivas para poli-GA. En el hipocampo, este número fue incluso mayor: 98%. Además, se analizó la colocalización de todos los poli-GP y poli-GR en CA4. De acuerdo con la impresión del análisis inmunohistoquímico (véase más arriba), poli-GA es la especie predominante de DPR. Esto sugiere fuertemente que poli-GA es el componente principal en inclusiones positivas para p62 y negativas para TDP-43 y, por lo tanto, una nueva proteína importante de la enfermedad en pacientes con *C9orf72* que permite el diagnóstico, profilaxis y tratamiento de la enfermedad.

Ejemplo 9 - Identificación de proteínas de unión específicas para la repetición del hexanucleótido GGGGCC

Para identificar proteínas, que podrían estar unidas por las repeticiones de hexanucleótidos GGGGCC, se realizaron ensayos de inmunoprecipitación pull-down usando ARN biotinilado transcrito *in vitro* que contienen 23 repeticiones GGGGCC o 17 AAAACC como control. Las sondas de ARN transcritas *in vitro* biotiniladas se incubaron con extractos nucleares de células HEK 293 (el extracto nuclear se preparó como se describió previamente (Dignam et al., 1983)). La concentración de proteína se determinó mediante el método BCA en ausencia o en presencia de un exceso de 50 veces de ARN competidor no biotinilado que contiene la repetición GGGGCC (Figura 9A). El ARN competidor evitó la unión como se indica mediante un flujo fuertemente mejorado (Figura 9A; carril 4). Usando concentraciones crecientes de sal para la elución de las proteínas unidas, se observó afinidad diferencial de unión de proteínas a las repeticiones GGGGCC/AAAACC tras la elución con NaCl 500 mM (Figura 9A). Las proteínas que se eluyen con NaCl 500 mM de las repeticiones GGGGCC o AAAACC se sometieron a LC-MS/MS. Esto permitió la identificación de 235 proteínas en tres réplicas. Se identificaron 188 proteínas al menos dos veces en los tres experimentos de inmunoprecipitación pull-down. De estas 188 proteínas, la unión de 127 proteínas podría competir con un exceso de sondas de GGGGCC no biotiniladas. 72 proteínas mostraron al menos una unión de 2 veces a GGGGCC sobre AAAACC. Finalmente se seleccionaron todas las proteínas con abundancia ≥ 20 en la muestra de GGGGCC. Estos estrictos criterios de

filtración dieron como resultado 20 proteínas candidatas principales, de las cuales la mayoría eran factores conocidos que interactúan con ARN, tales como ribonucleoproteínas heterogéneas (RNPhn), factores de empalme y proteínas de unión a ARNm (Figura 10). Luego se confirmó una selección de estas 20 proteínas para las cuales los anticuerpos específicos estaban disponibles comercialmente mediante transferencia Western de las fracciones de elución. Para RNPhn A3, RNPhn A2B1, SFPQ, ILF3, NONO, RNPhn L, IL2BP1, ILF-2 y FUS, la unión fuerte y selectiva a la repetición de GGGGCC se pudo confirmar en la fracción de NaCl 500 mM, y estas señales fueron completamente bloqueadas por un doble exceso de 50 veces de la sonda GGGGCC no marcada (Figura 9B).

Ejemplo 10: Identificación de anticuerpos que se unen específicamente a la repetición de GGGGCC

Para determinar si los anticuerpos se unen específicamente con alta afinidad a un polipéptido que consiste en repeticiones de dipéptidos de la secuencia (Gly-Ala), se generaron tres anticuerpos monoclonales de ratón diferentes (GA-5E9; GA-5F2 y GA-1A12). La generación de anticuerpos monoclonales con la estrategia publicada en Mori et al., Science 2013, fracasó porque predominantemente se indujeron IgM, que no son estables en las células de hibridoma y tienen poca afinidad. Solo la inmunización de péptidos GA₁₀ agregados dio como resultado anticuerpos IgG anti-GA. Los anticuerpos se generaron inmunizando ratones CBL por vía subcutánea e intraperitoneal con 30 µg de péptido recombinante agregado que contenía un GA₁₀ unido a polietilenglicol (C-PEG-(GA)₁₀) por vía subcutánea e intraperitoneal con una mezcla de 5 nmol del oligonucleótido CpG 2006 (Tib Molbiol, Berlín, Alemania), 150 µL de PBS y 150 µL de adyuvante de Freund incompleto. Los ratones se reforzaron con 30 µg de péptido en PBS después de 6 semanas. Las células hiperinmunes del bazo se fusionaron con la línea celular de mieloma de ratón P3X63Ag8.653 usando procedimientos estándar. Los sobrenadantes fueron seleccionados por ELISA, inmunotransferencia e inmunohistoquímica. Los clones GA 1A12 y GA 5E9 de la subclase IgG1 y el clon GA 5F2 de la subclase IgG2a se subclonaron de manera estable.

Partes alícuotas de proteínas de fusión GST purificadas que contenían GA-DPR (GA)₁₅, GP-DPR (GP)₁₅, GR-DPR (GR)₁₅, AP-DPR (AP)₁₅, PR-DPR (PR)₁₅ se sometieron a electroforesis y se realizó la transferencia Western usando los anticuerpos indicados (GA 5E9, GA 5F2, GA 1A12). Las proteínas de fusión GST que contenían GP-DPR, GR-DPR, AP-DPR, PR-DPR se usaron como control para la especificidad de unión a GA. Para obtener más información sobre la afinidad de los anticuerpos utilizados, se realizó un ensayo ELISA de acuerdo con el siguiente procedimiento experimental:

Tomados en conjunto, estos datos demuestran que la traducción de la expansión de repetición de GGGGCC intrónica en pacientes con FTL/ALS conduce a la acumulación de agregados de DPR insolubles. Poly-GA, la DPR más hidrófoba, se detectó de manera más sólida tras la transfección de repeticiones de GGGGCC en células HEK293 y en el ensayo de trampa de filtro de cerebelo humano. Además, los anticuerpos poli-GA tiñeron más del 90% de las enigmáticas inclusiones positivas para p62/negativas para TDP-43 encontradas exclusivamente en casos de FTL/ALS con mutación *C9orf72*. En menor medida, las proteínas poli-GP y poli-GR también fueron detectables en estos agregados. Ninguno de los productos está codificado por un marco de lectura abierto.

Una amplia evidencia sugiere un papel patogénico de las inclusiones de DPR en pacientes con FTL con expansión de repetición de hexanucleótidos de *C9orf72*. Primero, la patología de DPR es predominante en regiones cerebrales clínicamente relevantes (hipocampo y neocorteza frontotemporal) y probablemente precede a la patología de TDP-43, porque ocasionalmente se observó poli-GA en el núcleo de las inclusiones de TDP-43, pero nunca se encontró tinción de fosfo-TDP-43 dentro de las inclusiones de DPR. En segundo lugar, los pacientes con *C9orf72* mostraron atrofia cerebelosa que no ocurre en otras variantes genéticas o esporádicas de FTL/ALS que carecen de inclusiones DPR cerebelosas. En tercer lugar, la patología de DPR es una consecuencia directa de la expansión de hexanucleótidos, la causa genética más común de FTL/ALS. Finalmente, al menos un portador de la expansión de *C9orf72* tenía abundante patología de DPR y la variante de comportamiento clínico de FTL, pero no había patología detectable de inclusión de TDP-43. Por lo tanto, los anticuerpos específicos de DPR serán una herramienta valiosa para diseccionar aún más el papel de las abundantes inclusiones negativas de TDP-43 en diferentes regiones del cerebro.

[0146] Las proteínas de repetición de dipéptidos son la proteína principal de la enfermedad en las inclusiones características corticales y cerebelosas en pacientes con FTL/ALS con expansión de hexanucleótidos de *C9orf72* y están directamente relacionadas con la patología predominante. El mecanismo de traducción inusual y el producto altamente anormal facilitan enfoques diagnósticos y/o terapéuticos más selectivos que los posibles para otras enfermedades neurodegenerativas. Los enfoques diagnósticos y/o terapéuticos adecuados incluyen la aplicación y/o administración de oligonucleótidos, anticuerpos o fragmentos de unión a antígeno de los mismos, proteínas similares a anticuerpos y/o peptidomiméticos dirigidos contra las repeticiones de hexanucleótidos o la proteína DPR.

Referencias

1. Al-Sarraj S, King A, Troakes C et al., (2011) p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of *C9orf72*-linked FTL and MND/ALS. Acta Neuropathol 122 (6): 691-702.

2. Arai T, Hasegawa M, Akiyama H et al. (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun 351 (3): 602-611

3. Ashizawa T., Ranun L.P. (2012) GGCCTG repeats put a hex on Purkinje cells and motor neurons in SCA36, *Neurology* 79(4), 302-303.
- 5 4. Benajiba L, Le Ber I, Camuzat A et al., (2009) TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration. *Ann Neurol* 65 (4): 470-473.
5. Bigio EH, Weintraub S, Rademakers R, Baker M, Ahmadian SS, Rademaker A, Weitner BB, Mao Q, Lee KH, Mishra M, Ganti RA, Mesulam MM. (2012) Frontotemporal lobar degeneration with TDP-43 proteinopathy and chromosome 9p repeat expansion in C9ORF72: clinicopathologic correlation. *Neuropathology* doi: 10.1111/j.1440-1789.2012.01332.x
- 10 6. Boxer AL, Mackenzie IR, Boeve BF, Baker M, Seeley WW, Crook R, Feldman H, Hsiung GY, Rutherford N, Laluz V, Whitwell J, Foti D, McDade E, Molano J, Karydas A, Wojtas A, Goldman J, Mirsky J, Sengdy P, Dearmond S, Miller BL, Rademakers R. (2011) Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J Neurol Neurosurg Psychiatry* 82(2): 196-203.
- 15 7. DeJesus-Hernandez M, Mackenzie IR, Boeve BF et al. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72 (2): 245-256.
8. Desmyter et al. (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Structure Biol.* 3, 803-811
9. Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 11 (5): 1475-1489.
- 20 10. Gasser T, Hardy J, Mizuno Y (2011) Milestones in PD genetics. *Mov Disord* 26 (6): 1042-1048.
11. Geall et al. (2012) Non-viral delivery of self-amplifying RNA vaccines. *Proc Natl Acad Sci USA*, 109, 14604-14609.
12. Gijssels I, Van Langenhove T, van der Zee J et al. (2012) A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol* 11 (1): 54-65.
- 25 13. Haass C, Selkoe DJ (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat Rev Mol Cell Biol* 8 (2): 101-112.
14. Huston et al. (1988) Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 85, 5879-5883
- 30 15. Ivanov IP, Firth AE, Michel AM, Atkins JF, Baranov PV. (2011) Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res* 39, 4220.
16. Josephs KA, Hodges JR, Snowden JS, Mackenzie IR, Neumann M, Mann DM, Dickson DW (2011) Neuropathological background of phenotypical variability in frontotemporal dementia. *Acta Neuropathol* 122 (2): 137-153.
- 35 17. Kremmer E, Kranz BR, Hille A, Klein K, Eulitz M, Hoffmann-Fezer G, Feiden W, Herrmann K, Delecluse HJ, Delsol G, Bornkamm GW, Mueller-Lantzsch N, Grässert FA. (1995) Rat monoclonal antibodies differentiating between the Epstein-Barr virus nuclear antigens 2A (EBNA2A) and 2B (EBNA2B). *Virology* 208(1): 336-42.
18. Kufer P., Lutterbüse R., Baeuerle P.A. (2004) A revival of bispecific antibodies. *Trends Biotechnol.* 22, 238-244.
19. Li H, Wyman T, Yu ZX, Li SH, Li XJ. (2003) Abnormal association of mutant huntingtin with synaptic vesicles inhibits glutamate release. *Hum Mol Genet* 12, 2021.
- 40 20. Ma AS, Moran-Jones K, Shan J et al. (2002) Heterogeneous nuclear ribonucleoprotein A3, a novel RNA trafficking response element-binding protein. *J Biol Chem* 277 (20): 18010-18020.
21. Mackenzie IR, Rademakers R, Neumann M. (2010) TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol.* 9(10): 995-1007.
- 45 22. Mahoney CJ, Downey LE, Ridgway GR, Beck J, Clegg S, Blair M, Finnegan S, Leung KK, Yeatman T, Golden H, Mead S, Rohrer JD, Fox NC, Warren JD (2012) Longitudinal neuroimaging and neuropsychological profiles of frontotemporal dementia with C9ORF72 expansions. *Alzheimers Res Ther.* Septiembre 24 de 2012; 4(5): 41.
23. Neumann M, Sampathu DM, Kwong LK et al. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314 (5796): 130-133.

24. Peabody DS (1989) Translation initiation at non-AUG triplets in mammalian cells. *J Biol Chem* 264, 5031.
25. Rademakers R, Neumann M, Mackenzie IR (2012) Advances in understanding the molecular basis of frontotemporal dementia. *Nat Rev Neurol* 8 (8): 423-434
26. Ranum LP, Cooper TA (2006) RNA-mediated neuromuscular disorders. *Annu Rev Neurosci* 29: 259-277.
- 5 27. Renton AE, Majounie E, Waite A et al. (2011) A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* 72 (2): 257-268.
28. Sieben A, Van Langenhove T, Engelborghs S et al. (2012) The genetics and neuropathology of frontotemporal lobar degeneration. *Acta Neuropathol* 124 (3): 353-372.
- 10 29. Sreedharan J, Blair IP, Tripathi VB et al. (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 319 (5870): 1668-1672
30. Touriol C. et al. (2003) Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol Cell* 95, 169.
31. van der Zee J, Gijssels I, Dillen L et al. (2012) A Pan-European study of the C9orf72 Repeat Associated with FTL: Geographic Prevalence, Genomic Instability and Intermediate Repeats. *Hum Mutat* doi: 10.1002/humu.22244.
- 15 32. Vance C, Rogelj B, Hortobagyi T et al. (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323 (5918): 1208-1211
- 20 33. Whitwell JL, Weigand SD, Boeve BF, Senjem ML, Gunter JL, DeJesus-Hernandez M, Rutherford NJ, Baker M, Knopman DS, Wszolek ZK, Parisi JE, Dickson DW, Petersen RC, Rademakers R, Jack CR Jr, Josephs KA. (2012) Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain* 135(Pt 3): 794-806.
34. Zu T, Gibbens B, Doty NS et al. (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci USA* 108 (1): 260-265.

Listado de secuencias

- <110> DZNE Deutsches Zentrum für Neurodegenerative Erkrankungen e.V. (DZNE)
- 25 <120> Proteínas de repetición de dipéptidos como objetivos terapéuticos en enfermedades neurodegenerativas
- <130> 820-1 PCT
- <150> PCT/EP2013/000190
- <151> 2013-01-22
- <150> EP13185419.2
- 30 <151> 2013-09-20
- <160> 85
- <170> PatentIn versión 3.5
- <210> 1
- <211> 888
- 35 <212> ADN
- <213> Homo sapiens
- <220>
- <221> región de repetición
- <222> (321)..(327)
- 40 <223>/nota="el hexanucleótido ggggcc se repite n veces"
- <400> 1

ES 2 744 804 T3

```

acgtaaccta cggtgtcccg ctaggaaaga gaggtgcgtc aaacagcgac aagttccgcc      60
cacgtaaaaag atgacgcttg gtgtgtcagc cgtccctgct gcccggttgc ttctcttttg      120
ggggcgggggt ctagcaagag caggtgtggg tttaggagggt gtgtgttttt gtttttccca      180
ccctctctcc ccactacttg ctctcacagt actcgctgag ggtgaacaag aaaagacctg      240
ataaagatta accagaagaa aacaaggagg gaaacaaccg cagcctgtag caagctctgg      300
aactcaggag tcgcgcgcta ggggccgggg cgtggtcggg gcgggcccgg gggcggggcc      360
ggggcggggc tgcggttgcg gtgcctgcgc ccgcggcggc ggaggcgag gcggtggcga      420
gtgggtgagt gaggagggcg catcctggcg ggtggctggt tggggttcgg ctgccgggaa      480
gaggcgcggg tagaagcggg ggctctcctc agagctcgac gcatttttac tttccctctc      540
atctctctga ccgaagctgg gtgtcgggct ttcgcctcta gcgactggtg gaattgcctg      600
catccgggcc ccgggcttcc cggcggcggc ggcggcggcg gcggcgcagg gacaagggat      660
ggggatctgg cctcttctt gctttcccgc cctcagtacc cgagctgtct ccttcccggg      720
gaccgcctgg gagcctgcc gctgcgggct cgagaaaagg gagcctcggg tactgagagg      780
cctcgcctgg ggaagggccg gaggggtggc ggcgcgcggc ttctgcggac caagtccggg      840
ttcgctagga acccgagacg gtccctgccg gcgaggagat catgcggg      888

```

<210> 2

<211> 1196

<212> ADN

5 <213> Homo sapiens

<220>

<221> región de repetición

<222> (590)..(596)

<223> /nota="el hexanucleótido ggctg se repite n veces"

10 <400> 2

ES 2 744 804 T3

Gly Arg Gly Arg Gly Gly Pro Gly Gly Gly Pro Gly Ala Gly Leu Arg
1 5 10 15

Leu Arg Cys Leu Arg Pro Arg Arg Arg Arg Arg Arg Arg Trp Arg Val
20 25 30

Gly Glu

<210> 5

<211> 54

<212> PRT

5 <213> Homo sapiens

<400> 5

Gly Val Val Gly Ala Gly Pro Gly Ala Gly Pro Gly Arg Gly Cys Gly
1 5 10 15

Cys Gly Ala Cys Ala Arg Gly Gly Gly Gly Ala Gly Gly Gly Glu Trp
20 25 30

Val Ser Glu Glu Ala Ala Ser Trp Arg Val Ala Val Trp Gly Ser Ala
35 40 45

Ala Gly Lys Arg Arg Gly
50

<210> 6

<211> 23

10 <212> PRT

<213> Homo sapiens

<400> 6

Ser Ala Arg Leu Leu Ser Ser Arg Ala Cys Tyr Arg Leu Arg Leu Phe
1 5 10 15

Pro Ser Leu Phe Ser Ser Gly
20

<210> 7

15 <211> 6

<212> PRT

<213> Homo sapiens

<400> 7

Pro Leu Ala Arg Asp Ser
1 5

20 <210> 8

ES 2 744 804 T3

<211> 106

<212> PRT

<213> Homo sapiens

<400> 8

Arg Leu Arg Leu Arg Leu Pro Trp Glu Arg Val Pro Ala Asp Ala Glu
1 5 10 15

Val Ala Leu Thr Leu Ala Pro Arg Leu Pro Phe Gln Val Leu Leu His
20 25 30

Val Leu Phe Glu His Ala Val Gly Tyr Ala Leu Leu Ala Leu Glu Ile
35 40 45

Ser Leu Leu Gln Pro Gln Val Gly Glu Ile Arg Gly Leu Leu Trp Arg
50 55 60

Pro Arg Arg Pro Ser Ser Arg Gly Ser Gln His Ala Gln Arg Ala Pro
65 70 75 80

Thr Trp Pro Ala Glu Arg Phe Gly Ala Gly Arg Ala Trp Lys Ala Leu
85 90 95

5 Asp Pro Gly Val Phe Thr Leu Thr His Gly
100 105

<210> 9

<211> 20

<212> PRT

<213> Homo sapiens

10 <400> 9

Ala Cys Ala Cys Ala Cys Ala Cys Pro Gly Asn Gly Phe Arg Gln Thr
1 5 10 15

Leu Arg Leu Arg
20

<210> 10

<211> 16

<212> PRT

15 <213> Homo sapiens

<400> 10

Pro Ala Pro Ala Pro Ala Pro Ala Leu Gly Thr Gly Ser Gly Arg Arg
1 5 10 15

<210> 11

ES 2 744 804 T3

<211> 123

<212> PRT

<213> Homo sapiens

<400> 11

Ala Leu Ser Ala Ala Arg Thr Gln Ala Glu Thr Pro Thr Val Ala Ser
1 5 10 15

Pro Ala Pro Arg Asn His Ser Ser Pro Trp Arg Gln Leu Pro Gly Ser
20 25 30

Ala Arg Asn Ala Ala Ser Ala Leu Arg Arg Ala Glu Pro Gly Ser Ala
35 40 45

Pro Ala Pro Pro Thr Asn Arg Arg Pro Gly Thr Thr Pro Ser Pro Ala
50 55 60

Phe Gly Gly Ala Ala Gln Gly Gly Ile Ala Arg Ser Ala Gly Trp Val
65 70 75 80

Val Pro Ser Ala Gly Glu Arg Ala Phe Ser Gly Leu Cys Ser Cys Ala
85 90 95

Gly Gly Arg Gly Cys Arg Gly Arg Ala Ala Pro Gly Arg Ala Leu Gln
100 105 110

Ala Ala Lys Arg Leu Gly Lys Ala Leu Ser Gly
115 120

5

<210> 12

<211> 36

<212> PRT

<213> Homo sapiens

10

<400> 12

Pro Val Cys Gly Pro Asn Ala Gly Arg Asn Pro His Arg Arg Val Ala
1 5 10 15

Arg Ala Pro Gln Pro Leu Leu Thr Met Ala Pro Ala Pro Gly Phe Gly
20 25 30

Ser Gln Cys Gly
35

<210> 13

<211> 136

15

<212> PRT

ES 2 744 804 T3

<213> Homo sapiens

<400> 13

Pro Cys Leu Arg Pro Glu Arg Arg Pro Lys Pro Pro Pro Ser Arg Arg
1 5 10 15

Pro Arg Pro Ala Thr Thr Pro His His Gly Ala Ser Ser Arg Val Arg
20 25 30

Leu Ala Met Arg Leu Ala Arg Ser Gly Ala Arg Asn Gln Ala Pro Glu
35 40 45

Ala Thr Pro Pro Pro Pro Arg Pro Ile Gly Gly Arg Gly Arg His His
50 55 60

Leu Arg Pro Leu Glu Gly Pro Pro Arg Val Val Ser Pro Gly Ala Arg
65 70 75 80

Ala Gly Trp Cys Arg Ala Arg Ala Ser Gly Pro Phe Pro Ala Ser Ala
85 90 95

Ala Ala Arg Ala Gly Val Val Ala Gly Asp Ala Arg Pro Pro Val Gly
100 105 110

His Cys Arg Arg Arg Asn Asp Leu Gly Arg Pro Phe Arg Gly Asn Ser
115 120 125

Pro Val Pro Asp Val Leu Val Ile
130 135

<210> 14

5 <211> 10

<212> PRT

<213> Homo sapiens

<400> 14

Gln Ala Leu Glu Leu Arg Ser Arg Ala Leu
1 5 10

10 <210> 15

<211> 121

<212> PRT

<213> Homo sapiens

<400> 15

ES 2 744 804 T3

Gly Arg Glu Ser Lys Glu Glu Ala Arg Ser Pro Ser Leu Val Pro Ala
1 5 10 15

Pro Pro Pro Pro Pro Pro Pro Pro Gly Ser Pro Gly Pro Gly Cys Arg
20 25 30

Gln Phe His Gln Ser Leu Glu Ala Lys Ala Arg His Pro Ala Ser Val
35 40 45

Arg Glu Met Arg Gly Lys Val Lys Met Arg Arg Ala Leu Arg Arg Ala
50 55 60

Pro Ala Ser Thr Arg Ala Ser Ser Arg Gln Pro Asn Pro Lys Gln Pro
65 70 75 80

Pro Ala Arg Met Pro Pro Pro His Ser Pro Thr Arg His Arg Leu Arg
85 90 95

Leu Arg Arg Arg Gly Arg Arg His Arg Asn Arg Ser Pro Ala Pro Gly
100 105 110

Pro Pro Pro Gly Pro Pro Arg Pro Arg
115 120

<210> 16

<211> 25

<212> PRT

5 <213> Homo sapiens

<400> 16

Arg Leu Thr Arg Arg Lys Gln Gly Gly Lys Gln Pro Gln Pro Val Ala
1 5 10 15

Ser Ser Gly Thr Gln Glu Ser Arg Ala
20 25

<210> 17

<211> 59

10 <212> PRT

<213> Homo sapiens

<400> 17

Gly Glu Pro Pro Leu Leu Pro Ala Pro Leu Pro Gly Ser Arg Thr Pro
1 5 10 15

Asn Ser His Pro Pro Gly Cys Arg Leu Leu Thr His Pro Leu Ala Thr
20 25 30

ES 2 744 804 T3

Ala Cys Ala Ser Ala Ala Ala Gly Ala Gly Thr Ala Thr Ala Ala Pro
 35 40 45

Pro Arg Ala Arg Pro Arg Ala Arg Pro Asp His
 50 55

<210> 18

<211> 160

<212> PRT

5 <213> Homo sapiens

<400> 18

Ser Pro Arg Arg Gln Gly Pro Ser Arg Val Pro Ser Glu Pro Arg Leu
 1 5 10 15

Gly Pro Gln Lys Pro Arg Ala Ala His Pro Pro Ala Phe Pro Gln Ala
 20 25 30

Arg Pro Leu Ser Thr Arg Gly Ser Leu Phe Ser Ser Pro Gln Arg Gln
 35 40 45

Arg Ser Gln Arg Val Pro Gly Lys Glu Thr Ala Arg Val Leu Arg Ala
 50 55 60

Gly Lys Gln Ala Arg Met Gln Ala Ile Pro Pro Val Ala Arg Gly Glu
 65 70 75 80

Ser Pro Thr Pro Ser Phe Gly Gln Arg Asn Glu Arg Glu Ser Lys Asn
 85 90 95

Ala Ser Ser Ser Glu Glu Ser Pro Arg Phe Tyr Pro Arg Leu Phe Pro
 100 105 110

Ala Ala Glu Pro Gln Thr Ala Thr Arg Gln Asp Ala Ala Ser Ser Leu
 115 120 125

Thr His Ser Pro Pro Pro Ala Pro Pro Pro Pro Arg Ala Gln Ala Pro
 130 135 140

Gln Pro Gln Pro Arg Pro Gly Pro Ala Pro Gly Pro Ala Pro Thr Thr
 145 150 155 160

<210> 19

<211> 22

10 <212> PRT

<213> Homo sapiens

ES 2 744 804 T3

<400> 19

Gly Val Val Ala Gly Arg Gly Arg Arg Asp Gly Gly Gly Phe Gly Leu
 1 5 10 15

Arg Ser Gly Arg Arg Gln
 20

<210> 20

<211> 134

5 <212> PRT

<213> Homo sapiens

<400> 20

Gln Glu His Gln Glu Pro Asp Tyr Tyr Pro Glu Arg Ala Phe Pro Ser
 1 5 10 15

Arg Phe Ala Ala Cys Ser Ala Arg Pro Gly Ala Ala Arg Pro Arg Gln
 20 25 30

Pro Arg Pro Pro Ala Gln Leu Gln Arg Pro Glu Lys Ala Arg Ser Pro
 35 40 45

Ala Leu Gly Thr Thr Gln Pro Ala Leu Arg Ala Ile Pro Pro Trp Ala
 50 55 60

Ala Pro Pro Lys Ala Gly Asp Gly Val Val Pro Gly Leu Arg Leu Val
 65 70 75 80

Gly Gly Ala Gly Ala Trp Pro Leu Glu Pro Gly Ser Ala Arg Arg Ser
 85 90 95

Ala Leu Ala Ala Leu Arg Ala Glu Pro Gly Ser Trp Arg His Gly Glu
 100 105 110

Glu Trp Leu Arg Gly Ala Gly Asp Ala Thr Val Gly Val Ser Ala Cys
 115 120 125

Val Arg Ala Ala Asp Arg
 130

<210> 21

10 <211> 36

<212> PRT

<213> Homo sapiens

<400> 21

ES 2 744 804 T3

Pro His Cys Glu Pro Asn Pro Gly Ala Gly Ala Met Val Arg Ser Gly
1 5 10 15

Cys Gly Ala Arg Ala Thr Arg Arg Trp Gly Phe Arg Pro Ala Phe Gly
20 25 30

Pro Gln Thr Gly
35

<210> 22

<211> 67

5 <212> PRT

<213> Homo sapiens

<400> 22

Gly Ser Ala Gly Pro Pro Lys Glu Pro Thr Asp Leu Thr His Leu Arg
1 5 10 15

Leu Gln Gln Thr Asp Leu Leu His Phe Leu Gln Arg Gln Gln Arg Val
20 25 30

Ala Asp Arg Val Leu Lys Gln His Val Gln Gln His Leu Glu Arg Glu
35 40 45

Pro Gly Arg Glu Arg Gln Arg Asn Leu Ser Val Cys Arg Asn Pro Phe
50 55 60

Pro Gly Gln
65

<210> 23

10 <211> 53

<212> PRT

<213> Homo sapiens

<400> 23

ES 2 744 804 T3

Ser Pro Pro Leu Pro Ser Ala Pro Ala Ala Arg Ser Arg Pro Arg Ala
1 5 10 15

Gln Thr Ala Arg Ala Thr Ala Pro Gly Thr Gly Ala Gly Ala Arg Ala
20 25 30

Ser Thr Gln Pro Gln Arg Leu Pro Glu Pro Val Pro Arg Ala Gly Ala
35 40 45

Gly Ala Gly Ala Gly
50

<210> 24

<211> 41

<212> PRT

5 <213> Homo sapiens

<400> 24

Pro Thr Ala Cys Ser Asn Ser Thr Cys Asn Ser Thr Trp Asn Gly Ser
1 5 10 15

Arg Gly Ala Ser Val Asn Ala Thr Ser Ala Ser Ala Gly Thr Arg Ser
20 25 30

Gln Gly Arg Arg Arg Arg Arg Arg Arg
35 40

<210> 25

<211> 47

10 <212> PRT

<213> Homo sapiens

<220>

<221> REPETICIÓN

<222> (13)..(14)

15 <223> el dipéptido GA se repite m times

<400> 25

Gln Ala Leu Glu Leu Arg Ser Arg Ala Leu Gly Ala Gly Ala Gly Ala
1 5 10 15

Trp Ser Gly Arg Ala Arg Gly Arg Ala Arg Gly Gly Ala Ala Val Ala
20 25 30

Val Pro Ala Pro Ala Ala Ala Glu Ala Gln Ala Val Ala Ser Gly
35 40 45

<210> 26

<211> 40

<212> PRT

<213> Homo sapiens

<220>

5 <221> REPETICIÓN

<222> (3)..(4)

<223> el dipéptido GP se repite o veces

<400> 26

Gly Pro Gly Pro Gly Pro Gly Arg Gly Arg Gly Gly Pro Gly Gly Gly
 1 5 10 15

Pro Gly Ala Gly Leu Arg Leu Arg Cys Leu Arg Pro Arg Arg Arg
 20 25 30

Arg Arg Arg Trp Arg Val Gly Glu
 35 40

10 <210> 27

<211> 128

<212> PRT

<213> Homo sapiens

<220>

15 <221> REPETICIÓN

<222> (125)..(126)

<223> el dipéptido GP se repite o veces

<400> 27

ES 2 744 804 T3

Gly Arg Glu Ser Lys Glu Glu Ala Arg Ser Pro Ser Leu Val Pro Ala
1 5 10 15

Pro Pro Pro Pro Pro Pro Pro Pro Gly Ser Pro Gly Pro Gly Cys Arg
20 25 30

Gln Phe His Gln Ser Leu Glu Ala Lys Ala Arg His Pro Ala Ser Val
35 40 45

Arg Glu Met Arg Gly Lys Val Lys Met Arg Arg Ala Leu Arg Arg Ala
50 55 60

Pro Ala Ser Thr Arg Ala Ser Ser Arg Gln Pro Asn Pro Lys Gln Pro
65 70 75 80

Pro Ala Arg Met Pro Pro Pro His Ser Pro Thr Arg His Arg Leu Arg
85 90 95

Leu Arg Arg Arg Gly Arg Arg His Arg Asn Arg Ser Pro Ala Pro Gly
100 105 110

Pro Pro Pro Gly Pro Pro Arg Pro Arg Pro Gly Pro Gly Pro Gly Pro
115 120 125

<210> 28

<211> 86

5 <212> PRT

<213> Homo sapiens

<220>

<221> REPETICIÓN

<222> (29)..(30)

10 <223> el dipéptido GR se repite p veces

<400> 28

Arg Leu Thr Arg Arg Lys Gln Gly Gly Lys Gln Pro Gln Pro Val Ala
1 5 10 15

Ser Ser Gly Thr Gln Glu Ser Arg Ala Arg Gly Arg Gly Arg Gly Arg

<213> Homo sapiens

<220>

<221> REPETICIÓN

<222> (186)..(187)

5 <223> el dipéptido PR se repite r veces

<400> 30

Ser Pro Arg Arg Gln Gly Pro Ser Arg Val Pro Ser Glu Pro Arg Leu
1 5 10 15

Gly Pro Gln Lys Pro Arg Ala Ala His Pro Pro Ala Phe Pro Gln Ala
20 25 30

Arg Pro Leu Ser Thr Arg Gly Ser Leu Phe Ser Ser Pro Gln Arg Gln
35 40 45

Arg Ser Gln Arg Val Pro Gly Lys Glu Thr Ala Arg Val Leu Arg Ala
50 55 60

Gly Lys Gln Gly Arg Gly Gln Ile Pro Ile Pro Cys Pro Cys Ala Ala
65 70 75 80

Ala Ala Ala Ala Ala Ala Gly Lys Pro Gly Ala Arg Met Gln Ala Ile
85 90 95

Pro Pro Val Ala Arg Gly Glu Ser Pro Thr Pro Ser Phe Gly Gln Arg
100 105 110

Asn Glu Arg Glu Ser Lys Asn Ala Ser Ser Ser Glu Glu Ser Pro Arg
115 120 125

Phe Tyr Pro Arg Leu Phe Pro Ala Ala Glu Pro Gln Thr Ala Thr Arg
130 135 140

Gln Asp Ala Ala Ser Ser Leu Thr His Ser Pro Pro Pro Ala Pro Pro
145 150 155 160

Pro Pro Arg Ala Gln Ala Pro Gln Pro Gln Pro Arg Pro Gly Pro Ala
165 170 175

Pro Gly Pro Ala Pro Thr Thr Pro Arg Pro Arg Pro Arg Pro Leu Ala
180 185 190

Arg Asp Ser
195

<210> 31

<211> 139

<212> PRT

<213> Homo sapiens

5 <220>

<221> REPETICIÓN

<222> (25)..(26)

<223> el dipéptido GL se repite s veces

<400> 31

Gly Val Val Ala Gly Arg Gly Arg Arg Asp Gly Gly Gly Phe Gly Leu
1 5 10 15

Arg Ser Gly Arg Arg Gln Gly Leu Gly Leu Gly Leu Arg Leu Arg Leu
20 25 30

Arg Leu Pro Trp Glu Arg Val Pro Ala Asp Ala Glu Val Ala Leu Thr
35 40 45

Leu Ala Pro Arg Leu Pro Phe Gln Val Leu Leu His Val Leu Phe Glu
50 55 60

His Ala Val Gly Tyr Ala Leu Leu Ala Leu Lys Glu Val Glu Glu Ile
65 70 75 80

Ser Leu Leu Gln Pro Gln Val Gly Glu Ile Arg Gly Leu Leu Trp Arg
85 90 95

Pro Arg Arg Pro Ser Ser Arg Gly Ser Gln His Ala Gln Arg Ala Pro
100 105 110

Thr Trp Pro Ala Glu Arg Phe Gly Ala Gly Gly Arg Ala Trp Lys Ala
115 120 125

Leu Asp Pro Gly Val Phe Thr Leu Thr His Gly
130 135

10

<210> 32

<211> 160

<212> PRT

<213> Homo sapiens

15 <220>

<221> REPETICIÓN

<222> (137)..(138)

ES 2 744 804 T3

<223> el dipéptido AW se repite t veces

<400> 32

Gln Glu His Gln Glu Pro Asp Tyr Tyr Pro Glu Arg Ala Phe Pro Ser
1 5 10 15

Arg Phe Ala Ala Cys Ser Ala Arg Pro Gly Ala Ala Arg Pro Arg Gln
20 25 30

Pro Arg Pro Pro Ala Gln Leu Gln Arg Pro Glu Lys Ala Arg Ser Pro
35 40 45

Ala Leu Gly Thr Thr Gln Pro Ala Leu Arg Ala Ile Pro Pro Trp Ala
50 55 60

Ala Pro Pro Lys Ala Gly Asp Gly Val Val Pro Gly Leu Arg Leu Val
65 70 75 80

Gly Gly Ala Gly Ala Trp Pro Leu Glu Pro Gly Ser Ala Arg Arg Ser
85 90 95

Ala Leu Ala Ala Leu Arg Ala Glu Pro Gly Ser Trp Arg His Gly Glu
100 105 110

Glu Trp Leu Arg Gly Ala Gly Asp Ala Thr Val Gly Val Ser Ala Cys
115 120 125

Val Arg Ala Ala Asp Arg Ala Trp Ala Trp Ala Trp Ala Cys Ala Cys
130 135 140

Ala Cys Ala Cys Pro Gly Asn Gly Phe Arg Gln Thr Leu Arg Leu Arg
145 150 155 160

5 <210> 33

<211> 58

<212> PRT

<213> Homo sapiens

<220>

10 <221> REPETICIÓN

<222> (39)..(40)

<223> el dipéptido PG se repite u veces

<400> 33

ES 2 744 804 T3

Pro His Cys Glu Pro Asn Pro Gly Ala Gly Ala Met Val Arg Ser Gly
1 5 10 15

Cys Gly Ala Arg Ala Thr Arg Arg Trp Gly Phe Arg Pro Ala Phe Gly
20 25 30

Pro Gln Thr Gly Pro Gly Pro Gly Pro Gly Pro Ala Pro Ala Pro Ala
35 40 45

Pro Ala Leu Gly Thr Gly Ser Gly Arg Arg
50 55

<210> 34

<211> 199

<212> PRT

5 <213> Homo sapiens

<220>

<221> REPETICIÓN

<222> (70)..(71)

<223> el dipéptido AQ se repite w veces

10 <400> 34

ES 2 744 804 T3

Gly Ser Ala Gly Pro Pro Lys Glu Pro Thr Asp Leu Thr His Leu Arg
 1 5 10 15

Leu Gln Gln Thr Asp Leu Leu His Phe Leu Gln Arg Gln Gln Arg Val
 20 25 30

Ala Asp Arg Val Leu Lys Gln His Val Gln Gln His Leu Glu Arg Glu
 35 40 45

Pro Gly Arg Glu Arg Gln Arg Asn Leu Ser Val Cys Arg Asn Pro Phe
 50 55 60

Pro Gly Gln Ala Gln Ala Gln Ala Gln Ala Leu Ser Ala Ala Arg Thr
 65 70 75 80

Gln Ala Glu Thr Pro Thr Val Ala Ser Pro Ala Pro Arg Asn His Ser
 85 90 95

Ser Pro Trp Arg Gln Leu Pro Gly Ser Ala Arg Asn Ala Ala Ser Ala
 100 105 110

Leu Arg Arg Ala Glu Pro Gly Ser Arg Gly His Ala Pro Ala Pro Pro
 115 120 125

Thr Asn Arg Arg Pro Gly Thr Thr Pro Ser Pro Ala Phe Gly Gly Ala
 130 135 140

Ala Gln Gly Gly Ile Ala Arg Ser Ala Gly Trp Val Val Pro Ser Ala
 145 150 155 160

Gly Glu Arg Ala Phe Ser Gly Leu Cys Ser Cys Ala Gly Gly Arg Gly
 165 170 175

Cys Arg Gly Arg Ala Ala Pro Gly Arg Ala Leu Gln Ala Ala Lys Arg
 180 185 190

Leu Gly Lys Ala Leu Ser Gly
 195

<210> 35

<211> 94

<212> PRT

5 <213> Homo sapiens

<220>

<221> REPETICIÓN

ES 2 744 804 T3

<222> (56)..(57)

<223> el dipéptido GP se repite x veces

<400> 35

Ser Pro Pro Leu Pro Ser Ala Pro Ala Ala Arg Ser Arg Pro Arg Ala
1 5 10 15

Gln Thr Ala Arg Ala Thr Ala Pro Gly Thr Gly Ala Gly Ala Arg Ala
20 25 30

Ser Thr Gln Pro Gln Arg Leu Pro Glu Pro Val Pro Arg Ala Gly Ala
35 40 45

Gly Ala Gly Ala Gly Gly Pro Gly Pro Gly Pro Val Cys Gly Pro Asn
50 55 60

Ala Gly Arg Asn Pro His Arg Arg Val Ala Arg Ala Pro Gln Pro Leu
65 70 75 80

Leu Thr Met Ala Pro Ala Pro Gly Phe Gly Ser Gln Cys Gly
85 90

5 <210> 36

<211> 183

<212> PRT

<213> Homo sapiens

<220>

10 <221> REPETICIÓN

<222> (44)..(45)

<223> el dipéptido PR se repite y veces

<400> 36

Pro Thr Ala Cys Ser Asn Ser Thr Cys Asn Ser Thr Trp Asn Gly Ser
1 5 10 15

Arg Gly Ala Ser Val Asn Ala Thr Ser Ala Ser Ala Gly Thr Arg Ser
20 25 30

Gln Gly Arg Arg Arg Arg Arg Arg Arg Pro Arg Pro Arg Pro Arg Pro
35 40 45

ES 2 744 804 T3

Cys Leu Arg Pro Glu Arg Arg Pro Lys Pro Pro Pro Ser Arg Arg Pro
 50 55 60

Arg Pro Ala Thr Thr Pro His His Gly Ala Ser Ser Arg Val Arg Leu
 65 70 75 80

Ala Met Arg Leu Ala Arg Ser Gly Ala Arg Asn Gln Ala Pro Glu Ala
 85 90 95

Thr Pro Pro Pro Pro Arg Pro Ile Gly Gly Arg Gly Arg His His Leu
 100 105 110

Arg Pro Leu Glu Gly Pro Pro Arg Val Val Ser Pro Gly Ala Arg Ala
 115 120 125

Gly Trp Cys Arg Ala Arg Ala Ser Gly Pro Phe Pro Ala Ser Ala Ala
 130 135 140

Ala Arg Ala Gly Val Val Ala Gly Asp Ala Arg Pro Pro Val Gly His
 145 150 155 160

Cys Arg Arg Arg Asn Asp Leu Gly Arg Pro Phe Arg Gly Asn Ser Pro
 165 170 175

Val Pro Asp Val Leu Val Ile
 180

<210> 37

<211> 20

<212> ADN

5 <213> Homo sapiens

<400> 37

caattccacc agtcgctaga 20

<210> 38

<400> 38

10 000

<210> 39

<211> 18

<212> ADN

<213> Homo sapiens

15 <400> 39

aagaggcgcg ggtagaag 18

<210> 40

ES 2 744 804 T3

<211> 19
<212> ADN
<213> Homo sapiens
<400> 40
5 agtcgctaga ggcgaaagc 19
<210> 41
<211> 21
<212> ADN
<213> Homo sapiens
10 <400> 41
actggaatgg ggatcgagc a 21
<210> 42
<211> 27
<212> ADN
15 <213> Homo sapiens
<400> 42
accctgatct tccattctct ctgtgcc 27
<210> 43
<211> 19
20 <212> ADN
<213> Homo sapiens
<400> 43
ctgcggtgc ggtgcctgc 19
<210> 44
25 <211> 20
<212> ADN
<213> Homo sapiens
<400> 44
agctggagat ggcggtggc 20
30 <210> 45
<211> 23
<212> ADN
<213> Homo sapiens
<400> 45
35 tgaacaaaag acggaaggctg ctg 23
<210> 46
<211> 23

<212> ADN
 <213> Homo sapiens
 <400> 46
 tctgatagga tgggttggtt gca 23

5 <210> 47
 <211> 22
 <212> PRT
 <213> Homo sapiens
 <400> 47

	Gly	Gln	Ser	Arg	Gly	Gln	Ser	Arg	Gly	Arg	Gly	Arg	Gly	Arg	Gly	Arg
	1				5				10				15			

10 Gly Arg Gly Lys Gly Lys
 20

<210> 48
 <211> 20
 <212> ADN
 <213> Homo sapiens

15 <400> 48
 agtactcgct gaggtgaac 20
 <210> 49
 <211> 44
 <212> ADN

20 <213> Homo sapiens
 <400> 49
 cgtacgcac ccagtttgag agccccggcc ccggccccgg cccc 44
 <210> 50
 <211> 10

25 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Secuencia de consenso de CDR3 de cadena pesada
 <220>

30 <221> CARACTERÍSTICA NUEVA
 <222> (3)..(3)
 <223> cualquier aminoácido
 <220>
 <221> CARACTERÍSTICA NUEVA

<222> (4)..(4)

<223> cualquier aminoácido

<220>

<221> CARACTERÍSTICA NUEVA

5 <222> (9)..(9)

<223> cualquier aminoácido

<400> 50

Gly	Asp	Xaa	Xaa	Asn	Ser	His	Phe	Xaa	Tyr
1				5					10

<210> 51

10 <211> 10

<212> PRT

<213> Secuencia Artificial

<220>

<223> Secuencia de consenso de CDR2 de cadena pesada

15 <220>

<221> CARACTERÍSTICA NUEVA

<222> (2)..(2)

<223> cualquier aminoácido

<220>

20 <221> CARACTERÍSTICA NUEVA

<222> (8)..(8)

<223> cualquier aminoácido

<400> 51

Glu	Xaa	Leu	Pro	Gly	Ser	Gly	Xaa	Thr	Lys
1				5					10

25 <210> 52

<211> 10

<212> PRT

<213> Secuencia Artificial

<220>

30 <223> Secuencia de consenso de CDR1 de cadena pesada

<220>

<221> CARACTERÍSTICA NUEVA

<222> (3)..(3)

<223> cualquier aminoácido

35 <220>

<221> CARACTERÍSTICA NUEVA

<222> (5)..(5)

<223> cualquier aminoácido

<400> 52

5 Gly Tyr Xaa Phe Xaa Gly Tyr Trp Ile Glu
1 5 10

<210> 53

<211> 9

<212> PRT

<213> Mus musculus

10 <400> 53

Gln Gln Leu Val Glu Tyr Pro Leu Thr
1 5

<210> 54

<211> 7

<212> PRT

15 <213> Mus musculus

<400> 54

Leu Met Ser Thr Arg Ala Ser
1 5

<210> 55

<211> 16

20 <212> PRT

<213> Mus musculus

<400> 55

Arg Ser Ser Lys Ser Leu Leu Tyr Lys Asp Gly Lys Thr Tyr Leu Asn
1 5 10 15

<210> 56

25 <211> 17

<212> PRT

<213> Secuencia Artificial

<220>

<223> Longitud de la secuencia de consenso de CDR2 de cadena pesada

30 <220>

<221> CARACTERÍSTICA NUEVA

<222> (2)..(2)

<223> X es cualquier aminoácido

<220>

<221> CARACTERÍSTICA NUEVA

<222> (8)..(8)

<223> X es cualquier aminoácido

<220>

5 <221> CARACTERÍSTICA NUEVA

<222> (14)..(14)

<223> X es cualquier aminoácido

<220>

<221> CARACTERÍSTICA NUEVA

10 <222> (16)..(16)

<223> X es cualquier aminoácido

<400> 56

Glu Xaa Leu Pro Gly Ser Gly Xaa Thr Lys Tyr Asn Glu Xaa Phe Xaa
1 5 10 15

Gly

<210> 57

15 <211> 10

<212> PRT

<213> Mus musculus

<400> 57

Gly Asp Tyr Ser Asn Ser His Phe Ala Tyr
1 5 10

20 <210> 58

<211> 10

<212> PRT

<213> Mus musculus

<400> 58

Gly Asp Phe Thr Asn Ser His Phe Thr Tyr
1 5 10

25 <210> 59

<211> 10

<212> PRT

<213> Mus musculus

30 <400> 59

Gly Asp Tyr Ser Asn Ser His Phe Thr Tyr
1 5 10

<210> 60

<211> 10

ES 2 744 804 T3

<212> PRT

<213> Mus musculus

<400> 60

Gly Asp Phe Thr Asn Ser His Phe Ala Tyr
1 5 10

5 <210> 61

<211> 10

<212> PRT

<213> Mus musculus

<400> 61

Gly Asp Phe Ser Asn Ser His Phe Thr Tyr
1 5 10

10

<210> 62

<211> 10

<212> PRT

<213> Mus musculus

15 <400> 62

Gly Asp Phe Ser Asn Ser His Phe Ala Tyr
1 5 10

<210> 63

<211> 10

<212> PRT

20 <213> Mus musculus

<400> 63

Gly Asp Tyr Thr Asn Ser His Phe Ala Tyr
1 5 10

<210> 64

<211> 10

25 <212> PRT

<213> Mus musculus

<400> 64

Glu Ile Leu Pro Gly Ser Gly Ser Thr Lys
1 5 10

<210> 65

30 <211> 10

<212> PRT

<213> Mus musculus

<400> 65

ES 2 744 804 T3

Glu Asn Leu Pro Gly Ser Gly Ser Thr Lys
1 5 10

<210> 66

<211> 10

<212> PRT

5 <213> Mus musculus

<400> 66

Glu Asn Leu Pro Gly Ser Gly Thr Thr Lys
1 5 10

<210> 67

<211> 10

10 <212> PRT

<213> Mus musculus

<400> 67

Glu Ile Leu Pro Gly Ser Gly Thr Thr Lys
1 5 10

<210> 68

15 <211> 10

<212> PRT

<213> Mus musculus

<400> 68

Gly Tyr Thr Phe Thr Gly Tyr Trp Ile Glu
1 5 10

20 <210> 69

<211> 10

<212> PRT

<213> Mus musculus

<400> 69

Gly Tyr Lys Phe Ile Gly Tyr Trp Ile Glu
1 5 10

25 <210> 70

<211> 10

<212> PRT

<213> Mus musculus

30 <400> 70

Gly Tyr Lys Phe Thr Gly Tyr Trp Ile Glu
1 5 10

<210> 71

<211> 10

ES 2 744 804 T3

<212> PRT

<213> Mus musculus

<400> 71

Gly Tyr Thr Phe Ile Gly Tyr Trp Ile Glu
1 5 10

5 <210> 72

<211> 17

<212> PRT

<213> Mus musculus

<400> 72

Glu Ile Leu Pro Gly Ser Gly Ser Thr Lys Tyr Asn Glu Asn Phe Arg
1 5 10 15

10 Gly

<210> 73

<211> 17

<212> PRT

<213> Mus musculus

15 <400> 73

Glu Asn Leu Pro Gly Ser Gly Ser Thr Lys Tyr Asn Glu Lys Phe Lys
1 5 10 15

Gly

<210> 74

<211> 17

<212> PRT

20 <213> Mus musculus

<400> 74

Glu Asn Leu Pro Gly Ser Gly Thr Thr Lys Tyr Asn Glu Asn Phe Arg
1 5 10 15

Gly

<210> 75

<211> 17

25 <212> PRT

<213> Mus musculus

<400> 75

ES 2 744 804 T3

Glu Ile Leu Pro Gly Ser Gly Ser Thr Lys Tyr Asn Glu Asn Phe Lys
1 5 10 15

Gly

<210> 76

<211> 17

<212> PRT

5 <213> Mus musculus

<400> 76

Glu Ile Leu Pro Gly Ser Gly Thr Thr Lys Tyr Asn Glu Asn Phe Arg
1 5 10 15

Gly

<210> 77

<211> 17

10 <212> PRT

<213> Mus musculus

<400> 77

Glu Asn Leu Pro Gly Ser Gly Thr Thr Lys Tyr Asn Glu Lys Phe Lys
1 5 10 15

Gly

<210> 78

15 <211> 17

<212> PRT

<213> Mus musculus

<400> 78

Glu Asn Leu Pro Gly Ser Gly Thr Thr Lys Tyr Asn Glu Asn Phe Lys
1 5 10 15

Gly

20 <210> 79

<211> 110

<212> PRT

<213> Mus musculus

<400> 79

ES 2 744 804 T3

Ile Val Ile Ser Gln Asp Glu Leu Ser Asn Pro Val Thr Ser Gly Glu
1 5 10 15

Ser Val Ser Ile Ser Cys Arg Ser Ser Lys Ser Leu Leu Tyr Lys Asp
20 25 30

Gly Lys Thr Tyr Leu Asn Trp Phe Leu Gln Arg Pro Gly Gln Ser Pro
35 40 45

Gln Leu Leu Ile Tyr Leu Met Ser Thr Arg Ala Ser Gly Val Ser Asp
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Glu Ile Ser
65 70 75 80

Arg Val Lys Ala Glu Asp Val Gly Val Tyr Phe Cys Gln Gln Leu Val
85 90 95

Glu Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu
100 105 110

<210> 80

<211> 115

<212> PRT

5 <213> Mus musculus

<400> 80

Leu Gln Gln Ser Gly Ala Glu Leu Met Lys Pro Gly Ala Ser Val Lys
1 5 10 15

Leu Ser Cys Arg Ser Ser Gly Tyr Thr Phe Thr Gly Tyr Trp Ile Glu
20 25 30

Trp Val Lys Gln Arg Pro Gly His Gly Leu Glu Trp Ile Gly Glu Ile
35 40 45

ES 2 744 804 T3

Leu Pro Gly Ser Gly Ser Thr Lys Tyr Asn Glu Asn Phe Arg Gly Lys
 50 55 60

Ala Thr Phe Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Ile Arg Leu
 65 70 75 80

Ser Ser Leu Thr Thr Glu Asp Ser Ala Ile Tyr Phe Cys Ala Arg Gly
 85 90 95

Asp Phe Thr Asn Ser His Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val
 100 105 110

Thr Val Ser
 115

<210> 81

<211> 115

<212> PRT

5 <213> Mus musculus

<400> 81

Leu Gln Gln Ser Gly Ala Glu Val Met Lys Pro Gly Ala Ser Val Lys
 1 5 10 15

Leu Ser Cys Lys Val Thr Gly Tyr Thr Phe Thr Gly Tyr Trp Ile Glu
 20 25 30

Trp Val Lys Gln Arg Pro Gly His Gly Leu Glu Trp Ile Gly Glu Asn
 35 40 45

Leu Pro Gly Ser Gly Ser Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys
 50 55 60

Ala Thr Phe Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Met Gln Leu
 65 70 75 80

Ser Ser Leu Thr Ile Glu Asp Ser Ala Ile Tyr Tyr Cys Ala Arg Gly
 85 90 95

Asp Tyr Ser Asn Ser His Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val
 100 105 110

Thr Val Ser
 115

<210> 82

<211> 115

ES 2 744 804 T3

<212> PRT

<213> Mus musculus

<400> 82

Leu Gln Gln Ser Gly Val Glu Leu Met Lys Pro Gly Ala Ser Val Lys
1 5 10 15

Leu Ser Cys Lys Ser Thr Gly Tyr Lys Phe Ile Gly Tyr Trp Ile Glu
20 25 30

Trp Val Lys Gln Arg Pro Gly His Gly Leu Glu Trp Ile Gly Glu Asn
35 40 45

Leu Pro Gly Ser Gly Thr Thr Lys Tyr Asn Glu Asn Phe Arg Gly Lys
50 55 60

Ala Thr Phe Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Met Gln Leu
65 70 75 80

Ser Ser Leu Thr Ile Glu Asp Ser Ala Ile Tyr Tyr Cys Ala Arg Gly
85 90 95

Asp Tyr Ser Asn Ser His Phe Thr Tyr Trp Gly Gln Gly Thr Leu Val
100 105 110

Thr Val Ser
115

5 <210> 83

<211> 12

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> Ejemplo de secuencia de repetición de dinucleótidos

<400> 83

cgcgcgcgcg cg 12

<210> 84

<211> 0

15 <212> ADN

<213> Secuencia Artificial

<220>

<223> Ejemplo de secuencia de repetición de hexanucleótidos

<400> 84

000

<210> 85

<211> 12

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Ejemplo de secuencia de repetición de hexanucleótidos

<400> 85

ggtggcggtg gc 12

10

REIVINDICACIONES

- 5 1. Anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido que consiste en repeticiones de dipéptidos (Gly-Ala)_a, en las que a es un número entero de 16 o más en los que el anticuerpo o el fragmento de unión a antígeno del mismo comprende una secuencia de CDR3 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 53; una secuencia de CDR2 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 54; y una secuencia de CDR1 de cadena ligera que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 55
- y
- 10 (i) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende EILPGSGSTK de acuerdo con la SEQ ID NO: 64 y una secuencia de CDR3 que comprende GDFTNSHFAY de acuerdo con la SEQ ID NO: 60;
- o
- 15 (ii) una secuencia de CDR1 de cadena pesada que comprende GYTFTGYWIE de acuerdo con la SEQ ID NO: 68, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGSTK de acuerdo con la SEQ ID NO: 65 y una secuencia de CDR3 que comprende GDYSNSHFAY de acuerdo con la SEQ ID NO: 57;
- o
- (iii) una secuencia de CDR1 de cadena pesada que comprende GYKFIGYWIE de acuerdo con la SEQ ID NO: 69, una secuencia de CDR2 de cadena pesada que comprende ENLPGSGTTK de acuerdo con la SEQ ID NO: 66 y una secuencia de CDR3 que comprende GDYSNSHFTY de acuerdo con la SEQ ID NO: 59
- 20 2. Un producto farmacéutico que comprende el anticuerpo o fragmento de unión al mismo de acuerdo con la reivindicación 1 y/o un ácido nucleico que codifica dicho anticuerpo o fragmento de unión a antígeno del mismo.
- 25 3. Uso de un anticuerpo o fragmento de unión a antígeno del mismo de acuerdo con la reivindicación 1 en el diagnóstico *in vitro* de una enfermedad caracterizado por una expansión de repeticiones genómicas de GGGGCC, en las que la enfermedad se selecciona del grupo que consiste en esclerosis lateral amiotrófica (ALS), demencia frontotemporal (FTD) y esclerosis lateral amiotrófica-demencia frontotemporal (ALS-FTD).

Fig. 1

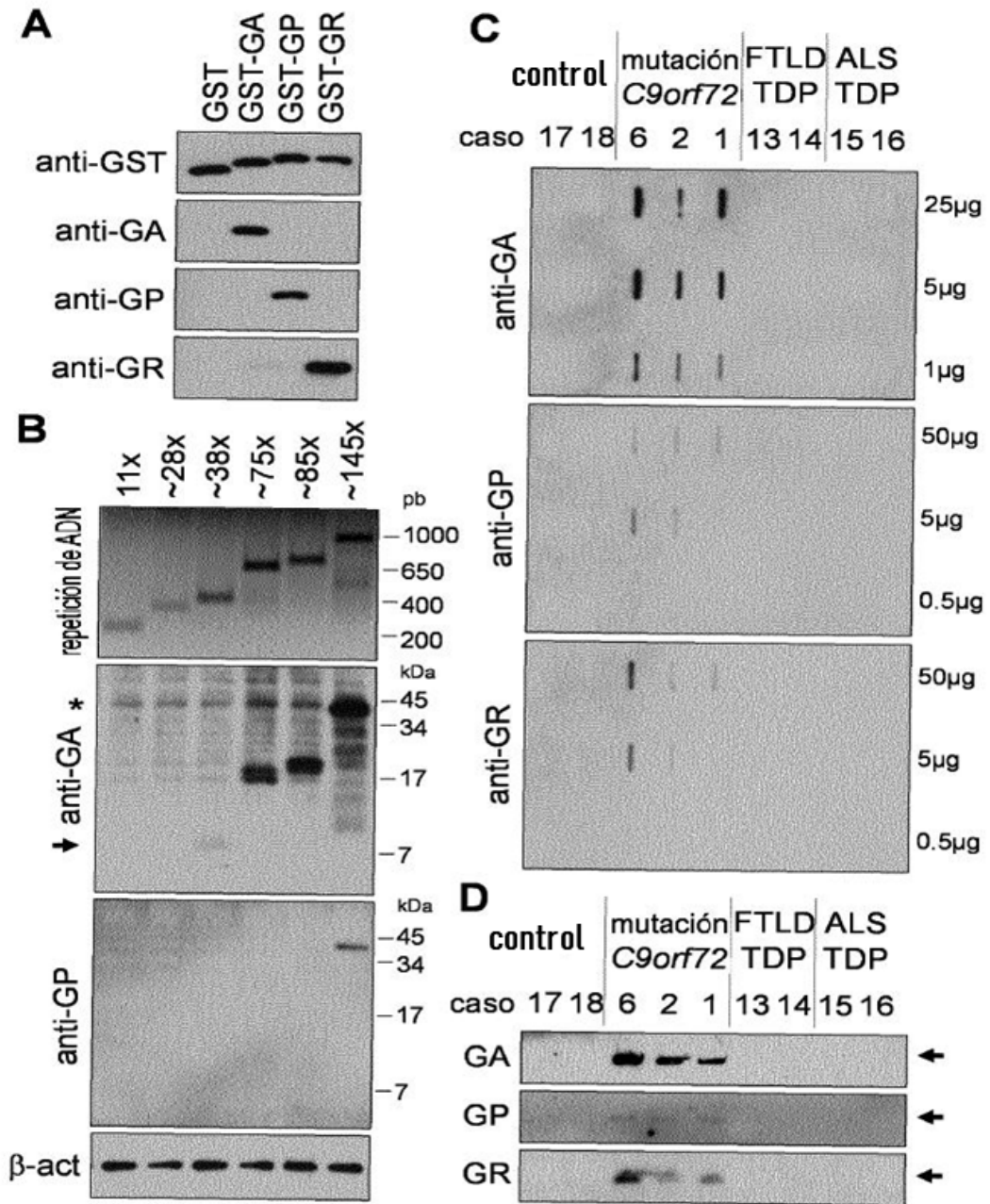


Fig. 2

Caso	Región	DPR			p62	TDP-43	Diagnóstico	Expansión de repetición de GGGGCC	Fuente
		GA	GP	GR					
TJ-1	Hipocampo	+	+	+	+	+	FTLD/ALS-TDP	+	LMU
	Cerebelo	+	+	+	+	-			LMU
TJ-2	Hipocampo	+	+	+	+	+	FTLD/ALS-TDP	+	LMU
	Cerebelo	+	+	+	+	-			LMU
TJ-3	Hipocampo	+	+	+	+	+	FTLD/ALS-TDP	+	LMU
	Cerebelo	+	+	+	+	-			LMU
TJ-4	Hipocampo	+	+	+	+	+	FTLD-TDP	+	LMU
	Cerebelo	+	+	+	+	-			LMU
TJ-5	Hipocampo	+	+	+	+	+	FTLD/ALS-TDP	+	LMU
	Cerebelo	+	+	+	+	-			LMU
TJ-6	Cerebelo	NA	NA	NA	NA	NA	FTLD/ALS-TDP	+	LMU
TJ-10	corteza	+	+	+	+	-	FTLD-UPS	+	VIB
TJ-11	Hipocampo	-	-	-	+	-	Huntington	NA	LMU
	corteza frontal	-	-	-	+	NE			LMU
TJ-12	Hipocampo	-	-	-	+	-	Huntington	NA	LMU
	corteza frontal	-	-	-	+	NE			LMU
TJ-13	Hipocampo	-	-	-	+	+	FTLD-TDP	-	LMU
	Cerebelo	-	-	-	-	-			LMU
TJ-14	Hipocampo	-	-	-	+	+	FTLD-TDP	-	LMU
	Cerebelo	-	-	-	-	-			LMU
TJ-15	Hipocampo	-	-	-	+	+	ALS-TDP	-	LMU
	Cerebelo	-	-	-	-	-			LMU
TJ-16	Hipocampo	-	-	-	+	+	ALS-TDP	-	LMU
	Cerebelo	-	-	-	-	-			LMU
TJ-17	Hipocampo	-	-	-	-	-	Control	-	LMU
	Cerebelo	-	-	-	NE	NE			LMU
TJ-18	Hipocampo	-	-	-	-	-	Control	-	LMU
	Cerebelo	-	-	-	NE	NE			LMU
TJ-21	corteza	-	-	-	-	-	Control	-	VIB

Fig. 3

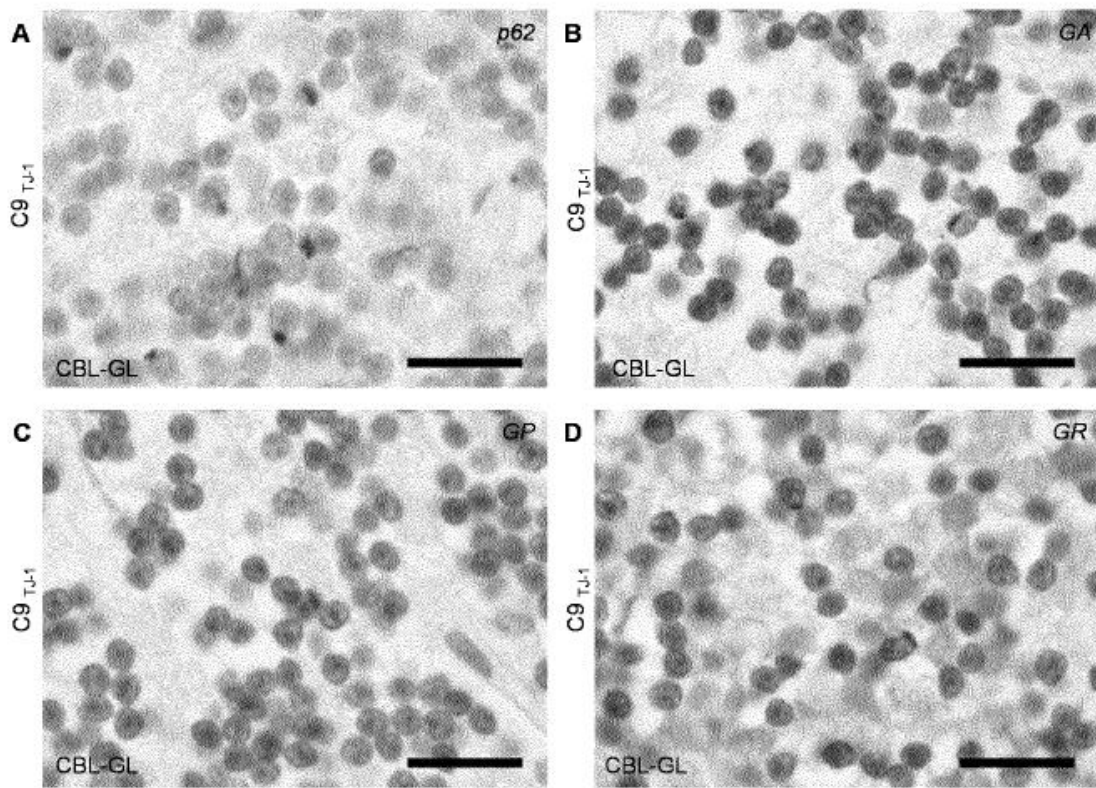


Fig. 4

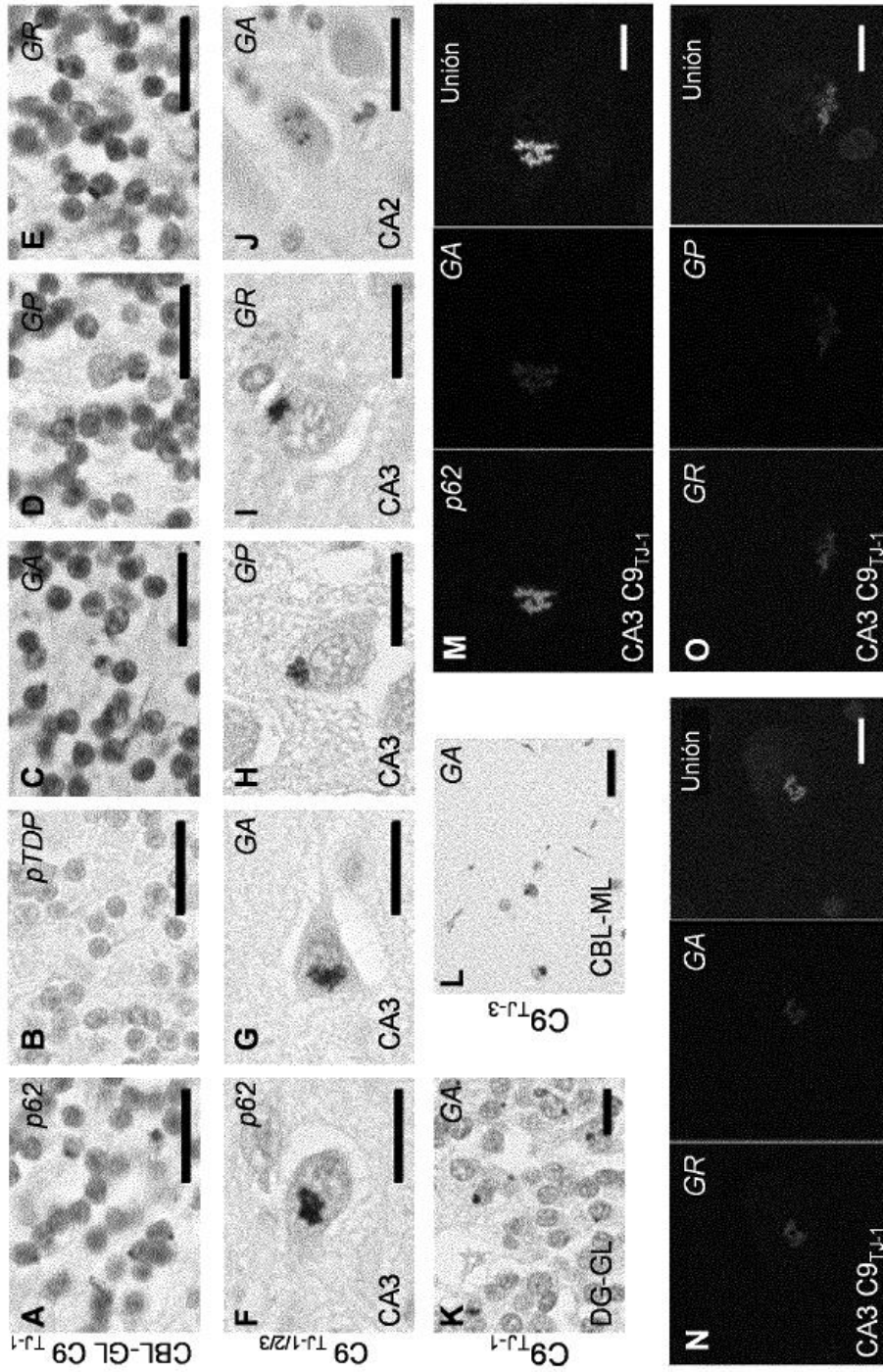


Fig. 5

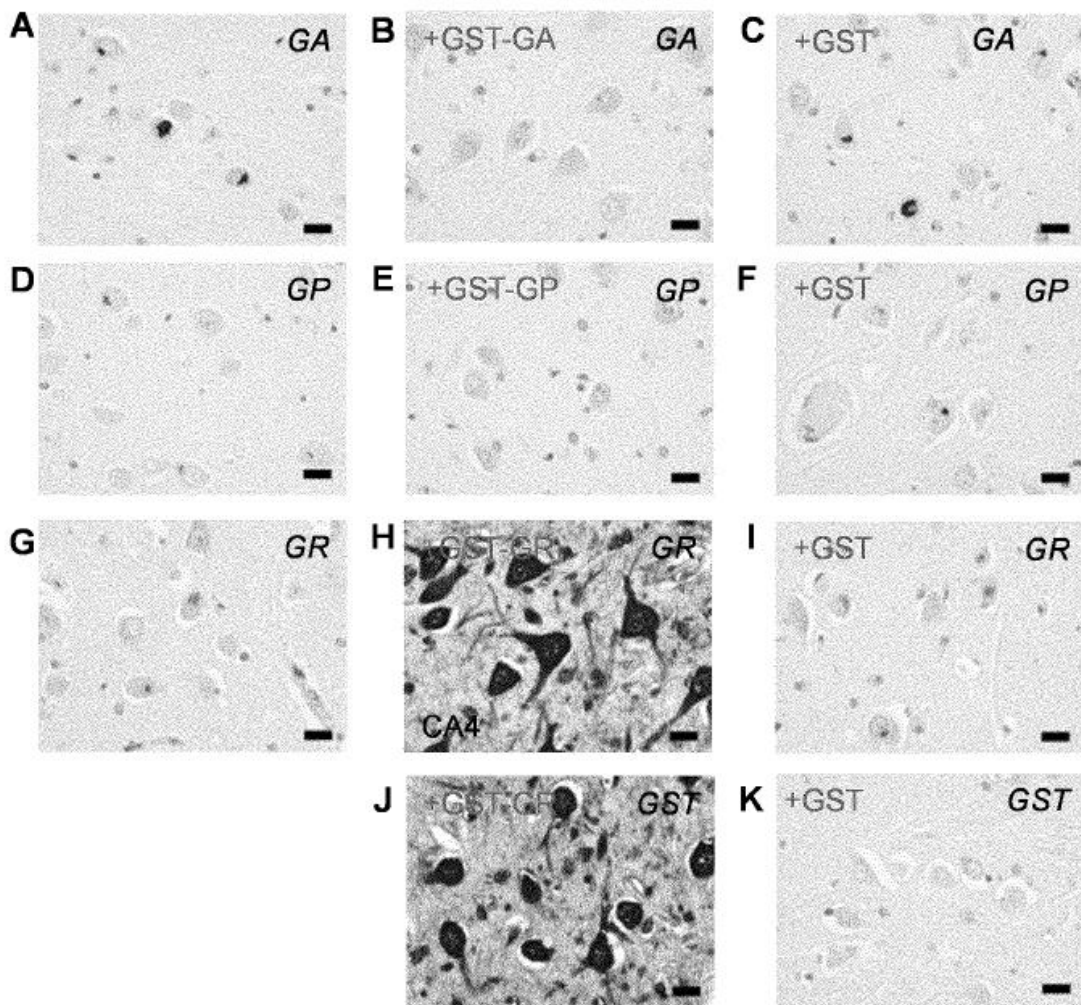


Fig. 6

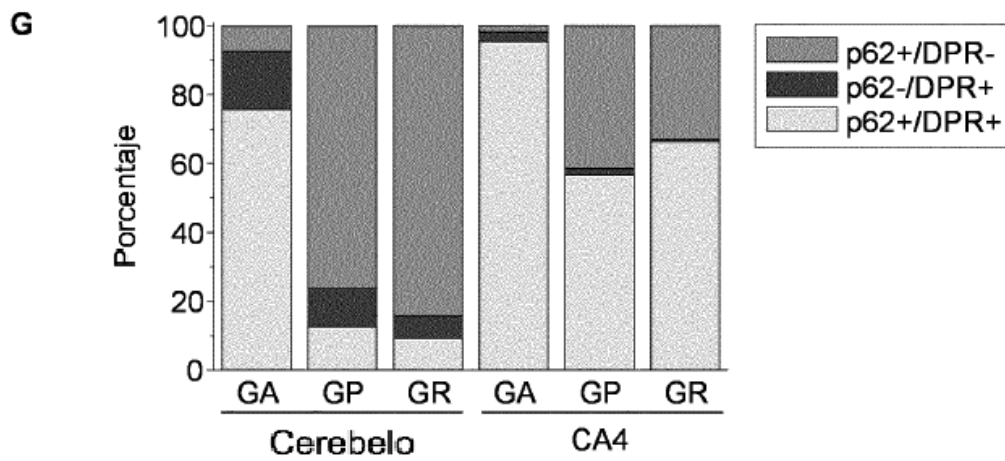
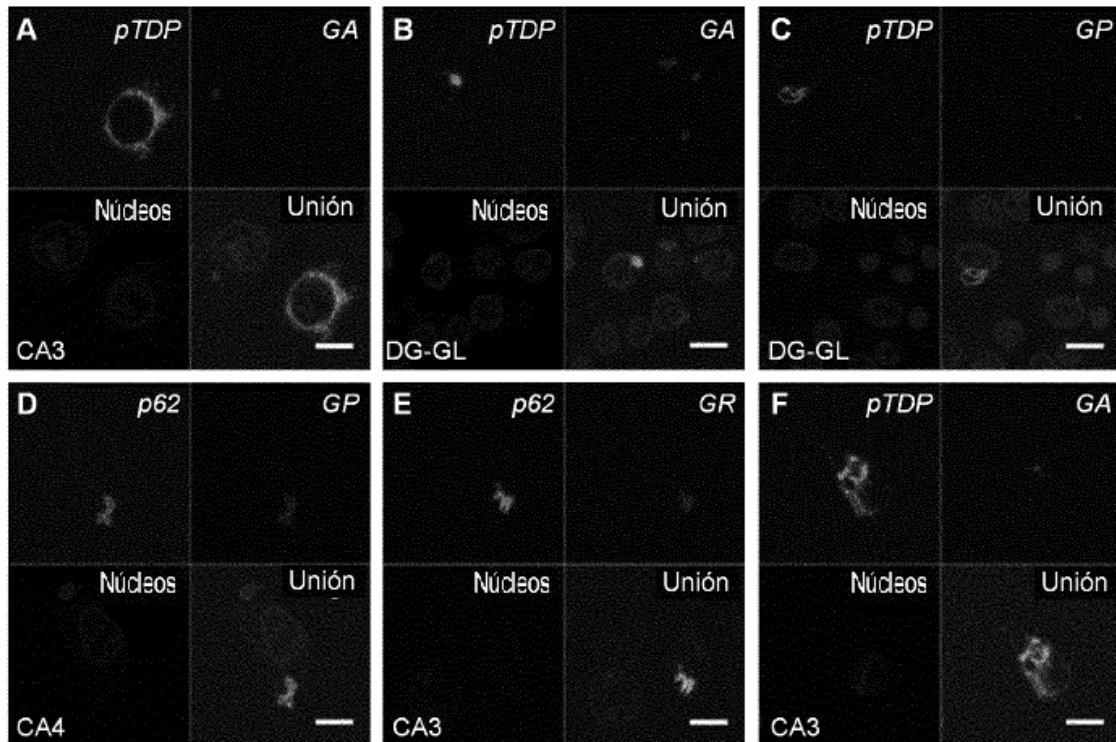
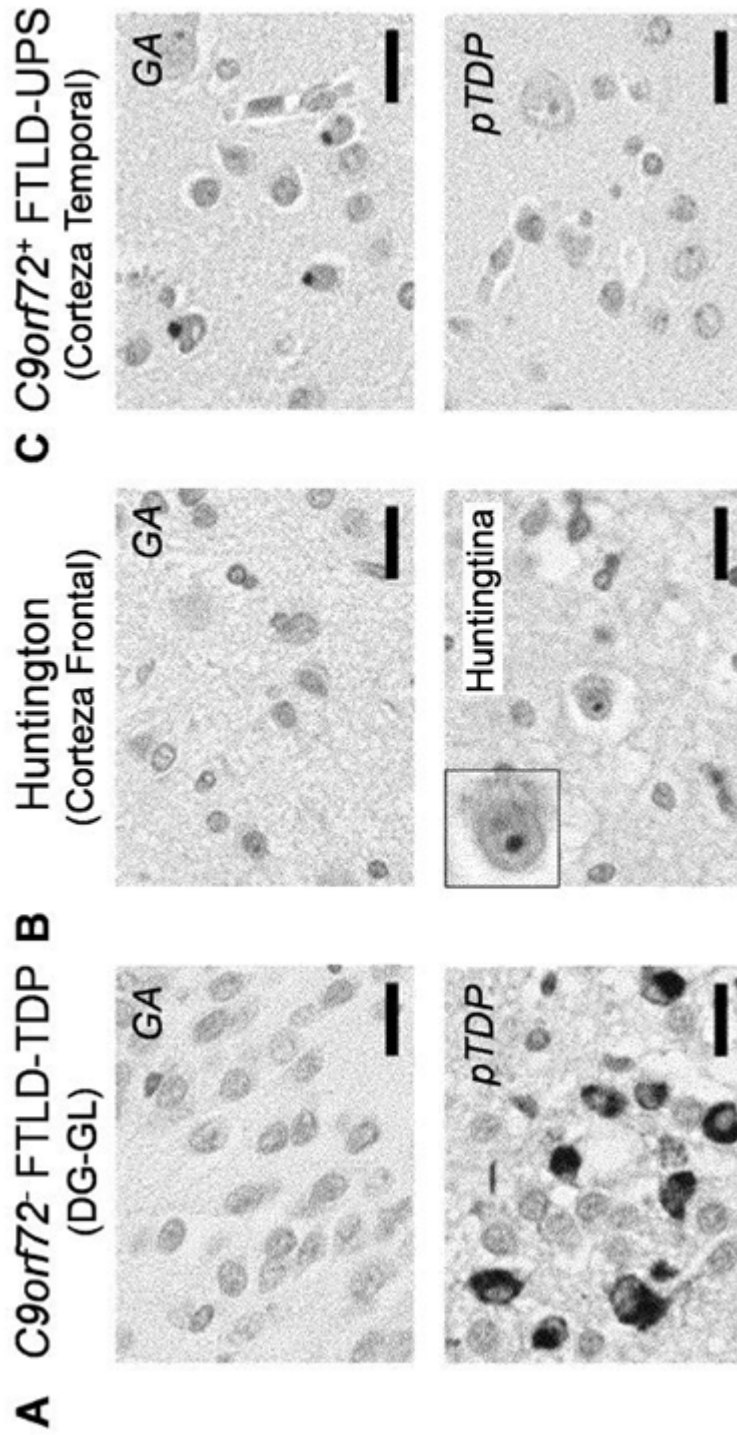


Fig. 7



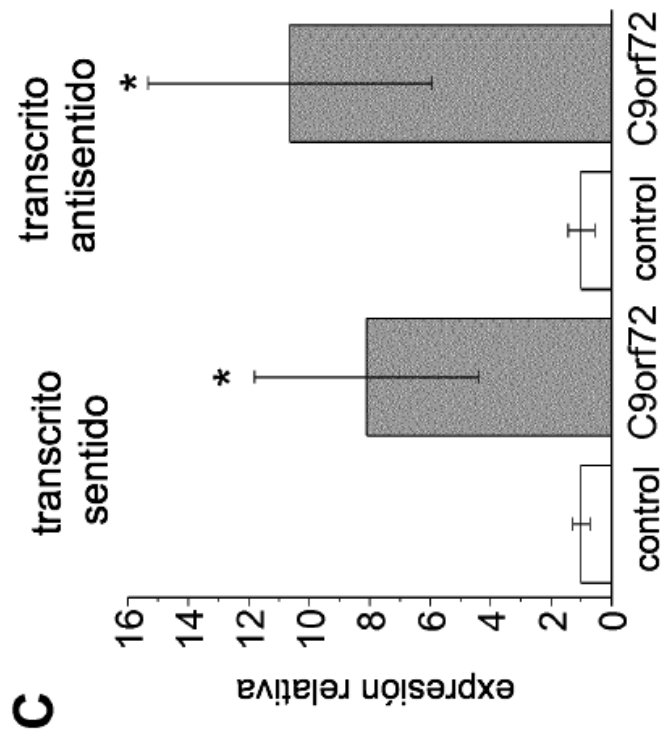
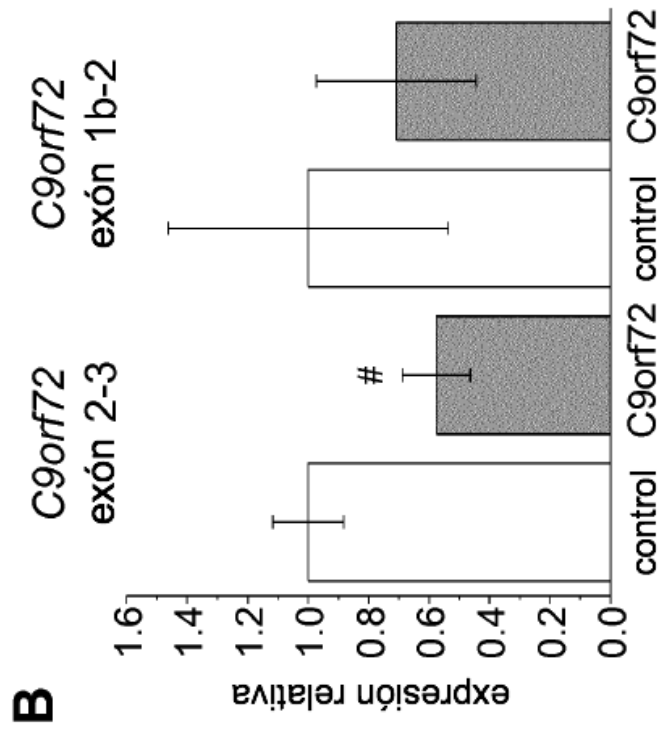
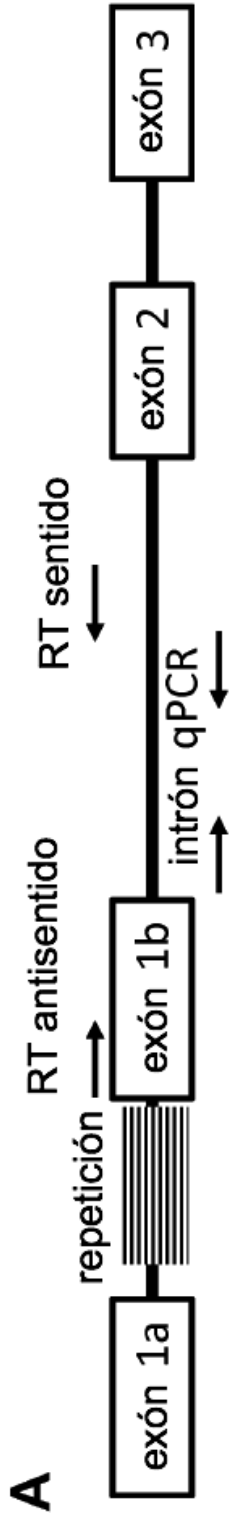


Fig. 8

Fig. 9

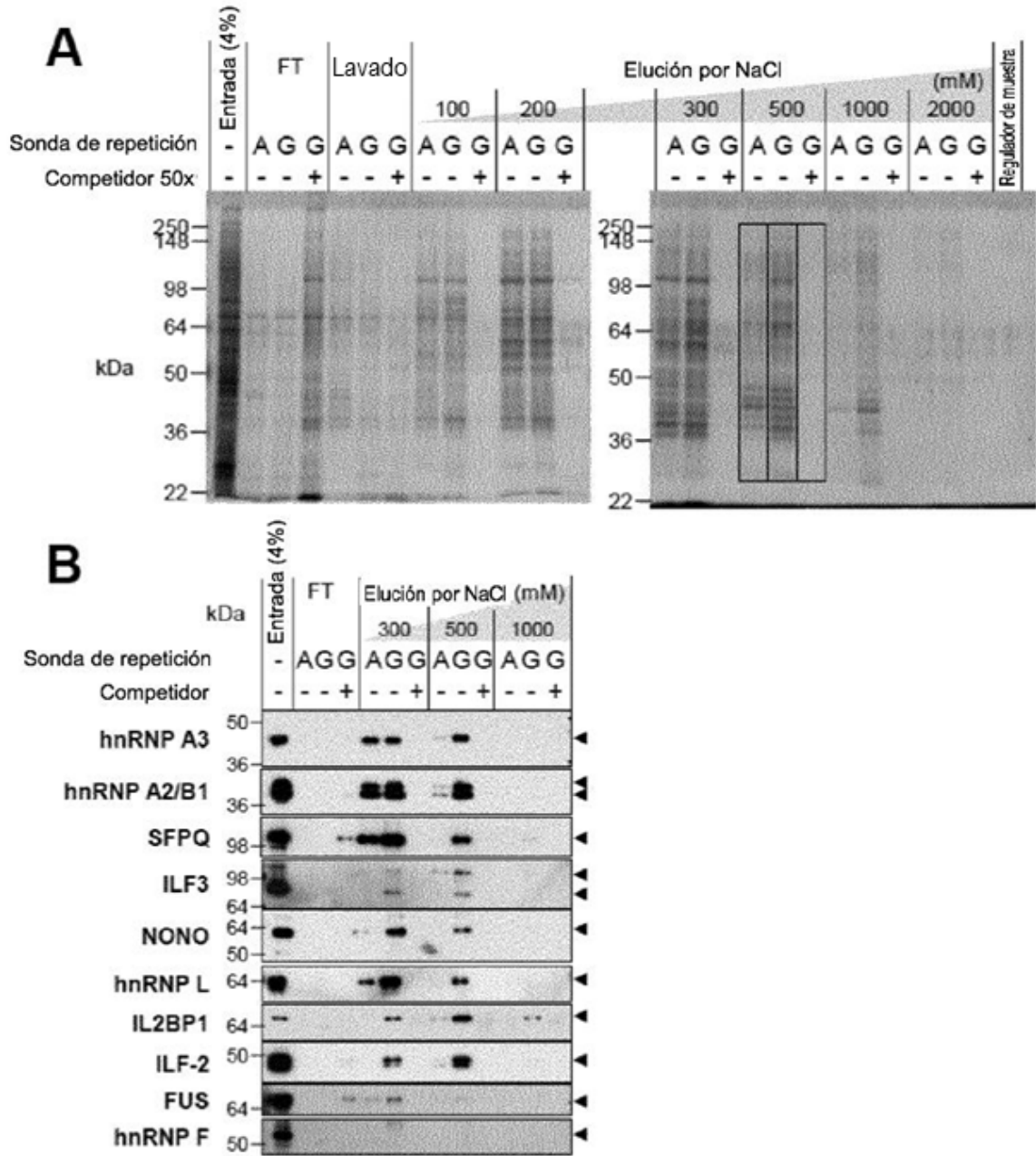


Fig. 10

	Proteínas Identificadas	Número de acceso	Valor Cuantitativo Promedio			Relación GC/AC
			AC	GC	Competición	
1	Ribonucleoproteínas nucleares heterogéneas A2/B1	ROA2_HUMANO	110,3	369,8	1,7	3,4
2	Factor de empalme, rico en prolina y glutamina	SFPQ_HUMANO	51,7	130,5	0,1	2,5
3	Factor de empalme 3B subunidad 3	SF3B3_HUMANO	51,1	118,8	6,3	2,3
4	Proteína 1 tipo ELAV; (Hu-antígeno R)	ELAV1_HUMANO	9,6	117,1	0,0	12,1
5	Factor 3 de unión al potenciador de interleuquina	ILF3_HUMANO	24,7	87,4	0,0	3,5
6	Proteína de unión al octámero que contiene el dominio no POU	NONO_HUMANO	13,4	80,4	0,0	6,0
7	Ribonucleoproteína nuclear heterogénea R	HNRPR_HUMANO	28,8	74,1	0,0	2,6
8	Ribonucleoproteína nuclear heterogénea A3	ROA3_HUMANO	23,8	70,6	0,0	3,0
9	Ribonucleoproteína nuclear heterogénea L	HNRPL_HUMANO	21,7	57,2	0,0	2,6
10	Factor B1 de unión al andamio	SAFB1_HUMANO	27,5	55,0	0,0	2,0
11	Proteína 1 de unión a ARNm del factor 2 de crecimiento tipo insulina (IMP1)	IF2B1_HUMANO	19,5	47,9	0,6	2,5
12	Factor B2 de unión al andamio	SAFB2_HUMANO	20,0	45,0	0,0	2,3
13	Ribonucleoproteína nuclear heterogénea A1	ROA1_HUMANO	20,8	44,9	0,1	2,2
14	Adenosina desaminasa específica de ARN de cadena doble (ADAR1)	DSRAD_HUMANO	16,5	42,3	0,3	2,6
15	ARN dependiente de ATP del factor de empalme de pre-ARNm putativo	DHX15_HUMANO	13,0	31,0	1,0	2,4
16	Factor 2 de unión al potenciador de interleuquina	ILF2_HUMANO	8,3	25,3	0,0	3,0
17	ARN helicasa DHX30 putativa dependiente de ATP	DHX30_HUMANO	6,3	23,5	0,0	3,8
18	Ribonucleoproteína nuclear K heterogénea	HNRPK_HUMANO	10,7	23,2	0,0	2,2
19	ARN helicasa 2 nucleolar	DDX21_HUMANO	7,5	23,0	0,7	3,1
20	Proteína FUS de unión a ARN	FUS_HUMANO	7,0	20,0	0,0	2,9

Fig. 11

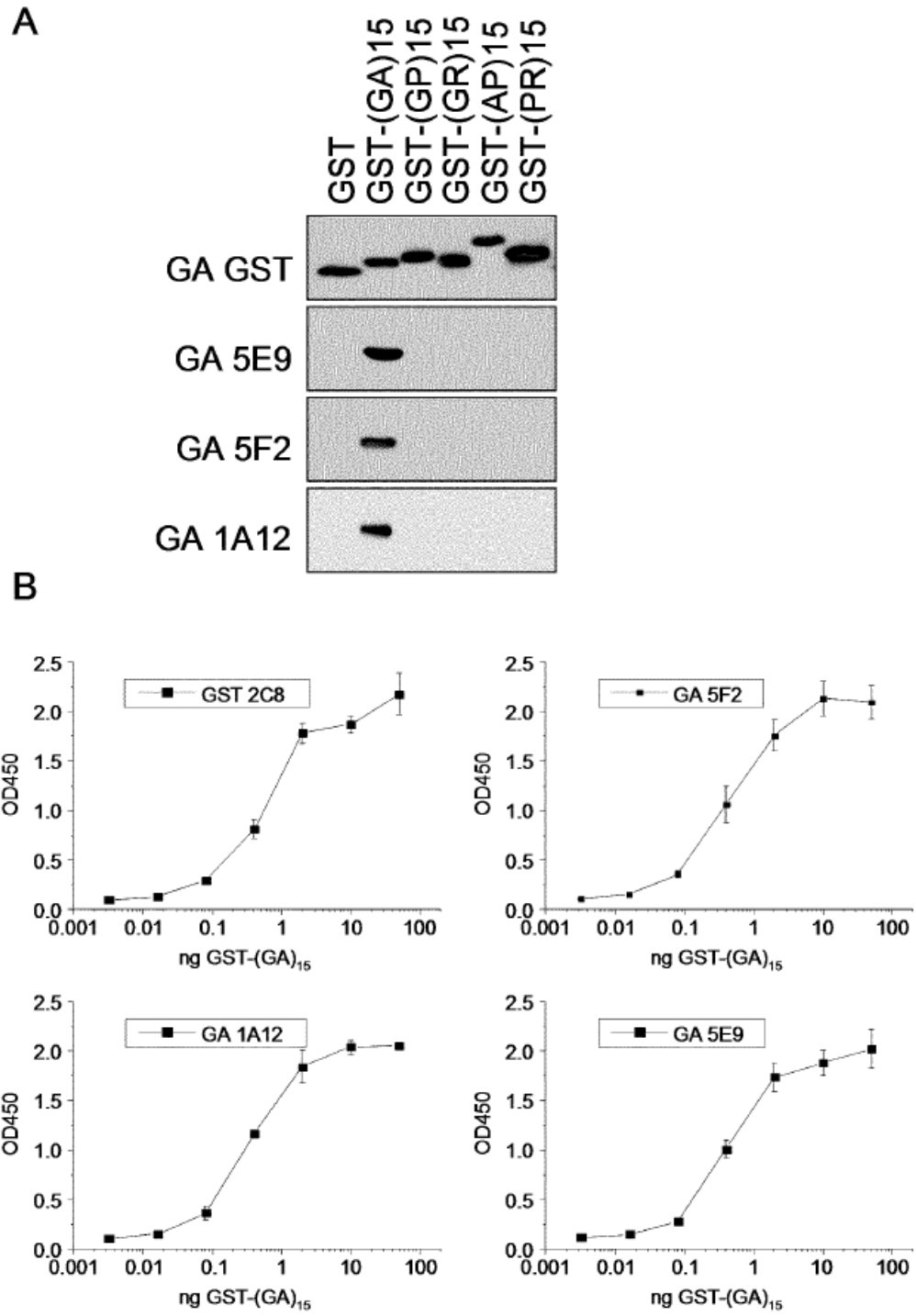


Fig. 12

Secuencias de las regiones variables de las cadenas pesada y ligera

Secuencia de la región variable de cadena ligera (LCVR):

ivisqdelsnpvtsgesvsiscRSSKSLLYKDGKTYLNwflqrpqspqlliyLMSTRASgvsdrgsgsgtdftl
eisrvkaedvgvyfcQQLVEYPLTfgagtklel

Secuencia de GA 1A12 de la región variable de cadena pesada (HCVR):

lqqsgaelmkpgasvklscrssGYTFTGYWIEwvkqrpghglewigEILPGSGSTKynenfrgkatftadtssntay
irlssltdedsaiyfcargDFTNSHFAYwgqgtlvtvs

Secuencia de GA 5E9 de la región variable de cadena pesada (HCVR):

lqqsgaevmkpgasvklscvtGYTFTGYWIEwvkqrpghglewigENLPGSGSTKynekfkgkatftadtssnt
aymqlssltdedsaiyyfcargDYSNSHFAYwgqgtlvtvs

Secuencia de GA 5F2 de la región variable de cadena pesada (HCVR):

lqqsgvelmkpgasvklscstGYKFIGYWIEwvkqrpghglewigENLPGSGTTKynenfrgkatftadtssnta
ymqlssltdedsaiyyfcargDYSNSHFTYwgqgtlvtvs