(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 883 998**

(51) Int. CI.:

C12N 9/12 (2006.01)
C12N 15/113 (2010.01)
A61K 48/00 (2006.01)
A61K 38/45 (2006.01)
A61K 38/00 (2006.01)
A61P 35/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **30.05.2017 PCT/GB2017/051539**

(87) Fecha y número de publicación internacional: **07.12.2017 WO17207979**

(96) Fecha de presentación y número de la solicitud europea: **30.05.2017 E 17730896 (2)**

(97) Fecha y número de publicación de la concesión europea: **03.03.2021 EP 3430143**

(54) Título: **Inhibidores de SRSF1 para tratar trastornos neurodegenerativos**

(30) Prioridad:

01.06.2016 GB 201609597

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
09.12.2021

(73) Titular/es:

UNIVERSITY OF SHEFFIELD (100.0%)
Firth Court, Western Bank
Sheffield, South Yorkshire S10 2TN, GB

(72) Inventor/es:

HAUTBERGUE, GUILLAUME;
AZZOUZ, MIMOUN;
WHITWORTH, ALEXANDER y
SHAW, PAMELA

(74) Agente/Representante:

FERNÁNDEZ POU, Felipe

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

Inhibidores de SRSF1 para tratar trastornos neurodegenerativos

Campo de la descripción

La presente descripción se refiere a antagonistas que se dirigen al factor de corte y empalme 1 rico en Serina/Arginina (SRSF1); vectores de expresión que comprenden antagonistas de SRSF1; y el uso de tales antagonistas en la terapia para el tratamiento de trastornos neurodegenerativos y cáncer.

Antecedentes de la descripción

La esclerosis lateral amiotrófica (ELA) y la demencia frontotemporal (DFT) son enfermedades neurodegenerativas de inicio en la edad adulta sin un tratamiento efectivo. La DFT es una causa común de demencia de inicio temprano caracterizada por una pérdida progresiva de células neuronales en los lóbulos frontal y temporal que conduce a alteraciones de la función cognitiva y características de la personalidad, lo que deja a los pacientes incapaces de cuidarse a sí mismos y provoca la muerte entre los 2-15 años de edad desde el inicio de la enfermedad. La ELA es la degeneración prematura de las neuronas motoras que conduce a la atrofia muscular y la parálisis que conduce a la muerte por insuficiencia respiratoria en un plazo de 2 a 3 años desde el inicio de los síntomas. La MND es una enfermedad degenerativa de las neuronas motoras de las células nerviosas que controlan los músculos. La MND conduce a una parálisis progresiva y a la muerte en un plazo de 2-5 años desde que aparecen los síntomas. La prevalencia de la enfermedad es de alrededor de 6-8/100 000 personas.

Las opciones de tratamiento neuroprotector son actualmente extremadamente limitadas y el agente anti-glutamatérgico riluzol prolonga la supervivencia en pacientes con ELA por solo aproximadamente 3 meses o alrededor de 6 meses de extensión de vida en pacientes con MND. La ELA y la FTD muestran una superposición clínica y patológica sustancial, el 40-50 % de los pacientes con ELA presentan disfunciones de la FTD, mientras que en aproximadamente el 5-10 % de los casos de pacientes con FTD desarrollan ELA y, por lo tanto, se propone que constituyan un espectro de la enfermedad.

Se han descubierto varios genes y se cree que son responsables de la ELA, tales como SOD1, TARDBP, FUS, OPTN y VCP. Además, se han identificado anticuerpos serina/arginina proteína quinasa (SRPK1) en suero de pacientes con enfermedad de Alzheimer (Daniilidou y otros. Journal of Neuroinmunology, 238 (2011) 67-72). Además, ambas enfermedades se caracterizan por la presencia de inclusiones de la proteína de unión a ADN (TDP-43) de respuesta de transactivación de proteínas en todo el sistema nervioso central. Sin embargo, la causa genética más comúnmente identificada de ELA y FTD implica expansiones de repetición polimórfica compuestas de cientos a miles de la secuencia de repetición de hexanucleótidos GGGGCC (de aquí en adelante abreviada G4C2) en el primer intrón del *gen C9ORF72*, con herencia autosómica dominante y penetrancia incompleta (*1*, *2*). El gen de C9ORF72 se encuentra en el brazo corto del marco de lectura abierto 72 del cromosoma 9 y se presenta en dos isoformas. Esta proteína se encuentra en muchas regiones del cerebro en el citoplasma de las neuronas, así como también en las terminales presinápticas, y también se encuentra mutada en enfermedades como la demencia lobar frontotemporal (FTLD), el trastorno tipo Huntington, la esclerosis lateral primaria, la atrofia muscular progresiva, el síndrome corticobasal, la enfermedad de Alzheimer y la demencia con cuerpos de Lewy. La fisiopatología implica potencialmente tres mecanismos ampliamente estudiados que pueden coexistir: (i) ganancia de función tóxica de ARN por secuestro de factores de unión a ARN (5-9), (ii) ganancia de función tóxica de proteína debido a la traducción de no ATG asociada a repeticiones (RAN) que ocurre en todos los marcos de lectura sentido y antisentido para producir cinco proteínas de repetición dipéptido (DPR (10-14) y (iii) haploinsuficiencia debido a la disminución de la expresión de la proteína C9ORF72 (1, 15, 16).

Típicamente, los enfoques terapéuticos se han dirigido a C9ORF72. El documento WO2016/024205 describe oligómeros complementarios del gen de C9ORF72 y, en particular, composiciones para dirigirse al ARN que contienen un número patológico de repeticiones de hexanucleótidos para su uso en el tratamiento de un trastorno neurológico seleccionado del grupo de ELA y FTD. El documento WO2014/062691 describe composiciones y métodos para reducir la expresión de ARNm de C9ORF72 y proteína en un animal para el tratamiento de ELA, FTD, síndrome de degeneración corticobasal, síndrome de Parkinson atípico y degeneración olivopontocerebelosa mediante el suministro de ARN antisentido dirigido contra los ácidos nucleicos de C9ORF27. Los ácidos nucleicos se administran en el sistema nervioso central por vía intratecal o intraventricular. Solo se han aprobado dos fármacos antisentido; fomivirsen para el tratamiento de la retinitis por citomegalovirus y mipomersen para el tratamiento de la hipercolesterolemia familiar homocigótica. Los vectores virales presentan una forma alternativa de vehículo de suministro del material genético. Los sistemas de vectores virales, basados en virus adenoasociados y lentivirus, son ideales para mediar la iARN porque pueden transducir de forma segura un amplio intervalo de tejidos y proporcionar niveles sostenidos de expresión génica. Además, se conocen pequeñas moléculas inhibitorias con función neuroprotectora como, por ejemplo, moléculas antisentido contra sefina1 (Crunkhorn, Nature Reviews. Drug Discovery, Vol 14, No 6, mayo de 2015), o contra SOD1 (WO2016/077687) se sabe que ambas proteínas están implicadas en la ELA.

La exportación nuclear de ARNm está mediada por la proteína NXF1 (factor de exportación nuclear 1) y se cree que los adaptadores de exportación como SRSF1 (factor de corte y empalme rico en serina/arginina 1) y ALYREF (factor de exportación Aly/REF) aumentan la afinidad por los ARNm maduros, lo que previene la exportación de transcritos no procesados. Se sabe que SRSF1 participa en la regulación de una variedad de vías y se encontró que se sobreexpresa en una variedad de tipos de cáncer y se cree que es una diana para la terapia contra el cáncer (Das y Krainer, Molecular Cancer Research, vol. 12, No. 9, mayo de 2014). Es más, Wee y otros. PLOS ONE, Vol. 9, No. 12, diciembre de 2014 describen moléculas de ARNip contra SRSF1 con función inhibitoria o de activación para su uso en el tratamiento de la atrofia muscular espinal. La inmunohistoquímica en tejido del sistema nervioso central de pacientes C9orf72+ con ELA demostró la colocalización de ARN repetido GGGGCC con SRSF2, hnRNP H1/F, ALYREF y hnRNP A1 en células granulares cerebelosas y con SRSF2, hnRNP H1/F y ALYREF en neuronas motoras, la diana principal de la patología en la ELA. También se confirmó la unión directa de proteínas al ARN repetido GGGGCC.

La presente descripción ha identificado que la unión excesiva de los adaptadores de exportación nuclear a los transcritos repetidos de G4C2 fuerza las interacciones con NXF1, lo que anula los mecanismos normales de retención nuclear y enseña que el agotamiento de factores como SRSF (también conocido como SF2 o ASF) 1 confiere neuroprotección.

Declaraciones de la invención

De acuerdo con un aspecto de la invención, se proporciona un casete de transcripción para su uso en el tratamiento de una enfermedad neurodegenerativa causada por la expansión polimórfica de G4C2 en el primer intrón del gen de C9ORF72 en donde dicho vector de expresión comprende:

i) un ácido nucleico antisentido, un ARN o ARNhc inhibidor que es complementario a e inhibe la expresión de un ácido nucleico que codifica un factor de corte y empalme rico en serina/arginina [SRSF1];
ii) una molécula de ácido nucleico que codifica un péptido que comprende o consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 15; o
iii) una molécula de ácido nucleico que codifica una proteína negativa dominante que comprende o consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 16 o 17;

en donde dicha molécula de ácido nucleico está unida operativamente a un promotor adaptado para expresar el agente de acuerdo con i), ii) o iii).

En una modalidad preferida de la invención, dicho ARN inhibidor, ácido nucleico antisentido o ARNhc, comprende o consiste en una secuencia de nucleótidos como se establece en la SEQ ID NO: 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14 o 47, 48, 50, 99, 154, 155, 156, 157, 158, 60, 97, 98, 100, 101, 102, 103, 104, 105, 106, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139 o 141.

En una modalidad preferida de la invención, dicho vector de expresión es un vector de expresión de base viral.

En una modalidad preferida de la invención, dicho vector de base viral es un virus adenoasociado [AAV].

En una modalidad preferida de la invención, dicho vector de base viral es AAV9.

En una modalidad preferida de la invención, dicho vector de base viral es un lentivirus.

En una modalidad preferida de la invención, dicha enfermedad neurodegenerativa se selecciona del grupo que consiste en: esclerosis lateral amiotrófica, demencia frontotemporal, enfermedad de la neurona motora, demencia lobar frontotemporal, trastorno tipo Huntington, esclerosis lateral primaria, atrofia muscular progresiva, síndrome corticobasal, enfermedad de Alzheimer y demencia con cuerpos de Lewy.

En una modalidad preferida de la invención, dicha enfermedad neurodegenerativa es una enfermedad de la neurona motora.

En una modalidad preferida de la invención, dicha enfermedad neurodegenerativa es la esclerosis lateral amiotrófica.

En una modalidad preferida de la invención, dicha enfermedad neurodegenerativa es la demencia frontotemporal.

Describimos un agente antagonista que inhibe la expresión de una molécula de ácido nucleico que codifica el factor de corte y empalme 1 rico en serina/arginina [SRSF1] o inhibe la actividad de una proteína SRSF1.

La inhibición de la expresión define la reducción de la expresión que varía del 1-100 % en comparación con la expresión del ácido nucleico/proteína que se encuentra en el tipo salvaje.

También describimos un casete de transcripción que comprende: una molécula de ácido nucleico que codifica un agente antagonista en donde dicho agente inhibe la expresión de un ácido nucleico que codifica un factor de corte y empalme rico en serina/arginina [SRSF1] o inhibe la actividad de una proteína SRSF1 en donde dicha molécula de ácido nucleico está unida operativamente a un promotor adaptado para expresar dicho agente.

En algunos ejemplos, SRSF1 está codificado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se establece en la SEQ ID NO: 1, o la variante de secuencia polimórfica que tiene 90 %-99 % de identidad de secuencia sobre la secuencia de nucleótidos de longitud completa como se establece en la SEQ ID NO:1.

En algunos ejemplos, SRSF1 está codificado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se establece en la SEQ ID NO: 3 o la variante de secuencia polimórfica que tiene una identidad de secuencia del 90 %-99 % sobre la secuencia de nucleótidos de longitud completa como se establece en la SEQ ID NO: 3.

En algunos ejemplos, SRSF1 está representado por la secuencia de aminoácidos como se establece en la SEQ ID NO:2, o una variante de secuencia polimórfica que tiene una identidad de secuencia del 90-99 % sobre la secuencia de aminoácidos de longitud completa como se indica en la SEQ ID NO: 2.

En algunos ejemplos, SRSF1 está representado por la secuencia de aminoácidos que se establece en la SEQ ID NO: 4, o una variante de secuencia polimórfica que tiene una identidad de secuencia del 90-99 % sobre la secuencia de aminoácidos de longitud completa como se establece en la SEQ ID NO: 4.

En algunos ejemplos, dicha molécula de ácido nucleico codifica un agente basado en ácido nucleico.

En algunos ejemplos, dicha molécula de ácido nucleico codifica un ácido nucleico antisentido.

En algunos ejemplos, dicha molécula basada en ácido nucleico es un ARN inhibidor.

En algunos ejemplos dicho ARN inhibidor es una molécula de ARNip o ARNhc o miARN.

Una técnica para eliminar específicamente la función génica es mediante la introducción de ARN bicatenario, también denominado ARN inhibidor pequeño o de interferencia (ARNip), en una célula que da como resultado la destrucción del ARNm complementario a la secuencia incluida en la molécula de ARNip. La molécula de ARNip comprende dos hebras complementarias de ARN (una hebra sentido y una hebra antisentido) hibridadas entre sí para formar una molécula de ARN bicatenario. La molécula de ARNip se deriva típicamente de exones del gen que se van a eliminar. Se está dilucidando el mecanismo de interferencia del ARN. Muchos organismos responden a la presencia de ARN bicatenario mediante la activación de una cascada que conduce a la formación de ARNip. La presencia de ARN bicatenario activa un complejo proteico que comprende ARNasa III que procesa el ARN bicatenario en fragmentos más pequeños (ARNip, aproximadamente 21-29 nucleótidos de longitud) que pasan a formar parte de un complejo ribonucleoproteico. El ARNip actúa como una guía para que el complejo de ARNasa escinda el ARNm complementario a la hebra antisentido del ARNip, lo que resulta en la destrucción del ARNm.

En algunos ejemplos, dicha molécula de ARN inhibitoria tiene una longitud entre 19 nucleótidos [nt] y 29 nt. Con mayor preferencia aún, dicha molécula de ARN inhibitoria tiene entre 21 nt y 27 nt de longitud. Preferentemente, dicha molécula de ARN inhibitoria tiene aproximadamente 21 nt de longitud.

En algunos ejemplos, dicho ARN inhibidor comprende o consiste en una secuencia de nucleótidos como se establece en las SEQ ID NO: 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14.

En algunos ejemplos, dicho ARN inhibidor comprende o consiste en una secuencia de nucleótidos como se establece en las SEQ ID NO: 97, 98, 99, 100, 101, 102, 103, 104, 105 o 106.

En algunos ejemplos, dicho ARN inhibidor comprende o consiste en una secuencia de nucleótidos como se establece en las SEQ ID NO: 107, 109, 111, 113, 115, 117, 119, 121, 123 o 125.

En algunos ejemplos, dicho ARN inhibidor comprende o consiste en una secuencia de nucleótidos como se establece en las SEQ ID NO: 127, 129, 131, 133, 135, 137, 139 o 141.

En algunos ejemplos, dicho ARN inhibidor comprende o consiste en una secuencia de nucleótidos como se establece en las SEQ ID NO: 47, 48, 50, 99, 154, 155, 156, 157, 158 o 60.

En algunos ejemplos, dicho agente basado en ácido nucleico comprende nucleótidos modificados.

En algunos ejemplos, dicho agente es un péptido.

En una modalidad preferida de la invención, dicho péptido comprende una secuencia de aminoácidos que tiene al menos 32 aminoácidos de longitud y comprende la secuencia de aminoácidos establecida en la SEQ ID NO: 15.

En una modalidad preferida de la invención, dicho péptido tiene al menos 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o al menos 100 aminoácidos de longitud pero menos que la secuencia de aminoácidos de longitud completa como se establece en las SEQ ID NO: 2 o 4.

En una modalidad preferida de la invención, dicho péptido consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 15.

En una modalidad preferida de la invención, dicho péptido se modifica, por ejemplo, dicho péptido se cicla.

En algunos ejemplos, dicha proteína es una proteína negativa dominante que comprende una modificación de la secuencia de aminoácidos como se establece en las SEQ ID NO: 2 o 4.

En algunos ejemplos, dicha proteína negativa dominante comprende o consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 2 o 4 en donde dicha secuencia de aminoácidos se modifica mediante la adición, deleción o sustitución de uno o más residuos de aminoácidos.

En una modalidad preferida de la invención, dicha proteína modificada comprende o consiste en la secuencia de aminoácidos como se establece en la SEQ ID NO: 16 o 17.

En algunos ejemplos, dicho agente antagonista es una proteína quinasa específica para SRSF1 para mantener el estado de fosforilación de SRSF1.

En algunos ejemplos, dicha proteína quinasa es SRPK1.

En algunos ejemplos, SRPK1 está codificado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se establece en la SEQ ID NO: 18.

En algunos ejemplos, SRPK1 comprende una secuencia de aminoácidos como se establece en la SEQ ID NO: 19.

En algunos ejemplos, dicha proteína quinasa es CDC tipo quinasa 1.

En algunos ejemplos, la CDC tipo quinasa 1 está codificada por una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se establece en la SEQ ID NO: 20.

En algunos ejemplos, la CDC tipo quinasa 1 comprende una secuencia de aminoácidos como se establece en la SEQ ID NO: 21.

En algunos ejemplos, dicho agente antagonista es un inhibidor de fosfatasa específico para la proteína fosfatasa-1 humana.

En algunos ejemplos, dicho antagonista inhibe la expresión o la actividad de la proteína fosfatasa-1 humana.

En algunos ejemplos, la proteína fosfatasa-1 humana está codificada por una secuencia de nucleótidos como se establece en la SEQ ID NO: 22.

En algunos ejemplos dicho antagonista es un ARN inhibidor.

En algunos ejemplos dicho promotor es un promotor constitutivo.

En algunos ejemplos, dicho promotor es un promotor regulado, por ejemplo, un promotor inducible o específico celular.

Describimos un vector de expresión que comprende un casete de transcripción de acuerdo con la invención.

Un número de virus se usan comúnmente como vectores para el suministro de genes exógenos. Los vectores comúnmente empleados incluyen virus de ADN y ARN con o sin envoltura modificados recombinantemente, por ejemplo, baculoviridiae, parvoviridiae, picornoviridiae, herpesveridiae, poxviridae, adenoviridiae, picornnaviridiae o retroviridae. También se pueden emplear vectores quiméricos que explotan elementos ventajosos de cada una de las propiedades del vector parental (ver, por ejemplo, Feng, y otros. (1997) Nature Biotechnology 15: 866-870). Dichos vectores virales pueden ser de tipo salvaje o pueden modificarse mediante técnicas de ADN recombinante para que sean deficientes en la replicación, de replicación condicional o competentes en la replicación. Los vectores virales de replicación condicional se usan para lograr la expresión selectiva en tipos celulares particulares lo que evita al mismo tiempo una infección de amplio espectro inapropiada. Ejemplos de vectores de replicación condicional se describen en Pennisi, E. (1996) Science 274: 342-343.; Russell y SJ (1994) Eur. J. de Cáncer 30A (8): 1165-1171.

Los vectores preferidos se derivan de los genomas adenovirales, virales adenoasociados y retrovirales.

En una modalidad preferida de la invención, dicho vector de expresión es un vector de expresión de base viral.

En una modalidad preferida de la invención, dicho vector de base viral es un virus adenoasociado [AAV].

En una modalidad preferida de la invención, dicho vector de base viral es AAV9.

En una modalidad alternativa preferida de la invención, dicho vector de base viral es un vector lentiviral.

Describimos una composición farmacéutica que comprende un vector de expresión de acuerdo con la invención y un excipiente o portador.

Las composiciones de agente o de vector de expresión de la presente invención se administran en preparaciones farmacéuticamente aceptables. Tales preparaciones pueden contener habitualmente concentraciones farmacéuticamente aceptables de sal, agentes tampones, conservantes, portadores compatibles, y agentes terapéuticos suplementarios. Las composiciones del agente o del vector de expresión de la invención pueden administrarse por cualquier vía convencional, incluida la inyección o por infusión gradual a lo largo del tiempo.

Las composiciones del agente o del vector de expresión de la invención se administran en cantidades eficaces. Una "cantidad efectiva" es aquella cantidad del agente o del vector de expresión que, solo o junto con dosis adicionales, produce la respuesta deseada. En el caso de tratar una enfermedad, la respuesta deseada es inhibir la progresión de la enfermedad. Esto puede involucrar disminuir solamente el progreso de la enfermedad temporalmente, aunque más preferentemente, involucra suspender el progreso de la enfermedad permanentemente. Esto se puede monitorear mediante métodos de rutina. Tales cantidades dependerán, por supuesto, de la condición particular que se trata, la gravedad de la condición, los parámetros individuales del paciente, que incluyen la edad, la condición física, el tamaño y el peso, la duración del tratamiento, la naturaleza de la terapia concurrente (si corresponde), la vía específica de administración y factores similares dentro del conocimiento y la experiencia del practicante de salud. Los factores son bien conocidos para las personas expertas en la técnica y pueden resolverse con no más que experimentación rutinaria. Se prefiere generalmente que se use una dosis máxima de los componentes individuales o sus combinaciones, o sea, la dosis segura más alta de acuerdo con el buen juicio médico. Sin embargo, los expertos en la técnica entenderán que un paciente puede insistir en una dosis más baja o una dosis tolerable por razones médicas, psicológicas o virtualmente por cualquier otra razón.

Las composiciones del agente o del vector de expresión usadas en los métodos precedentes son preferentemente estériles y contienen una cantidad efectiva del agente o del vector de expresión de acuerdo con la invención para producir la respuesta deseada en una unidad de peso o volumen adecuada para la administración a un paciente. Las dosis del agente o del vector administrada a un sujeto se pueden seleccionar de acuerdo con diferentes parámetros, en particular de acuerdo con el modo de administración usado y el estado del sujeto. Otros factores incluyen el período de tratamiento deseado. En el caso de que la respuesta en un sujeto sea insuficiente a las dosis iniciales aplicadas, pueden emplearse dosis más altas (o dosis más altas efectivas por una vía de suministro diferente, más localizada), hasta el límite que permita la tolerancia del paciente. Los expertos en la técnica conocerán otros protocolos para la administración de las composiciones del agente o del vector, en los que la cantidad de dosis, el programa de inyecciones, los lugares de las inyecciones, el modo de administración y los similares varían de los precedentes. La administración de las composiciones a mamíferos distintos de los seres humanos (por ejemplo, con fines de prueba o con fines terapéuticos veterinarios), se lleva a cabo sustancialmente en las mismas condiciones que las descritas anteriormente. Un sujeto, como se usa en la presente, es un mamífero, preferentemente un ser humano, e incluye un primate no humano, vaca, caballo, cerdo, oveja, cabra, perro, gato o roedor.

Cuando se administran, las composiciones del agente o del vector de expresión de la invención se aplican en cantidades farmacéuticamente aceptables y en composiciones farmacéuticamente aceptables. El término "farmacéuticamente aceptable" significa un material no tóxico que no interfiere con la eficacia de la actividad biológica del agente activo. Tales preparaciones pueden contener rutinariamente sales, agentes tamponadores, conservantes, vehículos compatibles y, opcionalmente, otros agentes terapéuticos (por ejemplo, los usados típicamente en el tratamiento de la indicación de la enfermedad específica). Cuando se usan en la medicina, las sales deben ser farmacéuticamente aceptables, pero las sales que no son farmacéuticamente aceptables se pueden usar convenientemente para preparar sales farmacéuticamente aceptables de estas y no se excluyen del alcance de la invención. Tales sales farmacológica y farmacéuticamente aceptables incluyen, pero no se limitan a, aquellas preparadas de los siguientes ácidos: clorhídrico, bromhídrico, sulfúrico, nítrico, fosfórico, maleico, acético, salicílico, cítrico, fórmico, malónico, succínico, y similares. Además, las sales farmacéuticamente aceptables se pueden preparar como sales de metales alcalinos o alcalinos térreos, tales como sales de sodio, potasio o calcio.

Las composiciones farmacéuticas que contienen los agentes o vectores de expresión de acuerdo con la invención pueden contener agentes tamponadores adecuados, que incluyen: ácido acético en una sal; ácido cítrico en una sal; ácido bórico en una sal; y ácido fosfórico en una sal. Las composiciones farmacéuticas pueden contener también, opcionalmente, conservantes adecuados, tales como: cloruro de benzalconio; clorobutanol; parabenos y timerosal.

Las composiciones del agente o del vector de expresión se pueden presentar convenientemente en forma de dosis unitarias y se puede preparar por cualquiera de los métodos bien conocidos en la técnica de farmacia. Todos los métodos incluyen la etapa de reunir el agente activo en asociación con un agente o vector que constituye uno o más de los ingredientes auxiliares. Esta preparación se puede formular de acuerdo con métodos conocidos mediante el uso de agentes humectantes o dispersantes adecuados y agentes de suspensión. La preparación inyectable estéril puede ser también una suspensión o solución inyectable estéril en un diluente o solvente parenteralmente aceptable no tóxico, por ejemplo, como una solución en 1,3-butanodiol. Entre los solventes aceptables que se pueden emplear están el agua, solución de Ringer, y solución de cloruro de sodio isotónica. En adición, los aceites fijos, estériles se emplean convencionalmente como solventes o medios de suspensión. Para este propósito cualquier aceite fijo suave se puede emplear e incluyen mono-o di-glicéridos sintéticos. Adicionalmente, los ácidos grasos tal como el ácido oleico se pueden usar en la preparación de inyectables. Una formulación portadora adecuada para administraciones oral, subcutánea, intravenosa, intramuscular, etc. se pueden encontrar en Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, PA.

Describimos un agente antagonista de acuerdo con la invención para su uso como medicamento.

Describimos un vector de expresión de acuerdo con la invención para su uso como medicamento.

Describimos un agente antagonista de acuerdo con la invención para su uso en el tratamiento de una enfermedad neurodegenerativa.

Describimos un vector de expresión de acuerdo con la invención para su uso en el tratamiento de una enfermedad neurodegenerativa.

En algunos ejemplos dicha enfermedad neurodegenerativa se selecciona del grupo que consiste en: enfermedad de la neurona motora, demencia lobar frontotemporal (FTLD), trastorno tipo Huntington, esclerosis lateral primaria, atrofia muscular progresiva, síndrome corticobasal, enfermedad de Alzheimer y demencia con cuerpos de Lewy.

En algunos ejemplos, dicha enfermedad neurodegenerativa es una enfermedad de la neurona motora.

Describimos un método para el tratamiento de una enfermedad neurodegenerativa que comprende administrar una cantidad terapéuticamente efectiva de un agente antagonista de acuerdo con la invención para prevenir y/o tratar dicha enfermedad neurodegenerativa.

Describimos un método para el tratamiento de una enfermedad neurodegenerativa que comprende administrar una cantidad terapéuticamente efectiva de un vector de expresión de acuerdo con la invención para prevenir y/o tratar dicha enfermedad neurodegenerativa.

Describimos un agente antagonista de acuerdo con la invención para su uso en el tratamiento del cáncer, en particular el cáncer metastásico.

Describimos un vector de expresión de acuerdo con la invención para su uso en el tratamiento del cáncer, en particular el cáncer metastásico.

El cáncer incluye referencia a los tumores. Por ejemplo, los adenocarcinomas incluyen cánceres malignos tales como la mayoría de los cánceres de colon, carcinoma de células renales, cáncer de próstata, tumores testiculares, carcinoma de pulmón de células no pequeñas, cáncer del intestino delgado y cáncer de esófago. El término "carcinoma" es reconocido en la técnica y se refiere a las malignidades de los tejidos epiteliales o endocrinos que incluyen los carcinomas del sistema respiratorio, carcinomas del sistema gastrointestinal, carcinomas del sistema genitourinario, carcinomas testiculares, carcinomas de mama, carcinomas prostáticos, carcinomas del sistema endocrino y melanomas. Los carcinomas ejemplares incluyen los que se forman a partir de tejido del cuello uterino, pulmón, próstata, mama, cabeza y cuello, colon y ovario. El término "carcinoma" también incluye carcinosarcomas, por ejemplo, que incluyen tumores malignos compuestos de tejidos carcinomatosos y sarcomatosos. Un "adenocarcinoma" se refiere a un carcinoma derivado de un tejido glandular o en el que las células tumorales forman estructuras glandulares reconocibles.

En general, las dosis de agentes antagonistas, por ejemplo, oligonucleótido antisentido, ARNip o ARNhc estarán entre 1 nM - 1 μM generalmente se formularán y administrarán de acuerdo con los procedimientos estándar. Preferentemente las dosis pueden variar entre 1 nM-500 nM, 5 nM-200 nM y 10 nM-100 nM. Además, se formulan y administran dosis de péptidos y/o antagonistas de proteínas en dosis entre 1 ng y 1 mg, y preferentemente entre 10 ng y 100 μg.

Describimos el uso de una molécula de ácido nucleico que codifica para SRSF1, o una proteína con la actividad asociada con SRSF1 en la identificación de un agente que inhibe la expresión o la actividad de SRSF1.

Describimos un método de cribado para la identificación de un agente que inhibe la expresión de SRSF1 o la actividad de una proteína con actividad asociada con SRSF1 que comprende las etapas:

i) proporcionar una célula que exprese SRSF1;

ii) poner en contacto la célula con uno o más agentes para ser probados para la actividad inhibitoria con respecto a un ácido nucleico que codifica para SRSF1 o una proteína SRSF1;

iii) monitorear el efecto de dicho (s) agente (s) sobre la expresión o la actividad de SRSF1 en comparación con una célula que no ha sido contactada con dicho (s) agente (s).

En algunos métodos, dicha célula es una célula nerviosa, por ejemplo, un astrocito.

En algún método, dicha célula se modifica para expresar SRSF1 de forma recombinante.

En algunos métodos, dicho agente es un agente basado en ácido nucleico.

En algunos métodos, dicho agente basado en ácido nucleico comprende nucleótidos modificados.

El término "modificado" como se usa en esta descripción describe una molécula de ácido nucleico en la que:

i) al menos dos de sus nucleótidos están unidos covalentemente mediante un enlace internucleotídico sintético (es decir, un enlace distinto del enlace fosfodiéster entre el extremo 5' de un nucleótido y el extremo 3' de otro nucleótido). Alternativamente o preferentemente, dicho enlace puede ser el extremo 5' de un nucleótido unido al extremo 5' de otro nucleótido o el extremo 3' de un nucleótido con el extremo 3' de otro nucleótido; y/o

ii) un grupo químico, como el colesterol, que normalmente no está asociado con los ácidos nucleicos, que se ha unido covalentemente al ácido nucleico monocatenario.

Los enlaces internucleotídicos sintéticos preferidos son fosforotioatos, alquilfosfonatos, fosforoditioatos, ésteres de fosfato, alquilfosfonotioatos, fosforamidatos, carbamatos, triésteres de fosfato, acetamidatos, péptidos y ésteres de carboximetilo.

El término "modificado" también abarca los nucleótidos con una base y/o azúcar modificados covalentemente. Por ejemplo, los nucleótidos modificados incluyen nucleótidos que tienen azúcares unidos covalentemente a grupos orgánicos de bajo peso molecular diferentes de un grupo hidroxilo en la posición 3' y de un grupo fosfato en la posición 5'. Por tanto, los nucleótidos modificados también pueden incluir azúcares sustituidos en 2' tales como 2'-O-metil-; 2-O-alquilo; 2-O-alilo; 2'-S-alquilo; 2'-S-alilo; 2'- fluoro-; 2'-halo o 2; azido-ribosa, análogos de azúcares carbocíclicos, azúcares α-anoméricos; azúcares epiméricos tales como arabinosa, xilosas o lixosas, azúcares piranosa, azúcares furanosa y sedoheptulosa.

Los nucleótidos modificados son conocidos en la técnica e incluyen purinas y/o pirimidinas alquiladas; purinas y/o pirimidinas aciladas; u otros heterociclos. Estas clases de pirimidinas y purinas son conocidas en la técnica e incluyen pseudoisocitosina; N4, N4-etanocitosina; 8-hidroxi-N6-metiladenina; 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo; 5-fluorouracilo; 5-bromouracilo, 5-carboximetilaminometil-2-tiouracilo; 5 carboximetilaminometiluracilo; dihidrouracilo; inosina; N6-isopentil-adenina; l-metiladenina; 1-metilpseudouracilo; 1-metilguanina; 2,2-dimetilguanina; 2-metiladenina; 2-metilguanina; 3-metilcitosina; 5-metilcitosina; N6-metiladenina; 7-metilguanina; 5-metilaminometiluracilo; 5-metoxi amino metil-2-tiouracilo; beta-D-manosilqueosina; 5-metoxicarbonilmetiluracilo; 5-metoxiuracilo; 2 metiltio-N6-isopenteniladenina; éster metílico del ácido uracil-5-oxiacético; pseudouracilo; 2-tiocitosina; 5-metil-2 tiouracilo, 2-tiouracilo; 4-tiouracilo; 5-metiluracilo; éster metílico del ácido N-uracil-5-oxiacético; ácido uracil 5-oxiacético; queosina; 2-tiocitosina; 5-propiluracilo; 5-propilcitosina; 5-etiluracilo; 5-etilcitosina; 5-butiluracilo; 5-pentiluracilo; 5-pentilcitosina; y 2,6,-diaminopurina; metilpseudouracilo; 1-metilguanina; 1-metilcitosina. Los ácidos nucleicos bicatenarios modificados también pueden incluir análogos de bases tales como bases modificadas con propino C-5 (ver Wagner y otros. Nature Biotechnology 14: 840-844, 1996.). El uso de nucleótidos modificados confiere, entre otras propiedades, resistencia a la digestión por nucleasas y estabilidad mejorada.

En algunos métodos, dicho agente es un agente peptídico o proteico.

En algunos métodos, dicho agente peptídico comprende uno o más aminoácidos modificados o son ciclados.

La ciclación es conocida en la técnica (ver Scott y otros. Chem Biol (2001), 8: 801-815.; Gellerman y otros. J. Peptide Res (2001), 57: 277-291; Dutta y otros. J. Peptide Res (2000), 8: 398-412; Ngoka y Gross J Amer Soc Mass Spec (1999), 10:360-363.

En algunos métodos, dicho agente es un agente orgánico pequeño.

A través de toda la descripción y las reivindicaciones de esta descripción, las palabras "comprenden" y "contienen" y las variaciones de las palabras, por ejemplo "que comprende" y "comprende", significan "que incluyen pero no se

limitan a", y no se pretende excluir (y no se excluyen) otras porciones, aditivos, componentes, enteros o etapas. "Consistir esencialmente" significa tener los enteros esenciales pero que incluyen los enteros que no afectan materialmente la función de los enteros esenciales.

A través de toda la descripción y las reivindicaciones de esta descripción, el singular abarca el plural, a menos que el contexto lo requiera de cualquier otra manera. En particular, cuando se usa el artículo indefinido, se entiende que la descripción contempla la pluralidad, así como también los singulares, a menos que el contexto lo requiera de cualquier otra manera.

Los rasgos, enteros, características, compuestos, porciones químicas o grupos descritos en conjunto con un aspecto particular, modalidad o ejemplo de la invención se entiende que son aplicables a cualquier otro aspecto, modalidad o ejemplo descrito en la presente descripción a menos que sea incompatible con éste.

Una modalidad de la invención será ahora descrita a manera de ejemplo solamente y con referencia a las siguientes figuras:

Figura 1: Vectores de agotamiento del ARNi de *SRSF1* humana modificados por ingeniería genética. (A) Diagramas de pCDNA6.2-GW/EmGFP-*SRSF1* miR1, EmGFP-*SRSF1* miR2 y EmGFP-*SRSF1* miR1+2. El pCDNA6.2-GW/EmGFP-*SRSF1* miR1 y EmGFP-*SRSF1* miR2 se construyeron por separado mediante el uso del kit de vector de expresión de ARNi miR de Pol II BLOCK-iT con EmGFP (ver métodos en línea). El casete de ARNi pre-miR2 para *SRSF1* humana se encadenó entonces mediante el subclonaje del fragmento cortado con BamHI/XhoI en los sitios BglII y XhoI de pcDNA6.2 GW/EmGFP-de SRSF1 humana miR1. (B) Diagrama del plásmido de EmGFP-de *SRSF1* humana miR1+2 clonado en el vector lentiviral autoinactivante (SIN-W-PGK) mediante el uso de los sitios de restricción SpeI y XhoI;

Figura 2: Cocultivos de HB9 de ratón: neuronas motoras GFP y astrocitos derivados de pacientes transducidos con SRSF1-ARNi. A) Evaluación de la eficiencia de la depleción del ARNi de *SRSF1* mediada por lentivirus en células HEK y iAstrocitos derivados de pacientes controles y con C9ORF72-ELA. *Los niveles del transcrito de SRSF1* se cuantificaron en células HEK transfectadas y en iAstrocitos transducidos con dosis aumentadas de MOI del LV-SRSF1-ARNi. Los niveles del transcrito *del ARNsn U1* se usaron para la normalización en experimentos por triplicado (Media ± SEM; ANOVA de dos vías, NS: no significativo, ****: p <0,001; N (reacciones qPCR) = 6); (b) Imágenes de alto contenido que muestran cómo el software de análisis Columbus reconoce las neuronas motoras Hb9:GFP y los axones que brotan de ellas sobre el fondo de astrocitos transducidos con SRSF1-ARNi;

Figura 3: El agotamiento de SRSF1 conduce a la reducción citoplasmática y a la acumulación nuclear de focos de ARN sentido. Imágenes representativas de focos de ARN sentido visualizados mediante el uso de Hibridación (rojo) *in situ* con fluorescencia Cy3-CCCCGG mediante microscopía confocal en i-Astrocitos no transducidos/transducidos. Los núcleos se tiñeron de azul mediante el uso de DAPI. Las flechas apuntan a focos de ARN. Las células con focos de ARN detectables representan aproximadamente el 15-40 % de la población celular, en dependencia de la línea de iAstrocitos derivada de pacientes individuales. La cuantificación se realizó en 20-25 células que contenían focos de ARN (ver las tablas de datos adicionales 1 y 4 para los conteos individuales y la figura 10E para el gráfico de barras);

Figura 4: Generación de construcciones repetidas ininterrumpidas de G4C2-sentido y C4G2-antisentido dependientes de RAN. (A) Geles de agarosa que confirman tanto la hibridación como la concatemerización de los oligonucleótidos G4C2x15. Las flechas apuntan a formas monoméricas y multiméricas de oligonucleótidos hibridados. (B) Se trataron oligonucleótidos triméricos con la nucleasa de frijol mungo para el clonaje romo. Gel de acrilamida al 8 % de digestiones HindIII/XhoI de construcciones pcDNA3.1/RAN que contienen repeticiones ininterrumpidas de G4C2-sentido y C4G2-antisentido con regiones flanqueantes 5' y 3'. (C) Curva estándar generada a partir del análisis del gel de acrilamida mediante el uso del software Gene Tools Image. (D) Tabla que muestra información sobre el tamaño de pares de bases teórico y experimental para las inserciones HindIII/XhoI. Las construcciones de RAN contienen al menos 38 repeticiones ininterrumpidas de G4C2-sentido o 39 repeticiones ininterrumpidas de C4G2-antisentido basadas en el tamaño experimental de las inserciones (ver las secuencias en E). (E) La secuenciación Sanger mediante el uso de betaína también se realizó en las direcciones 5' y 3' mediante el uso de cebadores de secuenciación del promotor T7 y T3, respectivamente. Cada lectura de secuencia cubrió la región flanqueante 5' o 3' y las repeticiones de 9-16 G4C2 o C4G2 antes de la interrupción. Las trazas de secuenciación están disponibles a solicitud. Los transcritos de ARN generados a partir de estas construcciones se resaltan en azul (regiones flanqueantes) y rojo (repeticiones sentido o antisentido). Las secuencias resaltadas en negro/subrayado corresponden al extremo 3' de la secuencia promotora y en negro/cursiva al inicio de la secuencia del terminador. Nótese la ausencia de codones de iniciación (ATG) tanto en los transcritos sentido como antisentido generados a partir de las construcciones de expresión de DPR dependientes de RAN. Los codones de parada se muestran en todos los marcos en negrita (SEQ ID NO 40 y 41). Las construcciones que expresan 15 repeticiones en orientación sentido o antisentido son idénticas excepto que solo contienen 15 repeticiones. El análisis de secuenciación y tamaño mostró además que el número de repeticiones se mantuvo estable durante múltiples rondas de transformación y replicación en *E. coli;*

Figura 5: Vectores de agotamiento del ARNi de *SRSF1*de ratón modificados por ingeniería genética. Diagramas del EmGFP-*SRSF1* miR1, EmGFP-*SRSF1* miR2 y EmGFP-*SRSF1* miR1 + 2 de ratón. El pCDNA6.2-GW/EmGFP-*SRSF1* miR1 y EmGFP-*SRSF1* miR2 se construyeron por separado mediante el uso del kit de vector de expresión de ARNi miR de Pol II BLOCK-iT con EmGFP (ver métodos en línea). El casete de ARNi pre-miR2 de *SRSF1* de ratón se encadenó luego mediante el subclonaje del fragmento cortado con BamHI/Xhol en los sitios BglII y Xhol de pcDNA6.2 GW/EmGFP-mouse *SRSF1* miR1;

Figura 6: El agotamiento de SRSF1 y la inhibición de la interacción SRSF1:NXF1 inhiben la exportación nuclear de los transcritos traducidos por RAN de repeticiones de hexanucleótidos y la producción de DPR. Cuantificación de DPR mediante transferencia Western. (A) Transferencia Western de células N2A cotransfectadas con un vector Ctrl o SRSF1-ARNi y un plásmido de cadena principal control (sin DPR Ctrl) o el mismo plásmido que expresa 38 repeticiones de G4C2-sentido ininterrumpidas (G4C2x38) o 39 repeticiones de C4G2-antisentido ininterrumpidas (C4G2x39) se cuantificaron en experimentos por triplicado (Media ± SEM; ANOVA de dos vías, ****: p <0,001; N = 3). En la Figura 11A. se presenta una réplica de las Transferencias Westerns. (B y C) Transferencia Western de células N2A cotransfectadas con plásmidos G4C2x38 (B) o C4G2x39 (C) y control (FLAG Ctrl) o SRSF1 aa11-196 de tipo salvaje marcado con FLAG (SRSF1), SRSF1-m2 o SRSF1-m4 se cuantificaron en experimentos por triplicado (Media ± SEM; ANOVA de dos vías, **: p <0,01, ****: p <0,001; N = 3). En la Figura 11C se presenta una réplica de cada transferencia Western; (d, e,) Transferencia Western mostrados en la Figura 11A para los paneles GP36 (d) y GA36 (e) se cuantificaron en experimentos por triplicado (promedio ± EEM; ANOVA bidireccional; N = 3). (f) Los extractos totales de células N2A transfectadas con FLAG control (FLAG ctrl) y aa11-196 de SRSF1 tipo salvaje marcado con FLAG (SRSF1), SRSF1-m2 o SRSF1-m4 se someten a inmunoprecipitación anti-FLAG. La coinmunoprecipitación de NXF1 endógeno se evaluó mediante el uso de anticuerpos anti-NXF1;

Figura 7: Agotamiento de SRSF1 en modelos de Drosophila que expresan DPR independientemente de las expansiones repetidas de G4C2. La expresión neuronal de repeticiones de dipéptidos de Gly-Arg (GR36) y Pro-Arg (PR36) causa déficits de escalada en adultos que no son rescatados por el agotamiento de SRSF1 (promedio ± EEM normalizada al Control (GAL4/Control-ARNi); Prueba de Kruskal-Wallis con prueba de comparación múltiple de Dunn, NS: no significativo, **: p <0,01, ****: p <0,001; N = Control: 188, GR36 + Ctrl-ARNi: 8, GR36 + SRSF1-ARNi: 7, PR36 + Ctrl-ARNi: 125, PR36 + SRSF1-ARNi: 119);

Figura 8: El agotamiento de SRSF1 o la inhibición de su secuestro por ARN repetido y su interacción con NXF1 inhiben la exportación nuclear de los transcritos traducidos por RAN para aliviar la neurotoxicidad. (a, b) Proteína: Ensayos de reticulación UV de ARN mediante el uso de proteínas recombinantes purificadas y sondas de ARN repetido de G4C2x5 (a) y C4G2x5 (b) radiomarcadas en el extremo 32P. Las proteínas se visualizan en SDS-PAGE teñidas con azul de Coomassie (paneles de la izquierda) y complejos de ARN: proteína unidos covalentemente mediante autorradiografía en Phosfolmages (paneles de la derecha). (c, d) Ensayos de inmunoprecipitación de ARN (RIP). Se añadió formaldehído al medio de células N2A vivas cotransfectadas con G4C2x15, G4C2x38 (c), C4G2x15 o C4G2x39 (d) y FLAG control (FLAG Ctrl), SRSF1 aa11-196 tipo salvaje marcado con FLAG (SRSF1) o SRSF1-m4 se sometieron a inmunoprecipitación anti-FLAG. El ARN purificado se analizó mediante qRT-PCR después de la normalización a los niveles de ARNnn U1 en experimentos por triplicado (promedio ± EEM; ANOVA de dos vías; N (reacciones qRT-PCR) = 6). (e) Control de calidad del fraccionamiento celular. Las células N2A cotransfectadas con G4C2x38 y los plásmidos Ctrl o SRSF1-ARNi o con G4C2x38 y aa11-196 de SRSF1 de tipo salvaje marcado con FLAG (SRSF1) o SRSF1-m4 se sometieron a fraccionamiento celular mediante el uso de condiciones de lisis hipotónica para producir fracciones citoplasmáticas. La calidad del fraccionamiento celular se comprobó mediante el uso de anticuerpos contra el factor SSRP1 de remodelación de la cromatina nuclear y alfa-tubulina. Esto se repitió para cada fraccionamiento y el ARN se mantuvo para el análisis de qRT-PCR si la fracción citoplasmática no mostraba contaminación nuclear. (g) Transferencia Western de Drosophila que expresan G4C2x36 y Ctrl o SRSF1-ARNi, sometidas a fraccionamiento celular mediante el uso de lisis hipotónica para producir fracciones citoplasmáticas. La histona H3 se usa para comprobar la contaminación nuclear potencial. (h) Los niveles de los transcritos de G4C2-sentido repetido citoplasmáticos y totales se normalizaron a los niveles de Tub84b en G4C2x36 + Ctrl-ARNi o G4C2x36 + SRSF1^{76}-ARNi *Drosophila* en experimentos por triplicado antes de graficar como una relación para tener en cuenta los cambios potenciales en la transcripción/estabilidad del ARNm (promedio ± EEM; prueba t pareada; N (reacciones qRT-PCR) = 3). En contraste con los niveles citoplasmáticos, los niveles totales de transcritos repetidos de hexanucleótidos no se alteraron significativamente tras la expresión de SRSF1-m4 o el agotamiento de SRSF1 en células o moscas (Figura 28B);

Figura 9: El agotamiento de SRSF1 previene *la neurodegeneración in vivo* y rescata la función locomotora. Los niveles de transcripción de(A) *SRSF1* (rojo y *ALYREF* (azul) se cuantificaron mediante análisis de qRT-PCR en líneas de noqueo independientes en moscas G4C2x36. *Los niveles del transcrito de Tub84b* se usaron para la normalización en experimentos por triplicado (promedio ± EEM; ANOVA de dos vías, NS: no significativo, ****: p <0,001; N (reacciones qRT-PCR) = 6). (B) La expresión dirigida de G4C2x36 provoca un fenotipo de ojo áspero que es completamente rescatado por *SRSF1*-noqueado. Se presentan los fenotipos oculares normales para las moscas Ctrl-ARNi y control G4C2x3. Las micrografías electrónicas de barrido representativas se presentan debajo de las imágenes de microscopía óptica. Barra de escala: 100 µm. (C y D) La expresión neuronal de G4C2x36 causa déficits de arrastre de larvas (C) y de la escalada de adultos (D) que son rescatados por el agotamiento de

SRSF1 (promedio ± EEM normalizada para control (Ctrl)-ARNi); prueba de Kruskal-Wallis con comparación múltiple de Dunn, NS: no significativo, **: p <0,01, ***: p <0,005, ****: p <0,001; N (larvas) = 10; N (adultos) = Control: 93, G4C2x3: 62, G4C2x36 + Ctrl-ARNi: 51, G4C2x36 + SRSF1-ARNi: 50, G4C2x36 + ALYREF-ARNi: 36) (e) Agotamiento de SRSF1 en modelos de Drosophila que expresan DPR independientemente de las expansiones de repetición de G4C2 y la traducción por RAN16. (f) La expresión neuronal de poli-Gly-Arg DPR (GR36) y poli-Pro-Arg DPR (PR36) provoca déficits de escalada en adultos que no se restaura con el agotamiento de SRSF1 (media ± 95 % CI normalizado al Control; N = Control (GAL4/luciferasa-ARNi): 239, GR36 + Ctrl-ARNi: 12, GR36 + SRSF1-ARNi: 7, PR36 + Ctrl-ARNi: 125, PR36 + SRSF1-ARNi: 119);

Figura 10: El agotamiento de SRSF1 suprime la toxicidad astrocítica mediada por C9ORF72 derivada del paciente y la muerte de neuronas motoras. (A) Las células HEK293T (UT) no transfectadas, control (Ctrl)-ARNi, *SRSF1*-ARNi o LV-*SRSF1*ARNi y SRSF1 etiquetadas con FLAG se analizaron 72 h después de la transfección mediante inmunotransferencia mediante el uso de anticuerpos anti-FLAG, anti-ALYREF y anti-α-tubulina como control de carga. (B) Los niveles de transcritos de *SRSF1* y *ALYREF* se cuantificaron mediante análisis qRT-PCR después de la normalización a los niveles de *U1* ARNsn en experimentos por triplicado (promedio ± EEM; ANOVA de dos vías, NS: no significativo, ****: p <0,001; N (reacciones qRT-PCR) = 6). Se usaron tres líneas de pacientes controles (Ctrl-pat) y tres C9ORF72-ELA (C9-ELA-pat) codificadas por colores. (C) Imágenes de microscopía de inmunofluorescencia representativas de cocultivos de iAstrocitos y neuronas motoras. Las flechas señalan ejemplos de axones de neuronas motoras. Barra de escala: 100 µm. (D) Las mismas líneas de pacientes controles (Ctrl-pat) y C9ORF72-ELA (C9-ELA-pat) codificadas por colores (panel B) se usan para la cuantificación de los conteos de neuronas motoras (MN) + Hb9-GFP en cuatro cocultivos de astrocitos y neuronas motoras por replicados (promedio ± EEM; ANOVA de una vía, NS: no significativo, ***: p <0,005, ****: p <0,001; N (células) = Ctrl-pat154: 567/589/582/500, Ctrl-pat154+SRSF1-ARNi: 620/543/504/349, Ctrl-pat155: 602/610/553/571, Ctrl-pat155+SRSF1-ARNi: 554/584/532/516, Ctrl-pat209: 519/599/584/535, Ctrl-pat209+SRSF1-ARNi: 617/486/425/572, C9-ELA-pat78: 352/279/294/258, C9-ELA-pat78+SRSF1-ARNi: 569/451/398/583, C9-ELA-pat183: 200/188/154/145, C9-ELA-pat183+SRSF1-ARNi: 480/420/380/399, C9-ELA-pat201: 201/243/261/224, C9-ELA-pat201+SRSF1-ARNi: 486/463/444/485). (E) Cuantificación de focos de ARN sentido nuclear y citoplasmático en *iAstrocitos transducidos con ARNi de SRSF1-* (promedio ± EEM; ANOVA de dos vías, NS: no significativo, **: p <0,01, ***: p <0,005, ****: p <0,001; N (células con 1-5 focos de ARN) = C9-ELA-pat78 + MOI0: 21, C9-ELA-pat78 + MOI5: 20, C9-ELA-pat78 + MOI7: 22, C9-ELA-pat183 + MOI0: 21, C9-ELA-pat183 + MOI5: 24, C9-ELA-pat183 + MOI7: 22, C9-ELA-pat201 + MOI0: 24, C9-ELA-pat201 + MOI5: 23, C9-ELA-pat201 + MOI7: 22). > 95% de las células con focos de ARN presentaron un total de 5 focos o menos;

Figura 11: El agotamiento de SRSF1 y la interacción SRSF1 bloqueada: TAP/NXF1 inhiben la exportación nuclear de los transcritos traducidos por RAN de repeticiones de C9ORF72 y la producción de DPR. (A) Transferencia Western de células N2A cotransfectadas con Ctrl o *el vector de ARNi de SRSF1*y el plásmido de cadena principal control (sin DPR Ctrl) o el mismo plásmido que expresa 38 repeticiones ininterrumpidas de G4C2-sentido (G4C2x38) o 39 repeticiones ininterrumpidas de C4G2-antisentido (C4G2x39). Las proteínas de poli-Gly-Pro sentido/antisentido y poli-Gly-Ala DPR sentido se producen mediante traducción por RAN de repeticiones internas en ausencia de un codón de inicio inicial (Figura 4) (secuencias de nucleótidos en la Figura 4e. Un control con especificidad de repeticiones de hexanucleótidos se proporcionó mediante la cotransfección de plásmidos que expresan poli-Gly-Ala (GAx36) o poli-Gly-Pro (GPx36) independientemente de las secuencias de nucleótidos con secuencias repetidas de G4C2/C4G2 en la Figura 26) y los vectores Ctrl o de ARNi de *SRSF1-*. (B) Células N2A cotransfectadas con G4C2x38 y plásmidos Ctrl o de ARNi de *SRSF1* (parte izquierda) y el tipo salvaje aa11-196 de SRSF1 marcado con FLAG (SRSF1) o SRSF1-m4 (parte derecha) se sometieron a fraccionamiento celular mediante el uso de lisis hipotónica para producir fracciones citoplasmáticas (figura 8e). Los niveles de los transcritos repetidos de G4C2 sentido de células completas (total) y citoplasmáticas se normalizaron a los niveles de *U1*ARNsn en experimentos por triplicado antes de representarlos como una relación para tener en cuenta los cambios potenciales en la transcripción/estabilidad del ARNm (promedio ± SEM; ANOVA de una vía, NS: no significativo, **: p <0,01; N (reacciones qRT-PCR) = 6). (**C**) Transferencia Western de células N2A cotransfectadas con plásmidos G4C2x38 o C4G2x39 y controles (FLAG Ctrl) o del tipo salvaje de aa11-196 de SRSF1 marcado con FLAG (SRSF1), SRSF1-m2 o SRSF1-m4. (**D**) El ensayo de proliferación celular MTT se realizaron en células N2A transfectadas con plásmidos G4C2x38 o C4G2x39 y Ctrl-ARNi, SRSF1-ARNi, FLAG Ctrl, SRSF1 o SRSF1-m4 en experimentos por triplicado (promedio ± SEM; ANOVA de una vía, NS: no significativo, **: p <0,01, ***: p <0,005, ****: p <0,001; N (OD_{650} valores) = 12);

Figura 12: Secuencias de ADN usadas para generar vectores de terapia génica recombinantes; A secuencias humanas de SRSF1; A1) Número de acceso a NCBI NM_006924.4 (secuencia de ARNm, variante de transcrito 1) - Secuencia de codificación CDS: 210-956; A2) Secuencia de aminoácidos codificada por NM_006924.4; A3) Número de acceso a NCBI NM_001078166.1 (secuencia de ARNm, variante de transcrito 2) - Secuencia de codificación CDS: 210-815; A4) Secuencia de aminoácidos codificada por NM_001078166.1; A5) Número de acceso a NCBI NR_034041.1 (secuencia de ARNm, variante de transcrito 3) - Considerado no codificante como sustrato potencial para la desintegración mediada por sin sentido; B) Secuencias de ratón de *SRSF1; B1)*_ número de acceso a NCBI BC046773.1 (secuencia de ARNm) - Secuencia de codificación CDS: 99-845; B2) Secuencia de aminoácidos codificada por BC046773.1: B3) Número de acceso a NCBI NM_173374.4 (secuencia de ARNm, variante de transcrito 1) - Secuencia de codificación CDS: 467-1213. Codifica la misma secuencia de aminoácidos

que BC046773.1; B4) Secuencia de aminoácidos codificada por NM_173374.4: B5) Número de acceso a NCBI NM_001078167.2 (secuencia de ARNm, variante de transcrito 2) - Secuencia de codificación CDS: 467-1072; B6) Secuencia de aminoácidos codificada por NM_001078167.2;

Figura 13 Expresión o fusión celular permeable de secuencias de SRSF1 que codifican un mutante negativo dominante que no interactúa con NXF1. Uso de cualquier secuencia que abarque la expresión de SRSF1 humano o secuencias ortólogas que codifiquen los dos motivos de reconocimiento de ARN (aminoácidos 11-196) y mutaciones de las argininas 90, 93, 117, 118 en la región enlazadora entre los 2 motivos de reconocimiento de ARN. Estos corresponden a los siguientes aminoácidos para SRSF1 humano, incluidas las mutaciones de las argininas 90, 93, 117, 118 en alanina (marcadas en negrita): Se pueden usar otros aminoácidos para la sustitución, por ejemplo, glicina, valina, etc. A) aminoácidos 1-196 de SRSF1 R90, 93, 117, 11, B) aminoácidos 11-196 de SRSF1 R90,93,117,118A; C; Aminoácidos 1-248 de SRSF1 R90,93,117,118A;

Figura 14 Expresión o fusión celular permeable de secuencias de SRSF1 que codifican péptidos antagonistas que interactúan con NXF1. Uso de cualquier secuencia de SRSF1 ortóloga que codifique péptidos que se unan a NXF1. Estos abarcan los aminoácidos 89-120 de SRSF1 humana;

Figura 15 Expresión de los aminoácidos 1-655 de SRPK1 para mantener el estado de fosforilación de SRSF1 para inhibir la interacción de SRSF1 con NXF1 A) Número de acceso a NCBI NM_003137.4 (secuencia de ARNm, variante de transcrito 1) - Secuencia de codificación CDS: 125-2092; B) Secuencia de aminoácidos codificada por NM_003137.4;

Figura 16a secuencia de nucleótidos de CLK1 o Clk/Sty (CDC tipo quinasa 1 (CLK1) de Homo sapiens, variante de transcrito 1, ARNm) - Secuencia de referencia en NCBI: NM_004071.3; Figura 16b secuencia de aminoácidos de CLK1;

Figura 17a secuencia de nucleótidos de PP1 (ARNm de la subunidad catalítica de la proteína fosfatasa-1 humana, cds completos) - GenBank: M63960.1; secuencia de aminoácidos de PP1 humana; La figura 17b es la secuencia de aminoácidos de PP1;

Figura 18 Secuencias promotoras usadas para generar vectores de terapia génica recombinantes; A) 6.1/ potenciador temprano de CMV / promotor híbrido de b-actina de pollo (CAG); B) promotor de CMV; C) promotor CBA (b-actina de pollo) D) promotor neuronal específico (Sinapsina 1); E) promotor específico glial (GFAP); F) promotor H1; G) promotor U6; y h) promotor H1_pLVTHM, j) promotor H1 en la cadena principal de scAAV, k) promotor U6, l) promotor de β-actina de pollo (CBA), m) promotor de CAG (promotor híbrido de β-actina de pollo), n) promotor de PGK de ratón de LV-LacZ, o) factor de elongación corto 1a (EFS1a), p) potenciador de CMV o q) promotor de CMV;

Figura 19 Serotipo del virus adenoasociado 1, 2, 5, 6, 8, 9, 10 A) AAV-CMV-GFP autocomplementario; B) pAAV2/9 - Para la construcción de scAAV9;

Figura 20 Las proteínas ALYREF y SRSF1 purificadas interactúan directamente con ARN sentido y antisentido con repeticiones de hexanucleótidos. (a, b) Proteína: Ensayos de reticulación de ARN UV mediante el uso de proteínas recombinantes purificadas y sondas repetidas de ARN G4C2x5 radiomarcadas en el extremo 32P (a) y C4G2x5 (b). Las proteínas se visualizan en SDS-PAGE teñidas con azul de Coomassie (paneles de la izquierda) y complejos de ARN: proteína unidos covalentemente mediante autorradiografía en Phosfolmages (paneles de la derecha). La exposición a rayos UV está indicada por +. (c) Las imágenes de microscopía confocal de fluorescencia muestran la colocalización de SRSF1 (marcado en verde por inmunofluorescencia) y los focos de ARN sentido (marcados en rojo mediante el uso de Hibridación In Situ de Fluorescencia) en neuronas motoras de tejidos de médula espinal posteriores a la muerte de dos casos de C9ORF72-ELA humanos. Los núcleos se tiñen en azul mediante DAPI. Barra de escala: 3 µm;

Figura 21 Generación de modelos de células neuronales que recapitulan la traducción dependiente de RAN de DPR sentido y antisentido. (a) representaciones esquemáticas de las construcciones. (b) Focos de ARN de G4C2 sentido teñidos con sonda C4G2 antisentido marcada con Cy3. Se usó DAPI para teñir los núcleos de células neuronales N2A en azul. Barra de escala: 5 µm. (c) Transferencia Western de células N2A transfectadas con un plásmido de cadena principal control (sin DPR Ctrl) o el mismo plásmido que expresa 15 repeticiones ininterrumpidas de G4C2 sentido (G4C2x15), 38 repeticiones ininterrumpidas de G4C2 sentido (G4C2x38), 15 repeticiones ininterrumpidas de C4G2 antisentido (C4G2x15) o 39 repeticiones ininterrumpidas de C4G2-antisentido (C4G2x39). Las membranas se probaron con anticuerpos contra poli-Gly-Pro DPR, poli-Gly-Ala DPR o D-tubulina como control de carga. (d) Ensayo de proliferación celular MTT realizado en células N2A transfectadas con un plásmido de cadena principal control (sin DPR Ctrl) o el mismo plásmido que expresa varias longitudes de transcritos repetidos ininterrumpidas de hexanucleótidos en experimentos por triplicado (promedio ± SEM; ANOVA de una vía; N (Valores de DO650) = 12);

Figura 22 El agotamiento de SRSF1 o la inhibición de su secuestro de ARN repetido y de su interacción con NXF1 inhiben la producción de DPR en neuronas primarias. (a) Microscopía de inmunofluorescencia de neuronas corticales de rata cultivadas. Se detectaron DPR en el canal rojo mediante el uso de anticuerpos anti-V5 y anti-ratón ALEXA594. Las construcciones Ctrl-ARNi y SRSF1-ARNi coexpresan GFP mientras que las proteínas SRSF1 etiquetadas con FLAG se tiñeron mediante el uso de un anticuerpo anti-FLAG conjugado con FITC que permite la detección y la cuantificación de neuronas transfectadas en el canal verde. Barra de escala: 5 µm. (b, c) La evaluación estadística de los conteos de neuronas corticales se realizó a partir de aproximadamente 100 neuronas transfectadas para cada grupo (prueba exacta de Fisher; N (neuronas transfectadas) = Ctrl-ARNi: 95, SRSF1-ARNi: 112, SRSF1: 106, SRSF1-m4: 121);

Figura 23 El agotamiento de SRSF1 inhibe específicamente la exportación nuclear de transcritos repetidos de C9ORF72 que retienen expansiones repetidas de hexanucleótidos en el intrón-1 en neuronas derivadas de pacientes. (a) La longitud axonal de los iMN derivados del paciente tratados o no con SRSF1-ARNi se evaluó mediante imágenes de alto contenido en experimentos por triplicado (promedio ± SEM; ANOVA de una vía, NS: no significativo; N (longitud media del axón/pocillo) = 9). (b) El área del cuerpo celular de los iMN derivados del paciente tratados o no con SRSF1-ARNi se evaluó mediante imágenes de alto contenido en experimentos por triplicado (promedio ± SEM; ANOVA de una vía; N (área promedio del cuerpo celular/pocillo) = 9). (c) La supervivencia de los iMN Ctrl o C9ORF72-ELA derivados del paciente tratados o no con SRSF1-ARNi se cuantificó en cocultivos con iAstrocitos Ctrl o C90RF72-ELA derivados del paciente en seis experimentos replicados el día 4 (promedio ± SEM; ANOVA de una vía; N (iMN) = Ctrl-pat209: 142/165/174/117/122/168, C9-ELA-pat78: 77/65/41/68/71/70; C9-ELA-pat78 + SRSF1-ARNi: 106/84/81/97/113/83; C9-ELA-pat201: 68/69/68/62/66/74; C9-ELA-pat201 + SRSF1-ARNi: 131/104/96/82/111/104). (d) Transferencia Western de iNeurons diferenciadas de pacientes control (Ctrl-pat 154, Ctrl-pat155) y C90RF72-ELA (C9-ELA-pat78, C9-ELA-pat183) tratados o no con LV-SRSF1-ARNi (MOI 0 o 5) se sometieron a fraccionamiento celular mediante el uso de lisis hipotónica para producir fracciones citoplasmáticas. El factor de remodelación de la cromatina SSRP1 se usa para comprobar la contaminación nuclear potencial en las fracciones citoplasmáticas. Se usó el agotamiento de actina en las fracciones nucleares para comprobar la calidad de las fracciones nucleares. La eficacia de SRSF1-ARNi también fue validada por qRT-PCR y se logró aproximadamente un 80 % de eliminación del ARNm de SRSF1 (Figura 29). (e) Los niveles totales, nucleares y citoplasmáticos de los transcritos de C9ORF72 empalmados en el intrón1 (medidos por la unión exón1-exón3) se cuantificaron en experimentos por duplicado mediante qRT-PCR después de la normalización a los niveles de ARNnn de U1 y al 100 % en pacientes controles a MOI0 (promedio ± SEM; ANOVA de una vía, NS: no significativo; N (reacciones qRT-PCR) = 8). (f) Los niveles de transcritos de C9ORF72 empalmados en el intrón1 normalizados a los niveles de ARNnn de U1 (panel d) se representaron como una relación SRSF1-ARNi MOI5 sobre MOI0 para evaluar el efecto específico debido al SRSF1-ARNi (promedio ± SEM; ANOVA de una vía; N (reacciones qRT-PCR) = 8). (g) Los niveles totales, nucleares y citoplasmáticos de los transcritos de C9ORF72 sin empalmar que retienen el intrón1 (medidos por la unión exón1-intrón1) se cuantificaron en experimentos por duplicado mediante qRT-PCR después de la normalización a los niveles de ARNnn de U1 y al 100 % para los pacientes controles a MOI0 (promedio ± SEM; ANOVA de una vía, NS: no significativo; N (reacciones qRT-PCR) = 8). (h) Los niveles de transcritos de C9ORF72 no empalmados que retienen el intrón1 normalizados a los niveles de ARNnn de U1 (panel f) se representaron como una relación SRSF1-ARNi MOI5 sobre MOI0 para evaluar el efecto específico debido al SRSF1-ARNi (promedio ± SEM; ANOVA de una vía; N (reacciones qRT-PCR) = 8);

Figura 24 Modelo para la exportación nuclear de transcritos repetidos de hexanucleótidos C9ORF72 patológicos y manipulación terapéutica. (a) La exportación nuclear de los transcritos de C9ORF72 sentido y antisentido que retienen las repeticiones de hexanucleótidos expandidos en el intrón1 depende específicamente del secuestro de SRSF1 y su interacción con el receptor de exportación nuclear NXF1. Por el contrario, la exportación nuclear de los transcritos de C9ORF72 empalmados en el intrón1 necesarios para la producción de la proteína C9ORF72 no implica la interacción de SRSF1 con NXF1, sin embargo, quedan por identificar el(los) adaptador(es) de exportación nuclear (NEA). (b) El agotamiento de SRSF1 inhibe específicamente la exportación nuclear de los transcritos de C9ORF72 que retienen las repeticiones de hexanucleótidos expandidos en el intrón1, probablemente debido a una reducción en el secuestro de SRSF1 endógeno en las repeticiones de hexanucleótidos de C9ORF72 y la incapacidad para remodelar anormalmente NXF1 en un modo de unión elevado al ARN, mientras que no afecta los niveles de expresión y de corte y empalme/retención del intron1. Además, el agotamiento de SRSF1 no afecta los niveles de expresión, el corte y empalme del intrón1 o la exportación nuclear de los transcritos de C9ORF72 cortados y empalmados en el intrón1 de tipo salvaje requeridos para la producción de la proteína C9ORF72. (c) La sobreexpresión de la proteína SRSF1-m4, que no interactúa eficazmente con NXF1, compite con la SRSF1 endógena por el secuestro en repeticiones de hexanucleótidos lo que previene a su vez las interacciones con NXF1 y la exportación nuclear de transcritos repetidos de C9ORF72;

Figura 25 El agotamiento parcial de SRSF1 conduce a una reducción prominente de las DPR poli-GP sentido y antisentido en Drosophila y neuronas motoras derivadas de pacientes con C90RF72-ELA. (a) Extractos de proteínas totales de larvas de Drosophila G4C2x3 + Ctrl-ARNi, G4C2x36 + Ctrl ARNi y G4C2x36 + SRSF1-ARNi se analizaron mediante transferencias puntuales mediante el uso de anticuerpos poli-GP y a-tubulina como control de carga. (b) Extractos de proteínas totales de células HEK transfectadas con plásmidos Ctrl o GP36 y iMN

derivados de pacientes transducidos (MOI =5) o no (MOI =0) con virus LV-SRSF1-ARNi se analizan mediante transferencias puntuales mediante el uso de anticuerpos poli-GP DPR y alfa-tubulina como control de carga;

Figura 26 Generación de construcciones sintéticas que expresan DPR independientemente de la traducción por RAN y los hexanucleótidos repetidos de G4C2. (a) Secuencia de nucleótidos que codifica poli-Gly-Pro x36 DPR. (b) Secuencia de nucleótidos que codifica poli-Gly-Ala x36 DPR. La secuencia DPR está resaltada en azul. El codón de inicio ATG se resalta en rojo mientras que el codón de terminación TAA se resalta en negrita. Una etiqueta V5 también está presente y resaltada en verde.

Figura 27 La pérdida parcial de sbr/NXF1 restaura los déficits locomotores en moscas que expresan G4C2x36. La expresión neuronal de G4C2x36 causa déficits de arrastre de larvas (a) y de la escalada de adultos (b) que son restaurados por el agotamiento de sbr (promedio ± 95 % de CI normalizado al Control; N (larvas) = 10; N (adultos) = Control (GAL4/luciferasa-ARNi): 105, G4C2x36 + Ctrl-ARNi: 70, G4C2x36 + sbr-ARNi: 72);

Figura 28 El agotamiento de SRSF1 o la inhibición de su secuestro e interacción con NXF1 alteran los niveles citoplasmáticos de los transcritos repetidos de hexanucleótidos, pero no sus niveles totales.

(a) Las células N2A cotransfectadas con G4C2x38 y los plásmidos Ctrl o SRSF1-ARNi (parte izquierda) y aa11-196 de SRSF1 tipo salvaje marcado con FLAG (SRSF1) o SRSF1-m4 (parte derecha) se sometieron a fraccionamiento celular mediante el uso de lisis hipotónica para producir fracciones citoplasmáticas (Figura 6e). Los niveles de transcritos repetidos de G4C2 sentido total y citoplasmático se normalizaron a niveles de ARNnn de U1 en experimentos por triplicado (promedio ± SEM; ANOVA de una vía; N (reacciones qRT-PCR) = 6).
(b) Drosophila que expresa G4C2x36 y control (Ctrl)-ARNi o SRSF1-ARNi. Se sometieron moscas completas a fraccionamiento celular mediante el uso de lisis hipotónica para producir fracciones citoplasmáticas (Figura 6g). Los niveles citoplasmáticos totales de transcritos repetidos de G4C2 sentido se normalizaron a los niveles de Tub84b en experimentos por triplicado (promedio ± SEM; prueba t pareada; N (reacciones qRT-PCR) = 3);

Figura 29 Secuencia del transcrito. Los recuadros representan los sitios de clonación HindIII (AAGCTT) y XhoI (CTCGAG) usados para clonar los oligonucleótidos hibridados G4C2x38 (Figura 3). Una construcción sintética que codifica las etiquetas 3x V5 (secuencias resaltadas en naranja, verde y violeta) con 3 codones de parada (TAA, subrayado/negrita) se clonó en una segunda etapa mediante el uso de los sitios NotI (GCGGCCGC) y XbaI (TCTAGA). Los transcritos de ARN generados a partir de esta construcción se resalta en azul (regiones flanqueantes) y rojo (38 repeticiones de G4C2 sentido) y naranja, verde y violeta (etiquetas 3x V5). Las secuencias resaltadas en negro/subrayadas corresponden al extremo 3' de la secuencia promotora y en negro/cursiva al comienzo de la secuencia del terminador. Nótese la ausencia de codones de iniciación (ATG) en el transcrito generado a partir de la construcción de expresión de DPR dependiente de RAN. El análisis de la secuenciación y del tamaño mostró además que el número de repeticiones permaneció estable durante múltiples rondas de transformación y replicación en NEB® 10-beta E. coli (New England Biolabs);

Figura 30 Diferenciación de iNPC de neuronas derivadas de fibroblastos de pacientes. Se realizó microscopía de inmunofluorescencia Tuj1 en neuronas diferenciadas de células progenitoras neurales inducidas (iNPC) derivadas de fibroblastos de pacientes controles (Ctrl-pat154) o C90RF72-ELA (C9-ELA-pat78) mediante el uso del canal rojo. Se usó DAPI para teñir los núcleos en azul;

Figura 31 Evaluación de la eficacia de la depleción de SRSF1-ARNi mediada por lentivirus en iNeuronas derivadas de pacientes controles y C90RF72-ELA. Los niveles de transcritos de SRSF1 se cuantificaron en células HEK transfectadas y en iAstrocitos transducidos con dosis aumentadas de MOI de LV-SRSF1-ARNi. Los niveles de transcritos de ARNnn de U1 se usaron para la normalización en dos líneas celulares controles (pat154, pat 155) o C9-ELA (pat78, pat183) por duplicado (promedio ± SEM; ANOVA de dos vías; N (reacciones qPCR) = 8); y

Figura 32 Un péptido permeable celular inhibitorio de SRSF1 inhibe la producción de DPR expresadas de una manera dependiente de RAN relevante para la enfermedad en células HEK humanas transfectadas con una construcción repetida de G4C2x38 que expresa etiquetas 3xV5 en todos los marcos. Secuencia del péptido: PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE GG GKPIPNPLLGLDST GG YGRKKRRQRRR.

Tabla 1

	MOI=0			MOI=5			MOI=7		
Focos de ARN/célula	C9 Pat. 78	C9 Pat. 183	C9 Pat. 201	C9 Pat. 78	C9 Pat. 183	C9 Pat. 201	C9 Pat. 78	C9 Pat. 183	C9 Pat. 201
Focos de ARN nuclear	0,86	0,95	0,88	1,20	1,08	1,48	1,82	1,68	1,95
Focos de ARN citoplasmático	0,71	0,71	0,58	0,35	0,50	0,35	0,23	0,27	0,23
SEM									
Focos de ARN nuclear		0,14	0,18	0,15	0,17	0,18	0,18	0,20	0,26	0,23
Focos de ARN citoplasmático		0,16	0,12	0,13	0,13	0,13	0,10	0,09	0,10	0,09

Tabla 2 Ejemplos de secuencias de miARN de SRSF1 humana

Éstos implican la expresión de casetes de pre-miARN que se procesan en la célula mediante el uso de la maquinaria fisiológica de ARNi (escisión de Dicer para generar miARN maduro de -22 nucleótidos e incorporación en RISC (complejo de silenciamiento inducido por ARN).

10 secuencias de ARNm de SRSF1 humano dirigidas por miARN maduro:

SEQ ID NO	número de miARN	*Región SRSF1*	Inicio	Secuencia diana (5' - 3') de 21 nt *de SRSF1*
5	1	RRM2	767	AACTGCCTACATCCGGGTTAA
6	2	RRM2	784	TTAAAGTTGATGGGCCCAGAA
7	3	RRM2	795	GGGCCCAGAAGTCCAAGTTAT
8	4	Dominio RS	900	AGCAGAGGATCACCACGCTAT
9	5	3'UTR	1422	TCAATAATGGAGGCAATGGTA
10	6	3'UTR	1436	AATGGTATGACTCCAAGTGCT
11	7	3'UTR	1695	GCTAATTTGTCACAGTGCTTA
12	8	3'UTR	1725	GTTAATGTGTGACCTGCTGTT
13	9	3'UTR	1804	ACTGCTAAATCTGCATGTCCT
14	10	3'UTR	1834	TGATAGAGCGTTGCTATTTCA

RRM2: Motivo de reconocimiento de ARN 2; RS: Rico en arginina-serina; 3'UTR: Región 3' no traducida Tabla 3 Ejemplos de secuencias de miARN de SRSF1 de ratón

10 secuencias de ARNm de SRSF1 de ratón dirigidas contra el miARN maduro:

SEQ ID NO	número de miARN	*Región SRSF1*	Inicio	Secuencia diana (5' - 3') de 21 nt *de SRSF1*
5	1	RRM2	656	AACTGCCTACATCCGGGTTAA
6	2	RRM2	673	TTAAAGTTGATGGGCCCAGAA
49	3	RRM2	684	GGGCCCAGAAGTCCAAGTTAT
8	4	Dominio RS	789	AGCAGAGGATCACCACGCTAT
51	5	3'UTR	1105	AATGTCTATTCTGCTCTGGTT
52	6	3'UTR	1843	AAATTGCAGATGGGAGCAATA
53	7	3'UTR	1846	TTGCAGATGGGAGCAATAGTT
54	8	3'UTR	1853	TGGGAGCAATAGTTTAGGTTT
55	9	3'UTR	1866	TTAGGTTTAGGTGGGTAGTAA
56	10	3'UTR	1867	TAGGTTTAGGTGGGTAGTAAT
RRM2: Motivo de reconocimiento de ARN 2; RS: Rico en arginina-serina; 3'UTR: región 3' no traducida				

Tabla 4: Número y distribución celular de focos de ARN de G4C2 en astrocitos derivados de iNPC tratados con MOI creciente de LV-SRSF1-ARNi

	Astrocitos C90RF72-ELA Paciente 78						Astrocitos C9ORF72-ELA Paciente 183						Astrocitos C9ORF72-ELA Paciente 201					
MOI de ARNide SRSF1-	0		5		7		0		5		7		0		5		7	
Focl	Nuc	Cito	Nuc	Cito	Nuc	Cito	Nuc	Cito	Nuc	Cito	Nuc	Cito	Nuc	Cito	Nuc	Nuc	Nuc	Cito
Conteos primarios	1	0	0	1	2	1	0	1	1	0	1	0	1	0	1	0	4	1
	2	0	1	0	1	0	2	0	1	0	1	0	2	0	1	0	2	1
	1	0	2	0	1	0	1	0	1	0	1	0	1	0	2	0	4	0
	1	1	0	2	2	0	1	0	0	1	2	1	1	0	1	0	1	1
	1	0	1	0	1	0	1	1	1	0	1	1	1	1	3	1	3	0
	0	2	1	0	2	1	0	1	2	0	1	0	3	1	1	0	1	0
	3	1	1	0	3	0	1	1	2	0	1	0	2	0	2	0	2	0
	0	1	1	0	2	0	1	1	3	0	1	0	0	1	2	0	1	0
	0	1	2	0	1	0	2	0	1	2	1	0	1	0	1	0	3	0
	1	0	1	0	1	0	0	1	0	1	3	0	1	0	1	1	3	0
	1	1	2	0	4	1	0	1	1	0	1	0	1	1	1	0	2	0
	1	1	1	0	1	0	1	0	1	0	1	0	1	0	1	1	3	0
	2	0	3	1	2	0	1	0	1	0	5	1	0	2	2	0	1	0
	1	2	2	0	1	0	0	1	1	2	4	1	1	0	1	0	2	0
	0	1	1	0	3	0	2	1	0	1	2	0	1	2	2	1	1	0
	1	0	1	1	4	1	1	2	1	1	1	0	0	1	1	0	3	1
	1	0	1	0	1	0	1	0	1	1	3	0	1	0	0	1	1	0
	0	2	0	1	1	0	1	1	1	0	1	0	1	0	2	0	1	0
	2	1	1	0	2	0	0	1	1	0	3	1	1	0	1	1	2	0
	1	0	2	1	1	0	1	1	1	0	2	0	0	1	1	1	2	0
	1	1			2	0	3	1		1	0	1	0	1	1	0	1	0
					2	1			1	0	1	0	1	1	2	0	0	1
									0	1			0	1	4	1		
									3	1			0	1				
Promedio	0,857	0,714	1,200	0,350	1,818	0,227	0,952	0,714	1,083	0,500	1,682	0,273	0,875	0,583	1,478	0,348	1.955	0.227
S.E.M.	0,143	0,156	0,172	0,131	0,204	0,091	0,178	0,122	0,180	0,135	0,258	0,097	0,151	0,133	0,176	0,102	0.232	0.091

Nuc. Nuclear; Cito = citoplasmático

Tabla 5 Listado y características de las células derivadas de pacientes usadas en este estudio.

Muestra del paciente	Etnia	Género	Tipo celular	Edad al colectar la biopsia
78	Caucásico	Masculino	C9ORF72	66
183	Caucásico	Masculino	C9ORF72	49
201	Caucásico	Femenino	C9ORF72	66
154	Caucásico	Femenino	Control	55
155	Caucásico	Masculino	Control	40
209	Caucásico	Femenino	Control	69

Tabla 6 Ejemplos de ARNhc que se dirigen a SRSF1 humano (NM_006924.4)

SEQ ID NO	número de ARNhc	*Región SRSF1*	Inicio	secuencia diana (5' - 3') de 21 nt de SRSF1
47	1	5'UTR	25	GGGATCGAATCGCTGTTTCCT
48	2	5'UTR	26	GCTGTTTCCTTCCGCTTCTCT
50	3	ORF	657	GCTGATGTTTACCGAGATGGC
99	4	ORF	689	GGAGTTTGTACGGAAAGAAGA
154	5	ORF	700	GGAAAGAAGATATGACCTATG
155	6	ORF	760	AGGGAGAAACTGCCTACATCC
156	7	3'UTR	1673	GGGACTAATGTGGGAAGAACT
157	8	3'UTR	2084	GCAACCACGAAACCTGTAATA
158	9	3'UTR	2470	GGGATCAGATTACCAGGAACA
60	10	3'UTR	3120	GCATCTGAAAGATAAGCTTCT

Materiales y procedimientos

Líneas de ARNi de horquilla corta inducible de Drosophila**.** Las líneas (UAS-ARNi) se obtuvieron del Centro de recursos de Drosophila de Viena: http://stockcenter.vdrc.at/control/main

Líneas de ARNi de SRSF1 (SF2/ASF):

v27775: ID base de mosca = FBst0457117
v27776: ID base de mosca = FBst0457118

Líneas de inserción independientes, ambas líneas llevan la siguiente secuencia de repetición invertida: (SEQ ID NO 57)

ATGCCGACGA TGCGGTGAAG GCGCGCGACG GCTACGACTA CGATGGGTAT
CGTCTGCGCG TGGAGTTCCC GCGGGGCGGT GGTCCTGGAA GCTACCGCGG
CGGCAACCGC AATGACCGAA GCCGCGACGG TGGGGGACGG ATGGGCGGAC
GCGGACCGCC AGCCAAGCGC TCGCAGTACC GCGTCATGGT TACTGGACTG
CCCGCCTCCG GATCGTGGCA AGATCTCAAG GATCACATGC GCGAGGCCGG
CGACGTCTGC TTCGCGGACA CTTACAAGGA TGGTTCCGGC GTCGTTGAGT
TCCTGCGCCA CGAGGACATG AAGTACGCAA TCAAAAAATT GGACGACTCT
CGCTTCCGA

Líneas de ARNi de ALYREF (Ref1):

v12301 (GD): ID base de mosca = FBst0450381 - la línea lleva la siguiente secuencia de repetición invertida (SEQ ID NO 58)

GGTCCGATAA AGAAGGCGGC AGTGCACTAC GATCGCTCCG GTCGCTCGTT
GGGCACCGCT GACGTGATTT TCGAACGTCG CGCCGACGCC TTGAAGGCCA
TTAAACAGTA CCATGGCGTA CCTTTGGACG GACGCCCTAT GACCATTCAG
CTGGCCGTCT CAGACGTGGC CGTGTTGACC CGTCCCGTAG CCGCCACCGA
TGTCAAGCGT CGCGTGGGTG GTACTGCACC AACTTCATTC AAGCGTGGTG
GTGGCCAAGC TGGTGGCACG GCGCGTCGCG GCTTCAAACG TCCGGTCGGT
GGCAAGCCGG CGGCAGGCGG CCAGCGACGG GAGCGCAAGG CCCCGCCCAC
TGCTGAGGAG CTGGACGCCG AACTGGACTC A

v104471 (KK): ID base de mosca = FBst0476329 - - la línea lleva la siguiente secuencia de repetición invertida (SEQ ID NO 59)

GTCGAACTTG ATAAAGCGCA TTTCTAAATA CAATAAATAC AGCATCAAAT
GTATTTCAGT TATCTTAACA TCCGCCGCAT TGGCAAAACT AACAATTAAT
GGATAAATGC GCAAGTGGTT GATTGATTTG ATGTCCGATG CTTTCAAAGA
TCTGCTCCTG GGCGCGGCGT TGTCGATGCG TTTGCATTTA TGTACCATGC
GGGGGGTGTC CATATGGTAG GCTTAAAACT ATAGATTGGG CTGCTCTTCT
ATTCTTGTTA GACTAATTCA GACTATTCAC TATTTAGATC TTCATGTCGT
TGATGTATGA GTCCAGTTCG GCGT

Vectores de AAV

Los vectores AAV adecuados son los vectores scAAV y AAV autocomplementarios de serotipos variados (por ejemplo, los serotipos 2, 5, 6, 9 y derivados de AAV9 como AAV-PHP.A y AAV-PHP. Otras alternativas son AAV9: (Valori y otros. Sci Transl Med. 2010; 2: 35ra42) y AAV-PHP-A y B) Deverman y otros. Nature Biotech 2016: 34: 204-9) (ver también la Figura 19)

Elementos promotores para impulsar la expresión de los antagonistas de SRSF1:

Los casetes SRSF1-ARNi se activarían mediante el uso de promotores H1 o U6; Expresión de proteínas o péptidos SRSF1 mediante el uso de CBA, CAG, PGK, EFS1a o CMV; En la figura 18 se enumeran más ejemplos.

Ensayos de cría y locomotores de Drosophila

Las Drosophilas se cultivaron en condiciones estándar de alimentos que consisten en agar, harina de maíz, melaza y levadura, a menos que se indique lo contrario. Todas las líneas transgénicas relacionadas con C9orf72 (14) fueron un regalo de Adrian Isaacs y Linda Partridge (University College de Londres). *GMR-GAL4* (#1104), *D42-GAL4* (#8816), *nSyb-GAL4* (# 51 635) *da-GAL4,* UAS-sbr-ARNi [P {TRiP.HM05135} attP2] y UAS-*luciferasa*-ARNi (#31 603), usados como ARNi controles, se obtuvieron del Centro de Abastecimiento de Drosophila *de Bloomington* (Bloomington, IN).

Los fenotipos oculares se analizaron mediante la inducción de la expresión transgénica mediada por*GMR-GAL4* elevada a 25 °C. Para los ensayos de rastreo de larvas, los transgenes se expresaron mediante *nSyb-GAL4* y los animales se cultivaron a 29 °C. Se recogieron de los viales larvas errantes del tercer estadio, se enjuagaron brevemente con agua destilada y se colocaron en placas de Petri con una matriz de agarosa al 1 %. Se observaron las larvas directamente durante 2 minutos y se registró el número de ondas peristálticas.

Los ensayos de escalada se realizaron como se describió previamente (32) con expresión transgénica inducida por *D42-GAL4*. Los cruces se iniciaron a 25 °C durante 3 días y se transfirieron a 29 °C durante 7 días más. Las moscas adultas se probaron 1-3 días después de la eclosión.

Para todos los experimentos solo se usaron moscas macho. Las Drosophilas se alojaron en un ciclo de luz: oscuridad 12:12.

Imágenes de microscopía óptica y microscopía electrónica de barridode Drosophila

Para la microscopía óptica de ojos de Drosophila, se colectaron pilas de imágenes en un microscopio de aumento estéreo SMZ motorizado Nikon equipado con una lente Apo 1x. Luego, se generaron imágenes de enfoque extendido mediante el uso del software Nikon Elements. La microscopía electrónica de barrido (SEM) se realizó de acuerdo con un protocolo estándar (33) y las imágenes se capturaron mediante el uso de un microscopio Philips XL-20 SEM. Todos los animales de un genotipo dado mostraron fenotipos esencialmente idénticos y se muestran imágenes representativas seleccionadas al azar.

Plásmidos

Los plásmidos SRSF1/ SF2/ASF marcados con FLAG se generaron como en la referencia (23) mediante la clonación de un fragmento de PCR de SRSF1/SF2/ASF humana de longitud completa en p3XFLAG-myc-CMV26 (*Sigma*).

Las secuencias de oligonucleótidos de miARN se diseñaron mediante el uso de la herramienta de diseño de ARNi "miR ARNi" Block-IT (*ThermoFisher*) en:https://rnaidesigner.thermofisher.com/rnaiexpress/ Se diseñaron dos horquillas de miARN contra SRSF1 humana (GenBank: NM_006924.4, ARNm) y SRSF1 de ratón (GenBank: BC046773.1, Secuencia de Referencia en NCBI del ARNm: NM_173374.4,) y SRSF1 de rata (Secuencia de Referencia en NCBI: NM_001109552.2, ARNm). La secuencia de SRSF1 dirigidas contra la horquilla 1 de miARN es idéntica en SRSF1 de humanos, ratones y ratas.

Las regiones en negrita en las secuencias a continuación representan las secuencias de miR ARNi maduras que se dirigen a las secuencias complementarias sentido en SRSF1 (cursiva):

La secuencia de *SRSF1* miR1 humana, de ratón y de rata dirigida (TTAAAGTTGATGGGCCCAGAA; SEC ID NO 6) comienza en 784 nt (Secuencia de Referencia en NCBI NM_006924.4 - región RRM2) y 673 nt respectivamente (región RRM2), 1,041 nt (Secuencia de Referencia en NCBI NM_173374.4 - región RRM2) y 699 nt (Secuencia de Referencia en NCBI: NM_001109552.2 - región RRM2):

SRSF1 humana/de ratón - miR1 - Cadena superior (SEQ ID NO 61)

5'- TGCTG**TTCTGGGCCCATCAACTTTAA**GTTTTGGCCACTGACTGACTTAAAG TTTGGGCCCAGAA -3'

SRSF1 humana/de ratón/de rata - miR1 - Cadena inferior: (SEQ ID NO 62)

5'- CCTGTTCTGGGCCCAAACTTTAAGTCAGTCAGTGGCCAAAAC*TTAAAGTTG ATGGGCCCAGAA*C -3'

La secuencia de SRSF1 miR2 humana dirigida (AATGGTATGACTCCAAGTGCT) (SEQ ID NO 10) comienza en 1436 nt (región 3'UTR):

SRSF1 humana - miR2 - Cadena superior: (SEQ ID NO 63)

5'- TGCTG**AGCACTTGGAGTCATACCATT**GTTTTGGCCACTGACTGACAATGGT ATCTCCAAGTGCT -3'

SRSF1 humana - miR2 - Cadena inferior: (SEQ ID NO 64)

5'- CCTGAGCACTTGGAGATACCATTGTCAGTCAGTGGCCAAAAC*AATGGTAT GACTCCAAGTGCT*C -3'

La secuencia de SRSF1 miR2 de ratón dirigida (AATGTCTATTCTGCTCTGGTT) (SEQ ID NO 65) comienza en 1,105 nt (región 3'UTR):

La secuencia de SRSF1 miR2 de ratón dirigida (AATGTCTATTCTGCTCTGGTT) (SEQ ID NO 66) comienza en 1,473 nt (Secuencia de Referencia en NCBI NM_173374.4 - región 3'UTR):

SRSF1 de ratón - miR2 - Cadena superior: (SEQ ID NO 67)

5'- TGCTG**AACCAGAGCAGAATAGACATT**GTTTTGGCCACTGACTGACAATGT
CTACTGCTCTGGTT -3'

SRSF1 de ratón - miR2 - Cadena inferior: (SEQ ID NO 68)

5'- CCTGAACCAGAGCAGTAGACATTGTCAGTCAGTGGCCAAAAC*AATGTCTA*
TTCTGCTCTGGTTC -3'

Los oligonucleótidos sintetizados *(Sigma)* se hibridaron y ligaron en pcDNA6.2 GW/EmGFP mediante el uso del kit de vector de expresión de miR ARNi de Pol II BLOCK-iT con EmGFP (*ThermoFisher,* número de catálogo K4936-00). Además, los casetes de pre-miR2 ARNi se enhebraron mediante el subclonaje de fragmentos cortados con BamHI/XhoI en los sitios BglII y XhoI de pcDNA6.2 GW/EmGFP-*SRSF1* miR1. Los fragmentos de PCR que abarcan EmGFP y el casete de pre-miARN de SRSF1 enhebrado se clonaron adicionalmente en el plásmido de expresión lentiviral SIN-PGK-cPPT-GDNF-WHV (9) (5) mediante el uso de las enzimas de restricción SpeI y XhoI. La restricción de SIN-PGK-cPPT-GDNF-WHV (9) por SpeI/XhoI permite la eliminación del inserto GDNF y la clonación del casete de ARNi de EmGFP- *SRSF1* humana o de ratón. (figuras 1 y 5).

Se construyeron repeticiones de C9ORF72 de hexanucleótidos GGGGCCx38 sentido (SEQ ID NO 69) y CCCCGGx39 antisentido (SEQ ID NO 70) ininterrumpidas mediante el uso de los oligonucleótidos sintéticos 5'-$(GGGGCC)_{15}$-3' y 5'-CCCC-$(GGCCCC)_{14}$-GG-3' (SEQ ID NO 71). Los oligonucleótidos se hibridaron mediante calentamiento a 99 °C durante 30 minutos y enfriamiento a 0,5 °C/min a temperatura ambiente con incubación a 70 °C durante 10 minutos. Los oligonucleótidos se fosforilaron con T4-polinucleótido quinasa (*New England Biolabs*), ligados mediante el uso de T4 DNA Ligasa (*ThermoFisher*) y tratados con la nucleasa de frijol mungo (*New England Biolabs*) para la ligadura roma. Las formas oligoméricas de oligonucleótidos hibridados se confirmaron tras el análisis en gel de agarosa (figura 4). A continuación, se cortó la banda correspondiente a los oligonucleótidos triméricos, se purificó en gel y se ligó en pcDNA3.1 *(Invitrogen)* con extremos EcoRI romos para permitir la clonación tanto en orientación sentido como antisentido. Las secuencias se presentan en la figura 4.

Las secuencias sintéticas que codifican para poli-Gly-Pro y poli-Gly-Ala x36 DPR independientemente de las repeticiones de G4C2 se clonaron primero en pcDNA3.1 *(Invitrogen)* mediante el uso de los sitios *EcoRI* y *NotI*. Las secuencias sintéticas que codifican para poli-Gly-Pro y poli-Gly-Ala x36 se subclonaron luego mediante el uso de BamHI/NotI en pCI-neo-V5-N mediante el uso de BclI/NotI. El sitio de restricción BclI se introdujo previamente en pCI-neo-V5-N mediante mutagénesis sitio dirigida mediante el uso de cebadores ¨sentido¨ actctagaggtaccacgtgatcattctcgagggtgctatccaggc (SEC ID NO 72) y antisentido gcctggatagcaccctcgagagaatgatcacgtggtacctctagagt (SEC ID No: 73) (QuikChange Lightning Site-Directed Mutagenesis Kit, *Agilent*).

Producción de lentivirus

La construcción de miARN se subclonó en un vector lentiviral autoactivante (SIN-W-PGK) mediante el uso de métodos de clonación estándar. Los lentivirus se propagaron en células HEK293T mediante el uso de la transfección con fosfato de calcio (34). Se transfectaron 13 µg de pCMVΔR8.92, 3,75 µg de pM2G, 3 µg de pRSV y 13 µg de SIN-CMV-miARN en células HEK293T. Cada una de las veinte placas de 10 cm sembradas con $3x10^6$ células HEK293T/placa se transfectó con 13 µg de pCMVΔR8.92, 3,75 µg de pM2G, 3 µg de pRSV y 13 µg de SIN-CMV-miARN mediante el uso de la transfección con fosfato de calcio[41]. Se permitió que las células produjeran los virus durante 72 horas, luego se colectó el sobrenadante, se filtró mediante el uso de un filtro de 0,45 µm y se centrifugó a 24 000 rpm durante 90 minutos a 4 ºC. Se descartó el sobrenadante y se resuspendió el sedimento viral en BSA al 1 % en PBS y se almacenó a -80 ºC. El título biológico del vector viral se determinó mediante la transducción de células HeLa con diluciones 10^{-2}, 10^{-3} y 10^{-4} del vector. 72 horas después de la transducción, las células se fijaron en paraformaldehído al 4 %, se lavaron en PBS y se midió el porcentaje de células positivas para GFP con un clasificador de células activado por fluorescencia (FACS, LSRII). El título biológico se expresa como el número de unidades transductoras por mL (TU/ml) y se puede calcular con la siguiente fórmula: Título del vector = [(% de células positivas x número de células durante la transducción) x factor de dilución x 2] TU/mL.

Cultivo de tejidos y transfección o transducción

Las células HEK293T se cultivaron en medio Eagle modificado de Dulbecco *(Sigma)* suplementado con suero bovino fetal al 10 % [FBS] (*de Gibco*) y 5U/mL de Penstrep de (*Lonza*). Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %. Para el análisis de qRT-PCR, se dividieron 50 000 células HEK en cada pocillo de placas de 24 pocillos y se transfectaron durante 72 h con 700 ng pcDNA6.2-GW/EmGFP-Control o las construcciones de miR ARNi de SRSF1 humana mediante el uso de 3,5 µg de medio PEI/mL y una décima parte del volumen del medio OptiMEM (*ThermoFisher*). Para el análisis de Transferencia Western, las células HEK se transfectaron durante 72 h

con 650 ng de pcDNA6.2-GW/EmGFP-Control miARN, pCDNA6.2-GW/EmGFP-human SRSF1 miR1 + 2 ARNi o las construcciones LV-EmGFP-human SRSF1 miR1 + 2 ARNi y 50 ng de p3XFLAG/human-SRSF1 mediante el uso de 3,5 µg de medio PEI/mL y una décima parte del volumen del medio OptiMEM.

Se cultivaron células Neuro-2a (ATCC) en medio Eagle modificado de Dulbecco (Sigma) suplementado con suero bovino fetal al 10 % [FBS] (Gibco), Penstrep 5U/mL (Lonza) y piruvato de sodio 5 mM. Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %. Se dividieron 75 000 células Neuro-2a en cada pocillo de placas de 24 pocillos y se transfectaron durante 72 h con 500 ng de pcDNA6.2-GW/EmGFP-Control miR ARNi (ThermoFisher) o pcDNA6.2-GW/EmGFP-mouse SRSF1 miR1 + 2 ARNi y 200 ng pcDNA 3.1/RAN-G4C2x38-sentido o RAN-C4G2x39-antisentido mediante el uso de 3 µg de PEI/1 µg DNA y un décimo del volumen del medio OptiMEM (ThermoFisher).

Se cultivaron células madre de ratón Hb9GFP como se describió anteriormente (35) y se diferenciaron en neuronas motoras con ácido retinoico 2 µM *(Sigma)* y el agonista suavizado (SAG) 1 µM (*Millipore*) durante 5 días. Los cuerpos embrioides luego se disociaron con papaína y se clasificaron mediante el uso de FACSAria ™ III (*BD Biosciences*). Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %.

Los astrocitos diferenciados derivados de pacientes humanos (iAstrocitos) se diferenciaron de las células progenitoras neuronales inducidas (iNPC) como se describió anteriormente (29) y se cultivaron en DMEM Glutamax (*Gibco*) con FBS (*al 10 % Sigma*) y N2 al 0,02 % (*Invitrogen*). Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %.

Para los cocultivos de neuronas motoras Hb9GFP + e iAstrocytres derivados de pacientes, se sembraron 20 000 células progenitoras neurales inducidas (iNPC) en placas de 6 pocillos en medio de astrocitos. El día después de la siembra en placa, los iAstrocitos fueron transducidos con adenovirus que expresan la proteína roja fluorescente (RFP) ya que los experimentos de cocultivo se realizaron mediante el uso de neuronas motoras GFP+ de células madre de ratón Hb9GFP+ y con lentivirus que coexpresan el ARNi de *SRSF1* humana y GFP a una MOI de 5, 7 o 10. La expresión de GFP en iAstrocitos se usó para monitorear la eficiencia de la transducción del ARNi de *SRSF1*. Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %.

Para los cultivos celulares derivados de pacientes, se obtuvo el consentimiento informado de todos los sujetos antes de la recolección de la muestra (Número de estudio STH16573, referencia del Comité de Ética en Investigación 12/YH/0330). Los astrocitos diferenciados derivados de pacientes humanos (iAstrocitos) se diferenciaron de las células progenitoras neurales inducidas (iNPC) como se describió anteriormente 52 y se cultivaron en DMEM Glutamax (Gibco) con FBS al 10 % (Sigma) y N2 al 0,02 % (Invitrogen). Las células se mantuvieron en una incubadora a 37 °C con CO2 al 5 %.

Las neuronas derivadas de los controles y de pacientes humanos (iNeurons) se diferenciaron de las iNPC establecidas previamente. Las iNPC se diferenciaron luego en neuronas mediante el uso de una versión modificada del protocolo descrito en la referencia[40]. En resumen, se sembraron 30 000 iNPC en una placa de 6 pocillos recubierta con fibronectina (Millipore) y se expandieron hasta una confluencia del 70-80 %. Una vez que alcanzaron esta confluencia, el medio iNPC se reemplazó con medio de diferenciación de neuronas (DMEM/F-12 con Glutamax suplementado con N2 al 1 %, B27 al 2 % (Gibco)). El primer día de diferenciación, las células se trataron con 2,5 µM de DAPT (Tocris) para promover la diferenciación hacia el linaje neuronal. El día tres, el medio de diferenciación neuronal se suplementa con ácido retinoico 1 µM (Sigma), agonista suavizado (SAG) 0,5 µM (Millipore) y forskolina 2,5 µM (Sigma) durante 7 días. Este protocolo conduce a rendimientos típicos de células positivas a µ-III tubulina (Tuj1) al 70%.

Para obtener Neuronas iMotoras, las iNeuronas se volvieron a sembrar en fibronectina y se cultivaron en ácido retinoico, SAG y forskolina durante 14 días más con la adición de BDNF, CNTF y GDNF (todos a 20 ng/mL) durante los últimos 10 días de diferenciación. Para la eliminación de SRSF1, las células se transdujeron con lentivirus que expresan GFP control o ARNi de SRSF1 humana que coexpresan GFP a una MOI de 5 en el día 14 junto con el adenovirus HB9:RFP.

Cocultivos de iAstrocitos derivados de pacientes humanos y neuronas motoras humanas o de ratón

Para los cocultivos de neuronas motoras Hb9GFP y de iAstrocitos derivados de pacientes, se sembraron en placas 20 000 células progenitoras neurales inducidas (iNPC) en placas de 6 pocillos en medio de astrocitos. El día después de la siembra en placa, las células se transdujeron con lentivirus que expresaban GFO control o ARNi de *SRSF1* humana con una MOI de 5, 7 o 10. El virus de ARNi de *SRSF1* humana también coexpresó GFP para permitir la evaluación de la eficiencia de la transducción. Dado que los experimentos de cocultivo se realizaron mediante el uso de neuronas motoras GFP+ de células madre de ratón Hb9GFP+, el mismo día las células también se transdujeron con un adenovirus que expresa la proteína roja fluorescente (RFP). Las células se mantuvieron en una incubadora a 37 °C con CO_2 al 5 %.

Siete días después de la transducción con Ad-RFP y LV-*SRSF1*-ARNi, los iAstrocitos se sembraron a una densidad de 10 000 células/pocillo. Al día siguiente, los cuerpos embrioides de Hb9GFP se disociaron y clasificaron para las células GFP+. Se colocaron en placa 10 000 neuronas motoras GFP+ sobre los astrocitos en medio de neuronas

motoras que constaba de DMEM/F12, suero de caballo al 2 % *(Invitrogen),* N2 al 2 %, B27 al 2 % más GDNF *(Invitrogen;* 10 ng/mL), BDNF *(Invitrogen;* 10 ng/mL), CNTF *(Invitrogen;* 10 ng/mL) e IGF-1 *(Invitrogen;* 10 ng/mL).

Se adquirieron diariamente nueve imágenes 10x/pocillo para cubrir toda la superficie del pozo durante 3 días mediante el uso del sistema de imágenes de alto contenido InCell 2000 (*GE Healthcare*), y se reunieron datos sobre el tamaño y número de células neuronales, la longitud axonal y la ramificación de neuritas. El análisis de datos se realizó mediante el uso del software Desarrollador *En Célula*. El análisis de datos se realizó mediante el uso del software *Columbus* (*PerkinElmer*). Los datos se presentan para 3 días de cocultivo. El programa diseñado para el análisis de cocultivos solo toma en cuenta las células GFP+ con al menos una proyección para excluir los conteos de restos celulares. Para los cultivos de iMN en iAstrocitos, los iAstrocitos se sembraron en placas de 384 pocillos 24 h antes de sembrar 1000 iMN clasificados con FACS. Los cultivos se mantuvieron durante 4 días. Los datos se presentan para 4 días de cocultivo.

Fraccionamiento citoplasmático

El fraccionamiento citoplasmático se realizó 72 h después de la transfección. Se retiraron de la placa 300 000 células de un pocillo de una placa de 6 pocillos mediante el uso de DEPC PBS y se sedimentaron mediante centrifugación a 800xg durante 5 minutos. Los sedimentos celulares se lavaron rápidamente con tampón de lisis hipotónico (HEPES 10 mM pH 7,9, MgCl 1,5 mM_2, KCl 10 mM, DTT 0,5 mM). A continuación, se lisaron los sedimentos celulares en tampón de lisis hipotónico que contenía 0,16 U/µl de inhibidores de ribosafe RNasa (Bioline), PMSF 2 mM (Sigma) y tabletas de un cóctel de inhibidores de proteasa SIGMAFAST™, sin EDTA (Sigma) de acuerdo con las instrucciones del fabricante. Las células se lisaron suavemente mediante el uso de una punta P1000 cortada y se aseguró de no ejercer fuerza física sobre el sedimento celular. El lisado luego se sometió a centrifugación diferencial (1500 g, 3 min, 4 °C, luego 3500 g rpm, 8 min, 4 °C y luego 17 000 g, 1 min, 4 °C) y se transfirió el sobrenadante a un tubo nuevo después de cada centrifugación. A continuación, se añadió el sobrenadante a PureZOL (Biorad) para extraer el ARN. El sedimento nuclear obtenido después de la centrifugación a 1500 g durante 3 min se lisó en tampón de lisis Reporter *(Promega)* durante 10 minutos en hielo antes de centrifugar a 17 000 g, 5 min, 4 °C.

Las fracciones totales se recogieron directamente en el tampón de lisis Reporter (*Promega*) que contiene 0,16 U/µL de inhibidores de ribosafe RNasa (*Bioline*), PMSF 2 mM (*Sigma*) y tabletas del cóctel de inhibidores de proteasa SIGMAFAST™, sin EDTA (Sigma) de acuerdo con las instrucciones del fabricante y lisadas durante 10 min en hielo antes de la centrifugación a 17 000 g, 5 min, 4 °C. A continuación, se añadió el sobrenadante a PureZOL (Biorad) para extraer el ARN.

Se resolvieron volúmenes iguales de lisados totales, nucleares y citoplasmáticos mediante el uso de SDS-PAGE, se sometieron a electrotransferencia en membranas de nitrocelulosa y se sondaron mediante el uso de anticuerpos anti-SSRP1 y anti-α-tubulina (ver la sección de análisis por Transferencia Western a continuación para obtener detalles sobre los anticuerpos).

Análisis por Transferencia Western

Las células HEK se transfectaron durante 72 h con 650 ng de pcDNA6.2-GW/EmGFP-Control miARN, pCDNA6.2-GW/EmGFP-human *SRSF1* miR1 + 2 ARNi o construcciones de ARNi LV-EmGFP-human SRSF1 miR1 + 2 y 50 ng de p3XFLAG/SRSF1 humana mediante el uso de 3,5 µg de medio PEI/mL y una décima parte del volumen del medio OptiMEM. Las células Neuro-2a se dividieron en cada pocillo de placas de 24 pocillos (75 000 células/pocillo) y se transfectaron durante 72 h con 350 ng de pcDNA6.2-GW/EmGFP-Control miR ARNi (ThermoFisher), pcDNA6.2-GW/EmGFP-ratón SRSF1 miR1 + 2 ARNi, p3xFLAG, p3xFLAG/SRSF1 (11-196), p3xFLAGSRSF1 (11-196)-m2 o p3xFLAG/SRSF1 (11-196)-m4 y 350 ng pcDNA 3.1/RAN-G4C2x38-sentido o RAN-C4G2x39 antisentido mediante el uso de 3 µg de PEI/1 µg de ADN y una décima parte del volumen del medio OptiMEM.

Las proteínas se extrajeron de células HEK o Neuro-2a 72 horas después de la transfección. Las células se lavaron brevemente en solución salina tamponada con fosfato (PBS) enfriada con hielo y luego se rasparon en tampón de lisis enfriado con hielo (Hepes 50 mM pH 7,5, NaCl 150 mM, glicerol al 10 %, tritón X-100 al 0,5 %, EDTA 1 mM, DTT 1 mM, cóctel de inhibidores de proteasa (*Sigma*)). Las células se dejaron lisar en hielo durante 10 minutos seguido de centrifugación a velocidad máxima a 4 ºC durante cinco minutos. Los extractos de proteínas se cuantificaron mediante el uso del reactivo de Bradford (*BioRAD*), se resolvió mediante SDS-PAGE, se sometió a electrotransferencia sobre una membrana de nitrocelulosa y se probó mediante el uso del anticuerpo primario correspondiente. SRSF1/SF2 humana/de ratón [dilución 1:1000] (*Señalización celular* # 8241) y poli-Gly-Pro [dilución 1: 10 000] (amablemente recibidos del profesor Stuart Pickering Brown) se detectaron con anticuerpos primarios con anticuerpo secundario de conejo conjugado con peroxidasa de rábano picante (HRP) [dilución 1:5000] (*Promega*), mientras que anticuerpos primarios α-tubulina [dilución 1:10 000] (*Sigma,* clon DM1A), FLAG [dilución 1:2000] (*Sigma* F1804, clon M2), ALYREF [dilución 1:2000] (*Sigma* A9979, clon 11G5), SSRP1 [dilución 1:500] (*Abcam* 26212, clon 10D7) y poli-Gly-Ala [dilución 1:500] (amablemente proporcionados por el profesor Dieter Edbauer) se detectaron mediante el uso del anticuerpo secundario de ratón conjugado con HRP [dilución 1:5000] (*Promega*). Para el análisis de transferencia puntual, se cargaron 50 µg de extractos de proteínas totales preparados en tampón de lisis helado en una membrana de

nitrocelulosa mediante el uso de un aparato de microfiltración (*Biorad*) y se analizó mediante inmunotransferencia western como se describió anteriormente.

Fraccionamiento citoplasmático

Las células Neuro-2a se dividieron en cada pocillo de placas de 6 pocillos (2 x 10^6 células/pocillo) y se transfectaron durante 72 horas con 1 µg de pcDNA6.2-GW/EmGFP-Control miR ARNi, pcDNA6.2-GW/EmGFP-mouse SRSF1 miR1 + 2 ARNi, p3xFLAG/SRSF1 (11-196), o p3xFLAG/SRSF1 (11-196)-m4 y 1 µg pcDNA 3.1/RAN-G4C2x38-sentido o RAN-C4G2x39-antisentido mediante el uso de 3 µg PEI/1 µg de ADN y un décimo del volumen del medio OptiMEM. las iNeuronas se cultivaron en placas de 6 pocillos y se transdujeron con 5MOI de lentivirus LV-*SRSF1*-ARNi humano durante 5 días.

El fraccionamiento citoplasmático se realizó como sigue. Las células se retiraron de la placa mediante el uso de DEPC PBS y se sedimentaron mediante centrifugación a 800 xg durante 5 minutos. Los sedimentos celulares se lavaron rápidamente con tampón de lisis hipotónico (HEPES 10 mM pH 7,9, MgCl 1,5 mM_2, KCl 10 mM, DTT 0,5 mM). A continuación, se lisaron los sedimentos celulares en tampón de lisis hipotónico que contenía 0,16 U/µl de inhibidores de ribosafe RNasa (Bioline), PMSF 2 mM (Sigma) y tabletas de un cóctel de inhibidores de proteasa SIGMAFAST™, sin EDTA (Sigma) de acuerdo con las instrucciones del fabricante. Las células se lisaron suavemente mediante el uso de una punta P1000 cortada para asegurar que no se ejerciera ninguna fuerza física sobre el sedimento celular. Para el tejido de mosca, *da-GAL4* se usó para dirigir la expresión del transgén en todos los tejidos, y se homogeneizaron y lisaron 10 larvas de tercer estadio mediante el uso del mismo tampón y un homogeneizador Dounce. El lisado luego se sometió a centrifugación diferencial (1500 g, 3 min, 4 °C, luego 3500 g rpm, 8 min, 4 °C y luego 17 000 g, 1 min, 4 °C) y se transfirió el sobrenadante a un tubo nuevo después de cada centrifugación. A continuación, se añadió el sobrenadante a PureZOL (*Biorad*) para extraer el ARN. El sedimento nuclear obtenido después de la centrifugación a 1500 g durante 3 min se lisó en tampón de lisis Reporter (*Promega*) durante 10 minutos en hielo antes de la centrifugación a 17 000 g, 5 minutos, 4 °C.

Las fracciones totales se recogieron directamente en tampón de lisis Reporter (*Promega*) que contiene 0,16 U/µL de inhibidores de ribosafe RNasa (*Bioline*), PMSF 2 mM (*Sigma*) y tabletas del cóctel de inhibidores de proteasa SIGMAFAST ™, sin EDTA (Sigma) de acuerdo con las instrucciones del fabricante y lisadas durante 10 min en hielo antes de centrifugar a 17 000 g, 5 minutos, 4 °C. A continuación, se añadió el sobrenadante a PureZOL (Biorad) para extraer el ARN.

Se resolvieron volúmenes iguales de lisados totales, nucleares y citoplasmáticos mediante el uso de SDS-PAGE, se sometieron a electrotransferencia en membranas de nitrocelulosa o PVDF y se sondaron mediante el uso de SSRP1 y α-tubulina (Neuro-2a), SSRP1 y actina (iNeurons), o α-tubulina y α- Histona H3 (*Drosophila*).

RT-PCR Cuantitativa

El ARN total se extrajo de 5-10 larvas o adultos de *Drosophila* triturados mediante el uso de 800 µL de PureZOL (*BioRAD*) y una jeringa con una aguja de calibre 21G para la homogeneización. El lisado se aclaró mediante centrifugación durante 10 minutos a 12 000 g a 4 ºC. Se añadieron 200 µL de cloroformo al sobrenadante y se agitó vigorosamente durante 15 segundos. Después de 10 minutos de incubación a temperatura ambiente, los tubos se centrifugaron a 12 000 g durante 10 minutos a 4 ºC y se colectaron los sobrenadantes (400 µL). El ARN se precipitó durante 30 minutos a temperatura ambiente con 2 µl de glucógeno (5 mg/mL, *Ambion*) y 500 µL de isopropanol y se sedimentaron a 12 000 g durante 20 minutos a 4 °C. Los sedimentos se lavaron con DEPC EtOH al 70 % y se resuspendieron en 40 µL de agua DEPC. Luego se trató el ARN total con DNasel (*Roche*) y se cuantificó mediante el uso de un Nanodrop (NanoDrop®Technologies).

Para las células HEK, se dividieron 50 000 células en cada pocillo de placas de 24 pocillos y se transfectaron con 700 ng de pcDNA6.2-GW/EmGFP-Control o construcciones de miR ARNi de SRSF1 humana mediante el uso de 3,5 µg de medio PEI/mL y un décimo del volumen del medio OptiMEM (*ThermoFisher*). Para los iAstrocitos, se sembraron 20 000 células progenitoras neurales inducidas (iNPC) en placas de 6 pocillos en medio de astrocitos. El día después de la siembra en placa, se transdujeron 3 pocillos con lentivirus que expresaban ARNi de*SRSF1* humana con una MOI de 5.

El ARN total se extrajo de las células HEK 72 horas después de la transfección o los iAstrocitos 5 días después de la transducción y el ARN se extrajo mediante el uso del Kit de Aislamiento de ARN Total EZ (*Geneflow*). Brevemente, las células se lavaron en PBS tratado con DEPC antes de la lisis directamente en la placa de cultivo a temperatura ambiente mediante el uso de la solución desnaturalizante. Las células lisadas se rasparon y sacaron de la placa de cultivo y se añadió tampón de extracción de igual volumen, se agitaron vigorosamente, se incubaron a temperatura ambiente durante 10 minutos y luego se centrifugaron durante 15 minutos a 4 ºC y 12 000 g. Posteriormente, se precipitó el ARN de la fase superior mediante el uso de volúmenes iguales de isopropanol durante la noche a -20 ºC, se sedimentó a 12 000 g, 4 ºC durante 15 minutos, se lavó con DEPC EtOH al 70% y se resuspendió en 22,5 µL de agua DEPC. Luego, el ARN se trató con DNasel (*Roche*) y se cuantificó mediante el uso de un Nanodrop (NanoDrop®Technologies).

Después de la cuantificación del ARN, 2 µg de ARN se convirtieron a ADNc mediante el uso de transcriptasa inversa BioScript™ (*Bioline*). Los cebadores de qRT-PCR se diseñaron mediante el uso de Primer-BLAST (*NCBI*) y se validaron mediante el uso de series de diluciones del molde seriadas 1 en 4 (curva estándar con R^2> 0,97). Las reacciones de qRT-PCR se realizaron por duplicado para 3 repeticiones biológicas independientes mediante el uso de mezcla maestra de QPCR Brilliant III Ultra-Fast SYBR® Green (*Agilent Technologies*) en un sistema de QPCR MX3000P (*Stratagene*). Los datos del qRT-PCR se analizaron mediante el uso de MxPro (*Stratagene*) y *GraphPad Prism* (Versión 6). Se usaron los siguientes cebadores para qPCR:

SF2 de Drosophila (diseñado mediante el uso de Primer-BLAST) (SEQ ID NO 74/75)

Dir: 5'-TACCGCGTCATGGTTACTGG-3'
Inv: 5'-GTACGCGAATGTAGGCAACC-3'

*Ref1*de Drosophila (diseñado mediante el uso de Primer-BLAST) (SEQ ID NO 76/77)

Dir: 5'- CGATATGTACGACGGACCGAA-3'
Inv: 5'- CGGACCAAAGTCGTTGAAGAG-3'

*Tub84b*de Drosophila (descrito en la referencia (36)) (SEQ ID NO 78/79)

Dir: 5'-TGGGCCCGTCTGGACCACAA-3'
Inv: 5'-TCGCCGTCACCGGAGTCCAT-3'

3'UTRdel C9 de *Drosophila* (descrito en la referencia[69]) (SEQ ID NO 80/81)

Dir: 5'-.TTCCAACCTATGGAACTGATGA-3'
Inv: 5'-.GGTTTTCCTCATTAAAGGCATTC-3'

SRSF1 humano (diseñado mediante el uso de Primer-BLAST) (SEQ ID NO 82/83)

Dir: 5'-CCGCATCTACGTGGGTAACT-3'
Inv: 5'- TCGAACTCAACGAAGGCGAA-3'

ALYREF/*THOC4*Humano (diseñado mediante el uso de Primer-BLAST) (SEQ ID NO 84/85)

Dir: 5'- TCTGGTCGCAGCTTAGGAAC-3'
Inv: 5'- CCACCTCTGTTTACGCTCTGT-3'

ARNsn de U1 humano (diseñado mediante el uso de Primer-BLAST) (SEQ ID NO 86/87)

Dir: 5'-CCATGATCACGAAGGTGGTT-3'
Inv: 5'-ATGCAGTCGAGTTTCCCACA-3'

SMN humano (descrito en la referencia[70]) (SEQ ID NO 88/89)

Dir 5'-CTTGTGAAACAAAATGCTTTTTAACATCCAT-3'
Inv 5'-GAATGTGAGCACCTTCCTTCTTTTT-3'

JUN humano (diseñado mediante el uso de Primer BLAST) (SEQ ID NO 90/91)

Dir 5'-GAACTGCACABCCAGAACAC-3'
Inv 5'TGGGTTGAAGTTGCTGAGG-3'

C9RAN (diseñado mediante el uso de Primer-BLAST). Los cebadores hibridan aguas abajo de las secuencias repetidas G4C2 o C4G2 en el 3'UTR del ARNm transcrito a partir de las construcciones pcDNA3.1. (SEQ ID NO 92/93)

Dir 5'-GGGCCCTTCGAACCCCCGTC-3'
Inv: 5'GGGAGGGGCAAACAACAGAT-3'

C9ORF72 Humano Exon-1 sentido (diseñado mediante el uso de Primer BLAST) (SEQ ID NO 94)
5'- TCAAACAGCGACAAGTTCCG-3'
*C9ORF72*Humano Exon-3 Reverse (diseñado mediante el uso de Primer BLAST) (SEQ ID NO 95)
5'-GTCGACATGACTGCATTCCA-3'
*C9ORF72*Humano Intron-1 Reverse (diseñado mediante el uso de Primer BLAST) (SEQ ID NO 96)
5'-GGAGAGAGGGTGGGAAAAAC-3'

Ensayo de proliferación celular MTT

Se dividieron células Neuro-2a (N2A) en cada pocillo de una placa de 24 pocillos (30 000 células/pocillo). Cada placa contenía 4 pocillos con solo medio para servir como un blanco y 4 pocillos/tratamiento. Las células se transfectaron durante 72 horas con 500 ng de pcDNA 3.1, pcDNA 3.1/RAN-G4C2x15 RANG4C2x38-sentido, RAN-C4G2x15 o RAN-C4G2x39-antisentido; o 250 ng pcDNA6.2-GW/EmGFP-Control miR ARNi, pcDNA6.2-GW/EmGFP-mouse SRSF1 miR1 + 2 ARNi, p3xFLAG/SRSF1 (11-196), o p3xFLAG/SRSF1 (11-196)-m4 y 250 ng de pcDNA 3.1, pcDNA 3.1/RAN-G4C2x38-sentido o RAN-C4G2x39-antisentido.

Se añadieron 250 mg del reactivo bromuro de tetrazolio azul de tiazolilo (MTT) a cada pocillo y se incubaron en la oscuridad a 37 ºC durante 1 hora. Posteriormente, las células se lisaron con tampón de lisis MTT de igual volumen (SDS al 20%, dimetilformamida al 50% (DMF)) y se incubaron, mediante agitación, a temperatura ambiente durante 1 hora. A continuación, se evaluó la absorbancia a 595 con un PHERAstar FS (BMG Labtech). Los experimentos se realizaron por triplicado para cada tratamiento. Los datos de absorbancia se analizaron mediante el uso de PHERAstar MARS (BMG Labtech) y *GraphPad Prism* (versión 6).

Inmunofluorescencia de neuronas corticales de rata

Las neuronas corticales se aislaron, cultivaron y transfectaron como se describió anteriormente. [36]. Brevemente, las neuronas se transfectaron mediante el uso de Lipofectamina LTX con el reactivo PLUS de acuerdo con las instrucciones del fabricante (Thermofisher; relación ADN:PLUS:LTX de 1:0,5:0,5 con 2 µg de ADN/100 000 células/cm^2). Después de 6 horas, la mezcla de transfección se reemplazó con medio condicionado.

La tinción por inmunofluorescencia de las neuronas corticales de rata se realizó 72 horas después de la transfección como se describió anteriormente[36]. Brevemente, las células cultivadas en cubreobjetos de vidrio se fijaron con formaldehído al 3,7 % en PBS a temperatura ambiente durante 20 minutos. Las células se lavaron una vez con PBS y el formaldehído residual se inactivó con NH 50 mM_4Cl en PBS a temperatura ambiente durante 10 minutos. Posteriormente, las células se lavaron en PBS y se permeabilizaron con Triton X-100 al 0,2 % en PBS a temperatura ambiente durante 5 minutos. Las células se lavaron en PBS para eliminar el exceso de Triton X-100, se bloquearon con suero de cabra al 4 % en PBS durante 2 horas a temperatura ambiente y luego se incubaron con el anticuerpo V5 [dilución 1:1000] (*ThermoFisher Scientific* # R96025) en PBS que contiene suero de cabra al 4 % durante la noche a 4 °C. Las células se lavaron 3 veces con PBS que contenía suero de cabra al 4 % y se incubaron durante 1 hora con PBS que contenía suero de cabra al 4 % y anticuerpo secundario antirratón de cabra, Alexa Fluor 594 [dilución 1:1000] (*ThermoFisher Scientific*). Las células transfectadas con pcDNA6.2-GW/EmGFP-Control o las construcciones de miR ARNi de SRSF1 humana se tiñeron posteriormente con Hoechst 33342 durante 10 minutos a temperatura ambiente, se lavaron 3 veces con PBS y se montaron en medio de montaje de fluorescencia (Dako). Después de la incubación en el anticuerpo secundario, las células transfectadas con p3xFLAG/SRSF1 (11-196) o p3xFLAG/SRSF1-m4 se lavaron 3 veces con PBS que contenía suero de cabra al 4 %, se incubaron a temperatura ambiente durante una hora con PBS que contenía suero de cabra al 4 % y anticuerpo anti-FLAG®M2-FITC [10 µg/mL] (*Sigma-Aldrich* # F4049) y posteriormente se tiñeron con Hoechst 33342. Luego, las células se lavaron 3 veces con PBS y se montaron en medio de montaje de fluorescencia (*Dako*).

Co-inmunoprecipitación

Las células se dividieron en placas de 1 x 10 cm / tratamiento (1,5 x 10^6 células/placa) y se transfectaron con 15 µg de p3xFLAG, p3xFLAG/SRSF1 (11-196), p3xFLAG/SRSF1 (11-196)-m2 o p3xFLAG/SRSF1 (11-196)-m4 mediante el uso de 3 µg de PEI/1 µg de ADN y una décima parte del volumen del medio OptiMEM.

Las proteínas se extrajeron de las células Neuro-2a 48 horas después de la transfección. Las células se lavaron brevemente en PBS enfriado con hielo, se rasparon en 500 µL de tampón de lisis enfriado en hielo, se pasaron a través de una aguja de calibre 21G 10 veces y se dejaron lisar en hielo durante 10 minutos. Las células lisadas se aclararon mediante centrifugación a velocidad máxima a 4 °C durante cinco minutos y los extractos de proteínas se cuantificaron mediante el uso del reactivo de Bradford. Se incubaron 2 mg de proteínas totales en 1 mL de tampón de lisis con 30 µL de la suspensión de resina con afinidad por anti-FLAG®M2 *(Sigma* A2220) (que se había bloqueado durante la noche con BSA al 1 % en tampón de lisis IP) durante 2 horas a 4 °C en una rueda giratoria. Los complejos proteicos capturados con la resina de afinidad por anti-FLAG®M2 se lavaron 5 veces con tampón de lisis helado y se eluyeron en 50 µL de tampón de lisis IP suplementado con 100 µg/mL del péptido 3xFLAG *(Sigma* # F4799) durante 30 minutos a 4 °C en una rueda giratoria. Se resolvieron mediante SDS-PAGE 30 µg de proteína total y 15 µl de complejos de proteína, se sometieron a electrotransferencia en una membrana de nitrocelulosa y se sondaron mediante el uso de FLAG, clon 53H8 de NXF1 [1: 2000] (*Abcam* ab50609) y □-tubulina.

ARN: ensayos de reticulación UV de proteínas

Proteínas recombinantes expresadas en 1,51 de BL21 (DE3)-RP de *E. coli* (*Novagen*) los cultivos celulares se purificaron mediante cromatografía IMAC en perlas TALON/Cobalt (Clontech) en tampones que contenían NaCl 1 M para prevenir la posible copurificación del ARN de *E. coli* (tampón de lisis: TRIS-HCl 50 mM pH 8,0, NaCl 1 M, Triton

X-100 al 0,5 %; tampón de lavado: TRIS-HCl 50 mM pH 8,0, NaCl 1 M, Triton X-100 al 0,5 %, imidazol 5 mM). La elución se logró en la etapa en la que el tampón que contenía imidazol 200 mM (TRIS-HCl 50 mM pH 8,0, NaCl 500 mM, imidazol 200 mM) y L-Arg y L-Glu 50 mM para prevenir la precipitación de proteínas mientras se retiene la interacción con ARN y NXF1[38,39]. [32]Sondas radiomarcadas con P (oligonucleótidos de ARN sintéticos G4C2x5 o C4G2x5 adquiridos de *Dharmacon*) durante 10 min a temperatura ambiente y 10 min en hielo antes de la reticulación con UV o no (10 min, 1,5 J/cm^2). Las reacciones de unión se resolvieron en SDS-PAGE antes del análisis mediante tinción de Coomassie y placa de fósforo fotoestimulable.

Ensayos de inmunoprecipitación de ARN (RIP)

Las células se dividieron en 1 x matraces T-175 / tratamiento (5 x 10^6 células/placa) y transfectadas con 30 µg de p3xFLAG, p3xFLAG/SRSF1 (11-196) o p3xFLAG/SRSF1 (11-196)-m4 y 10 µg de pcDNA 3.1/RAN-G4C2x15-sentido, RAN-G4C2x38-sentido, RAN- G4C2x15-antisentido o RAN-C4G2x39-antisentido mediante el uso de 3 µg de PEI/ 1 µg de PEI y un décimo del volumen de OptiMEM.

Los complejos de proteína-ARN se extrajeron de células Neuro-2a 48 horas después de la transfección. Los complejos de proteína-ARN se reticularon mediante el uso de formaldehído al 1 % durante 10 minutos, y mediante agitación a temperatura ambiente. El formaldehído residual se inactivó con glicina 250 mM a temperatura ambiente durante 5 minutos, mediante agitación. Las células reticuladas se lavaron posteriormente en PBS tratado con DEPC enfriado con hielo y se rasparon en tampón de lisis libre de RNasa enfriado con hielo (agua tratada con DEPC que contenía Hepes 50 mM pH 7,5, NaCl 150 mM, glicerol al 10 %, tritón X-100 al 0,5 %, EDTA 1 mM, DTT 1 mM, 1 µL de inhibidor de ARNasa, cóctel de inhibidores de proteasa). Las células se pasaron a través de una aguja de calibre 21G 10 veces y se dejaron lisar en hielo durante 10 minutos, seguido de centrifugación a velocidad máxima a 4 ºC durante cinco minutos y cuantificación mediante el uso del reactivo de Bradford. Se incubaron 2,5 mg de proteína total a una concentración de 1 mg/mL con 40 µl de suspensión de resina de afinidad anti-FLAG®M2 (que se había bloqueado durante la noche con BSA al 1 % y 5 µL/mL de ADNss) durante la noche a 4 °C en un rueda giratoria. Un 15avo del extracto de proteína total se retuvo para una muestra de entrada.

Los complejos de proteína-ARN capturados con resina de afinidad anti-FLAG®M2 se lavaron 5 veces con tampón de lisis libre de ARNasa enfriado con hielo. Posteriormente, los complejos se reticularon en sentido inverso y se eluyeron de la resina en tampón desnaturalizante de extracción de ARN EZ durante 1 hora a 70 ºC, y se resuspendió la resina cada 10 minutos. Las reticulaciones de formaldehído se invirtieron mediante calentamiento las muestras durante 1 hora a 70 °C y se extrajo el ARN mediante el uso de PureZOL (para muestras totales) o el Kit de Aislamiento de ARN Total EZ (para complejos eluidos) como se describe en la sección qRT-PCR. Las muestras de ARN extraídas se resuspendieron en 25 µL de agua libre de ARNasa.

Las muestras de ARN extraídas se trataron con DNasa y 10 µL de ARN de entrada o eluido se convirtieron en ADNc como se describió previamente en la sección de RT-PCR cuantitativa. Los cebadores qRT-PCR se diseñaron mediante el uso de Primer-BLAST y se validaron mediante el uso de series de diluciones moldes seriadas 1 en 4 (curva estándar con R^2> 0,97). Las reacciones qRT-PCR se realizaron por duplicado para 3 repeticiones biológicas independientes mediante el uso de la mezcla maestra Brilliant III Ultra-Fast SYBR® Green QPCR (*Tecnologías Agilent*) en un sistema en tiempo real CFX 96 ™ (*BIO-RAD*). Los datos de qRT-PCR se analizaron mediante el uso de CFX Manager 3.1 (*BIO-RAD*) y *GraphPad Prism* (Versión 6). Se calculó la concentración relativa de ADNc para cada muestra y las muestras de entrada se multiplicaron por 15 para obtener un valor de entrada total. Las muestras de eluato se expresaron como un porcentaje de los valores de entrada totales de entrada. Los cebadores de qPCR se describieron anteriormente en la sección de qRT-PCR.

Hibridación por Fluorescencia de ARN *in situ* (FISH) y visualización de focos de ARN

Para visualizar focos de ARN sentido, Hibridación por Fluorescencia de ARN *in situ* (FISH) se realizó como se describió anteriormente (9). Brevemente, las células iAstrocytos se cultivaron en cubreobjetos de 13 mm durante 5 días después de la transducción de LV-SRSF1-ARNi, antes de ser fijadas y permeabilizadas en paraformaldehído al 4 % y Tween-20 al 0,2 %. *(Sigma)* durante 10 minutos. Las células fijadas se lavaron posteriormente en PBS tratado con DEPC y se bloquearon con solución de hibridación [formamida al 50 %, citrato de sodio salino (SSC) 2 x, 100 mg/mL de sulfato de dextrano y fosfato de sodio 50 mM pH 7,0] durante 1 ha 66 °C. A continuación, las células se incubaron con 400 ng/mL de sonda desnaturalizada (una sonda de ADN incorporada con LNA (fluorescente de 16 mer) marcada con 5' TYE-563 contra la repetición de hexanucleótido de ARN C9orf72 sentido (Exiqon, Inc.)) en solución de hibridación durante la noche a 66 °C. Después de la hibridación, las células se lavaron una vez en SSC 2x y Tween-20 al 0,1 % a temperatura ambiente durante 15 minutos, seguido de tres lavados de 15 minutos en SSC 0,1 x a 65 ºC. Los cubreobjetos se montaron mediante el uso de un medio de montaje que contenía DAPI (Vector Labs, Inc.). Todas las soluciones se prepararon con agua tratada con DEPC.

Los focos de ARN se visualizaron mediante el uso de un sistema de microscopio confocal Leica SP5 y una lente de objetivo de inmersión en aceite 63/1,4. La presencia de focos se evaluó en alta resolución (848 µm^2 por imagen, 393 x 393 píxeles) mediante el uso de pilas z de 0,9 µm en todo el volumen de la celda.

Microscopía de inmunofluorescencia confocal en tejido de médula espinal post mortem humana de pacientes con C90RF72-ELA

Este estudio fue aprobado por el Comité de Ética de Investigación de South Sheffield y se obtuvo el consentimiento informado para todas las muestras. Los tejidos del cerebro y la médula espinal se donaron al Sheffield Brain Tissue Bank para su investigación, con el consentimiento de los familiares. La hibridación ¨in situ¨ de la fluorescencia de IHC y ARN (FISH) se realizó en tejidos embebidos en parafina (FFPE) fijados con formalina de cuatro pacientes C9ORF72 + con ELA. Después de la eliminación de la cera, la recuperación del antígeno se realizó mediante microondas durante 10 min en EDTA 0,8 mM pH 9,5. Se usó una sonda de ADN incorporada con LNA (fluorescente de 16 mer) marcada con 5' TYE-563 contra la repetición de hexanucleótido de ARN sentido (*Exiqon, Inc.;* número de lote 607323) como se describe en la sección anterior. Luego, los portaobjetos se transfirieron inmediatamente a PBS/BSA al 5 % para la tinción de proteínas mediante el uso del anticuerpo anti-SRSF1 (*Señalización celular* # 8241) a una dilución de 1:200. Después de la incubación con los anticuerpos primarios durante la noche a 4 °C en PBS tratado con DEPC/BSA al 5 %, los portaobjetos se lavaron en DEPC PBS y se incubaron con los anticuerpos secundarios fluorescentes anticonejo Alexa Fluor 488 (*Abcam* ab150077). Los portaobjetos montados se visualizaron mediante microscopía confocal como se describe en la sección anterior.

Análisis estadístico de los datos

Hemos usado ANOVA (análisis de varianza) de una vía o de dos vías para evaluar estadísticamente y graficar nuestros datos mediante el uso de la versión 6 del *GraphPad Prism* (*GraphPad Software Inc.,* San Diego, CA). Todos los archivos y puntos de datos están disponibles a pedido. Los focos de ARN se contaron de forma ciega y varios investigadores llevaron a cabo el análisis. Varios investigadores también participaron en la producción de datos de qRT-PCR y transferencia Western que muestran que el agotamiento de SRSF1 o la inhibición de la interacción de SRSF1 con NXF1 conducen a la alteración de la exportación nuclear de los transcritos repetidos de C0ORF72 y DPR tanto sentido como antisentido.

En la mayoría de los experimentos se usó ANOVA (análisis de varianza) de una vía o de dos vías con la corrección de Tukey para comparaciones múltiples, con las siguientes excepciones: El análisis DPR en neuronas primarias usó la prueba exacta de Fisher; la capacidad de escalado de las moscas adultas se analizó mediante la prueba no paramétrica de Kruskal-Wallis con la corrección de Dunn para comparaciones múltiples; y el análisis de los transcritos de G4C2x36 en *Drosophila* usó la prueba t de dos colas pareada. No se usó aleatorización en los estudios con animales. Los datos se graficaron mediante el uso de *GraphPad Prism* 6. El significado se indica de la siguiente manera; NS: no significativo, *: p <0,05, **: p <0,01, ***: p <0,001, ****: p <0,0001; Todos los archivos y puntos de datos están disponibles a pedido. Los focos de ARN, las neuronas DPR positivas, los ensayos de arrastre y escalada se analizaron de forma ciega y varios investigadores llevaron a cabo el análisis. Varios investigadores también participaron en la producción de datos de qRT-PCR y transferencia Western que muestran que el agotamiento de SRSF1 o la inhibición de la interacción de SRSF1 con NXF1 conducen a la alteración de la exportación nuclear de transcritos repetidos de *C9ORF72* y la reducción de DPR sentido y antisentido.

Vectores de terapia génica de ARNi y de expresión de SRSF1

Diseño de ARNhc para la eliminación de SRSF1

El diseño inicial de ARNhc se realizó mediante el uso del diseñador de ARNi Invitrogen Block-it (https://rnaidesigner.thermofisher.com/rnaiexpress/). La secuencia de entrada fue el número de acceso al NCBI: NM_001078166.1 - correspondiente a la variante 2 del transcrito de la isoforma corta del ARNm de SRSF1, lo que garantiza que ambas isoformas de SRSF1 sean dirigidas. El ORF se seleccionó como la región para el diseño dirigido y el contenido de GC se estableció en los valores predeterminados de 35-55 %. La orientación de la hebra de la secuencia de salida fue sentido-bucle-antisentido. La secuencia de bucle seleccionada fue CGAA pero esto se puede modificar para incorporar diferentes secuencias de bucle si es necesario.

Las 10 secuencias de salida principales clasificadas por probabilidad de eliminación de la diana se muestran en la Tabla 7.

Tabla 7 - Salida de diseño de Invitrogen ARNi

SEQ ID NO	Rango	Inicio	ARNhc sentido		GC %
97	1	259	GCATCTACGTGGGTAACTTAC	ORF	47,62
98	2	244	GGAACAACGATTGCCGCATCT	ORF	52,39
99	3	689	GGAGTTTGTACGGAAAGAAGA	ORF	42,86
100	4	603	GGAAGTTGGCAGGATTTAAAG	ORF	42,86
101	5	761	GGTAGGTTATACACGTATTCT	ORF	38,1
102	6	639	GCAGGTGATGTATGTTATGCT	ORF	42,86
103	7	657	GCTGATGTTTACCGAGATGGC	ORF	52,39
104	8	604	GAAGTTGGCAGGATTTAAAGG	ORF	42,86
105	9	701	GAAAGAAGATATGACCTATGC	ORF	38,1
106	10	632	GCGTGAAGCAGGTGATGTATG	ORF	52,39

Estas 10 secuencias mejor clasificadas fueron interrogadas por su potencial de desvío de la diana mediante análisis in silico de la hebra antisentido mediante el uso de la herramienta en línea siSPOTR (https://sispotr.icts.uiowa.edu/sispotr/tools/lookup/evaluate.html). En siSPOTR, las secuencias de ARNhc tienen un potencial de puntuación de desvío de la diana (POTS) donde una puntuación de ≤ 30 se considera ideal. Una puntuación más alta sugiere una alta probabilidad de desvío de la diana. Los resultados de este análisis se muestran en la Tabla 8.

Tabla 8 - Análisis siSPOTR del ARNhc diseñado por Invitrogen

SEQ ID NO	Secuencia	Antisentido	POTS	% peor	SEQ ID NO	Semillas
107	1	GTAAGTTACCCACGTAGATGC	343,584	28,50 %	108	TAAGTTA
109	2	AGATGCGGCAATCGTTGTTCC	40,8429	87,04 %	110	GATGCGG
111	3	TCTTCTTTCCGTACAAACTCC	588,488	5,44 %	112	CTTCTTT
113	4	CTTTAAATCCTGCCAACTTCC	1794	0,11 %	114	TTTAAAT
115	5	AGAATACGTGTATAACCTACC	171,811	65,49 %	116	GAATACG
117	6	AGCATAACATACATCACCTGC	217,02	57,33 %	118	GCATAAC
119	7	GCCATCTCGGTAAACATCAGC	463,716	12,93 %	120	CCATCTC
121	8	CCTTTAAATCCTGCCAACTTC	773,609	2,09 %	122	CTTTAAA
123	9	GCATAGGTCATATCTTCTTTC	167,649	66,20 %	124	CATAGGT
125	10	CATACATCACCTGCTTCACGC	465,8	12,79 %	126	ATACATC

siSPOTR también contiene una función que permite el diseño de secuencias de ARNhc mínimas de desvío de la diana independientemente de la eficacia potencial. Mediante el uso del CDS de la isoforma SRSF1 corta como una entrada, se generaron 8 secuencias con una puntuación POTS inferior a 30 y se muestran en la Tabla 9.

Tabla 9 - Secuencias de ARNhc con bajo potencial de desvío diseñadas por siSPOTR

SEQ ID NO	rango	posición	POTS	percentile_peor	ARNhc sentido	SEQ ID NO	Semillas	porcentaje_del contenido de_GC
127	1	11 9	22	96,99 %	GCGCTATCCG CGACATCGAC CT	128	GGU CGA U	63,64 %
129	2	24	30	92,95 %	TGGCCCCGCA GGGAACAACG AT	130	UCG UUG U	63,64 %
131	3	21 9	33	91,33 %	TCGCGACGGC TATGATTACGA T	132	UCG UAAU	50 %
133	4	11 7	34	90,84 %	CGGCGCTATC CGCGACATCG AC	134	UCG AUG U	68,18 %
135	5	33	37	88,84 %	AGGGAACAAC GATTGCCGCA TC	136	AGOSTO CGG C	54,55 %
137	6	20 8	39	88,01 %	GCGGTGTATG GTCGCGACGG CT	138	GCC GUC G	68,18 %
139	7	25 0	52	81,52 %	CTGCGGGTGG AGTTTCCTCG AA	140	UCG AGG A	59,09 %
141	8	25 6	54	80,66 %	GTGGAGTTTC CTCGAAGCGG CC	142	GCC GCU U	63,64 %

Como el algoritmo siSPOTR no clasifica las secuencias basado en su capacidad para eliminar el transcrito diana, sería sensato elegir una serie de secuencias de ARNhc de ambas herramientas de diseño y validar su capacidad para silenciar el ARNm de SRSF1 *in vitro.* siSPOTR indica que varios de los ARNhc de la herramienta de diseño de Invitrogen tienen una puntuación de POTS extremadamente alta muy superior a la puntuación recomendada de ≤30, quizás excluyéndolos de una pantalla de eliminación inicial a favor de secuencias de puntuación más baja. Referencia

siSPOTR: Boudreau, R.L. y otros. 2013. a tool for designing highly specific and potent siRNAs for human and mouse. Nucleic Acids Research, 41(1), p.e9.

Ejemplo 1

Para obtener información funcional sobre la dependencia de la exportación nuclear de los transcritos repetidos de G4C2, probamos si la reducción de los niveles de expresión de los adaptadores de exportación nuclear conservados SRSF1 (26) y ALYREF (27) podría rescatar la neurodegeneración en C90RF72-ELA *de Drosophila* que exhibe neurotoxicidad mediada tanto por ARN como por DPR (14). Se cruzaron moscas que expresaban 36 repeticiones ininterrumpidas (G4C2x36 puro) con dos líneas de ARNi transgénicas independientes, cada una dirigida a SRSF1 o REFALY (28), y se logró en general una reducción del 70-80 % en los niveles de expresión de ARNm (Figura 9A). La expresión dirigida de G4C2x36 altera el ojo compuesto y se altera mínimamente por la coexpresión de un ARNi control (14) (Figura 9B). Por el contrario, la coexpresión de dos secuencias de ARNi de*SRSF1*(también llamadas SF2/ASF) previnieron completamente la neurodegeneración, mientras que dos líneas de eliminación de *REFALY* (también llamada Ref1 o ALYREF) solo mostró un modesto rescate del fenotipo neurodegenerativo. Validamos la eliminación exitosa de SRSF1 y ALYREF en las moscas correspondientes que muestran una reducción del 70-80 % en los niveles de expresión de ARNm (Figura 9b). Además, la expresión neuronal de G4C2x36 causa déficits locomotores tanto en larvas como en adultos en este modelo[44]. Consistente con los resultados anteriores, el agotamiento de *SRSF1*restauró la función locomotora mientras el ARNi de*Ref1*no proporcionó rescate del fenotipo conductual (Figuras 9, C y D).

Este efecto es específico de SRSF1 ya que el agotamiento de ALYREF no mostró ningún efecto, lo que está de acuerdo con los fenotipos de ojo rugoso. Los efectos de neurotoxicidad observados en el modelo G4C2x36 C90RF72-ELA de Drosophila se atribuyeron principalmente a la expresión de DPRs16. En consecuencia, ahora mostramos que el agotamiento de SRSF1 conduce a una reducción prominente en la producción de DPR poli-GP tanto sentido como antisentido (Figura 25a).

Ejemplo 2

Para probar la especificidad de expansión repetida de hexanucleótidos de los fenotipos rescatados de ARNi de SRSF1, usamos las moscas GR36 y PR36 previamente establecidas 16 que expresan respectivamente 36 repeticiones de poli-Gly-Arg y poli-Pro-Arg DPR mediante el uso de codones alternativos. Como se informó en el estudio original 16, las moscas GR36 tienen una alta tasa de mortalidad y solo unas pocas moscas GR36 cruzadas con Ctrl o ARNi de SRSF1 sobrevivieron hasta la edad adulta. No obstante, el agotamiento parcial de SRSF1 no mejoró significativamente los fenotipos de ojos rugosos (Figura 9e) o los déficits locomotores (Figura 9f) inducidos por la expresión independiente de G4C2 de DPR en moscas que expresan GR36 y PR36. Estos resultados indican que el agotamiento parcial de SRSF1 ejerce neuroprotección a través de efectos directos sobre la expansión de la repetición de hexanucleótidos C9ORF72 en lugar de un efecto indirecto sobre la alteración de la expresión génica o la acumulación aguas abajo de DPR.

El agotamiento parcial de SRSF1 previene la neurotoxicidad astrocítica derivada del paciente y la muerte de las neuronas motoras. Para aplicar nuestros hallazgos *in vivo* de ELA relacionada con C9ORF72 humana, buscamos reducir SRSF1 en modelos celulares derivados de pacientes. Los plásmidos de atenuación génica de *SRSF1* humana que coexpresan un reportero de GFP y un casete de pre-miARN fueron diseñados para producir lentivirus de ARNi de SRSF1 recombinante (LV-*SRSF1*-ARNi) (Figura 1). Las células HEK293T cotransfectadas con construcciones de ARNi de *SRSF1*y un plásmido de expresión de SRSF1 etiquetado con FLAG mostraron un agotamiento eficiente y específico de SRSF1 (Figura 10A). La supervivencia de las neuronas derivadas de pacientes C90RF72-ELA no se ve afectada *in vitro* (*6,* 7) pero recientemente se informó que la función microglial alterada contribuye *in vivo* a la neurodegeneración relacionada con C9ORF72 (16). La supervivencia y morfología de las neuronas derivadas de pacientes C90RF72-ELA es indistinguible de las neuronas controles 45, 46.

Por lo tanto, evaluamos la supervivencia de las neuronas motoras en cocultivos con astrocitos derivados de pacientes mediante el uso de nuestro ensayo desarrollado recientemente que recapitula la neurotoxicidad mediada por astrocitos observada en la ELA para neuronas primarias derivadas de ratones y de humanos (29).

La transducción de astrocitos humanos derivados de iNPC (iAstrocitos) mediante el uso de una multiplicidad viral de infección (MOI) de 5 condujo a una eliminación eficaz del transcrito comparable a los niveles alcanzados *in vivo* en G4C2x36 + *SRSF1*-ARNi *de Drosophila neuroprotegido* (Figura 2). Se colocaron en placa neuronas motoras GFP-Hb9+ de ratón sobre astrocitos transducidos con LV-*SRSF1*-ARNi derivados de tres controles y tres líneas de fibroblastos de pacientes C90RF72-ELA (tabla 5) y la supervivencia se cuantificó mediante el uso de un sistema de imágenes de alto rendimiento (métodos en línea, Figura 2b, métodos).

La cuantificación de los niveles de ARNm de *SRSF1* y *ALYREF* confirmaron la eliminación específica y parcial de los transcritos de SRSF1 en los iAstrocitos controles y C90RF72-ELA (Figura 10B). Los iAstrocitos controles transducidos con un adenovirus RFP (rojo) apoyan eficazmente el crecimiento de las neuronas motoras GFP-Hb9+ (verde), mientras que menos neuronas motoras sobrevivieron cuando se cultivaron conjuntamente con astrocitos derivados de fibroblastos de pacientes con C90RF72-ELA (29) (Figura 10C). La cuantificación de las neuronas motoras

supervivientes a partir de cuatro experimentos repetidos mostró que, si bien el agotamiento de SRSF1 no es perjudicial para los cocultivos controles, la muerte de las neuronas motoras se previno mediante el agotamiento de SRSF1 en cocultivos derivados de tres casos separados de C9ORF72-ELA (figura 10D).

Ejemplo 3

Para investigar la posible alteración de la exportación nuclear de los transcritos repetidas de G4C2, cuantificamos focos de ARN sentido nuclear y citoplasmático en los iAstrocitos C9ORF72-ELA. Las imágenes representativas (figura 3) y los recuentos individuales se informan en la tabla de datos adicionales 1. Tras el agotamiento de SRSF1, el número de focos de ARN citoplasmático disminuyó, mientras que los focos de ARN nuclear aumentaron concomitantemente (Figura 11E) de acuerdo con la inhibición de la exportación nuclear predicha de los transcritos repetidas de G4C2.

Debido a la detección deficiente de DPR en moscas o iAstrocitos e iNeuronas derivados del paciente (no se muestra), diseñamos plásmidos que expresan 38 transcritos de repeticiones ininterrumpidas de G4C2-sentido o 39 ininterrumpidas de C4G2-antisentido que son sustratos para la traducción por RAN específica de DPR (Figura 4). La transfección de células N2A neuronales de ratón con construcciones de repeticiones sentido o antisentido condujo a la producción de DPR específica de poli-Gly-Pro (expresado a partir de transcritos sentido y antisentido y poli-Gly-Ala (expresado a partir de transcritos sentido) (Figura 11A). Sin embargo, la cotransfección de un plásmido *SRSF1*-ARNi en ratón (Figura 5) condujo a una reducción marcada en la traducción de DPR por RAN tanto sentido como antisentido (Figura 11A, paneles de la izquierda; cuantificación en la Figura 6A). Esto no depende de la actividad de corte y empalme reducida de SRSF1 ya que las construcciones de RAN se diseñaron sin sitios de corte y empalme. Además, es específico para las secuencias de repetición de hexanucleótidos, ya que la expresión sintética de poli-Gly-Ala x36 o poli-Gly-Pro x36 mediante el uso de codones alternativos (GGA/T, GCA/T, CCA/T) no se altera con el agotamiento de SRSF1 (Figura 11A, paneles de la derecha). Como apoyo a esto, *SRSF1*-ARNi no mejoró los déficits locomotores conferidos por la expresión de DPR en *Drosophila* (Figura 7). Estos datos son consistentes con nuestros hallazgos informados anteriormente que mostraron la unión específica y directa de SRSF1 en transcritos repetidos de G4C2 (9).

Ejemplo 4

Para identificar el deterioro potencial en la exportación nuclear de los transcritos repetidas de C9ORF72, medimos los niveles de expresión citoplasmática de los transcritos de G4C2 en presencia de construcciones Ctrl o *SRSF1*-ARNi. Los niveles de transcritos repetidos citoplasmáticos normalizados a niveles totales se reducen notablemente tras la exposición a SRSF1-ARNi (figura 11B, imagen izquierda), lo que indica que la exportación nuclear de transcritos repetidos de C9ORF72 se inhibe por el agotamiento de SRSF1. La calidad del fraccionamiento celular se comprobó mediante inmunotransferencia con anticuerpos contra el factor SSRP1 remodelador de cromatina, lo que mostró la ausencia de contaminación nuclear en las fracciones citoplasmáticas (figura 8). SRSF1 media la exportación nuclear de ARNm mediante la unión a NXF1 (*26, 30*). Anteriormente mostramos que cuatro residuos de arginina que se encuentran en el enlazador no estructurado entre los dos motivos de reconocimiento de ARN de SRSF1 (aminoácidos 11-196) son necesarios tanto para la exportación nuclear de ARN como para la interacción con NXF1, mientras que las mutaciones de solo dos residuos de arginina conducen a una unión a NXF1 reducida (*23*). La cotransfección del mutante puntual cuádruple de SRSF1 (SRSF1-m4) que no interactúa con TAP/NXF1 (*23*) condujo a una marcada reducción en los niveles de transcritos repetidos citoplasmáticos normalizados (Figura 10B, imagen derecha) y la producción de DPR sentido y antisentido, mientras que la región de tipo salvaje o la misma secuencia que lleva solo dos mutaciones de arginina (SRSF1-m2) respectivamente tenía poco efecto o ninguno (Figura 11C, cuantificación en las Figuras 6 B y C) que demuestra que los transcritos repetidos de C9ORF72 se exportan desde el núcleo a través de un mecanismo que requiere la interacción de SRSF1 y NXF1. Por consiguiente, tanto el agotamiento de SRSF1 como la expresión del mutante dominante negativo SRSF1-m4 suprimen la neurotoxicidad mediada por la expresión de los transcritos repetidos de C9ORF72 en células neuronales N2A (Figura 11D).

En este informe, mostramos por primera vez que la exportación nuclear de transcritos repetidos de C9ORF72 y la posterior traducción por RAN depende de la interacción del adaptador de exportación nuclear SRSF1 con el receptor de exportación NXF1. La inhibición de esta interacción o el agotamiento de SRSF1 conduce a una inhibición marcada de la exportación nuclear de transcritos repetidos de C9ORF72 y una traducción por RAN reducida de DPR tanto sentido como antisentido. Además, el agotamiento parcial de SRSF1 previene *in vivo* la neurodegeneración y rescata los déficits locomotores en un modelo de *Drosophila* de C9ORF72-ELA, consistente con trabajos previos que demostraron que las DPR causan neurodegeneración en *Drosophila (14)*. El agotamiento de *SRSF1* también suprime la neurotoxicidad en cocultivos de neuronas motoras y astrocitos C9ORF72-ELA derivados de humano y en células neuronales N2A. Por lo tanto, esto representa una perspectiva prometedora para el desarrollo de una estrategia neuroprotectora efectiva en la ELA relacionada con C9ORF72. Es de destacar que la depleción de ARNi de *SRSF1* se usó como un enfoque terapéutico para prevenir la transformación oncogénica tanto *in vitro* e *in vivo* en ratones (31). Las vías celulares que causan la neurodegeneración después de la exportación nuclear de secuencias repetidas de C9ORF72 G4C2 y el mecanismo o mecanismos precisos de neuroprotección conferidos por el agotamiento parcial de SRSF1 quedan por dilucidar en estudios futuros.

Ejemplo 5

ALYREF y SRSF1 se unen directamente al ARN repetido de G4C2 y C4G2. Realizamos ensayos de reticulación UV in vitro mediante el uso de proteínas recombinantes purificadas y sondas de ARN sintéticas G4C2x5 y C4G2x5 para investigar las interacciones directas proteína: ARN. ALYREF humana recombinante etiquetada con hexahistidina, los aminoácidos de SRSF1 11-196 que retienen la capacidad del tipo salvaje para unirse al ARN y NXF136, y MAGOH, una proteína control que no se une al ARN49, se purificaron mediante cromatografía de afinidad de iones metálicos en alta salinidad para interrumpir las interacciones potenciales con el ARN bacteriano. Las proteínas purificadas se incubaron con sondas de ARN G4C2x5 (Figura 20a) o C4G2x5 (Figura 20b) radiomarcadas con 32P en el extremo 5' antes de la irradiación con UV donde se indica (+) y se resolvieron mediante SDS-PAGE. Como se muestra en las placas de fósforo fotoestimulable, las moléculas de ARN unidas covalentemente permanecieron asociadas con ALYREF y SRSF1 visualizadas en los paneles teñidos con Coomassie durante la electroforesis desnaturalizante. Estos datos demuestran interacciones directas de ALYREF y SRSF1 con el ARN repetido sentido y antisentido de acuerdo con nuestros estudios anteriores [42, 43]- Las interacciones son específicas ya que no se detectó unión del ARN en ausencia de irradiación UV o con la proteína control negativo MAGOH. Estas interacciones directas también son consistentes con nuestra colocalización previamente informada de ALYREF con focos de ARN en neuronas motoras de pacientes con C9ORF72-ELA[42]. Además, mostramos que SRSF1 colocaliza con focos de ARN en neuronas motoras de tejido de médula espinal humana post mortem de casos de C9ORF72-ELA (Figuras 20c).

Ejemplo 6

Generación de modelos de células neuronales que recapitulan la traducción dependiente de RAN de DPR sentido y antisentido. Para investigar si la unión de SRSF1 a secuencias de ARN repetidas de G4C2 sentido y C4G2 antisentido tiene la capacidad de desenhebrar la exportación nuclear de transcritos repetidos, generamos construcciones de expresión de mamíferos sintéticos con longitudes crecientes de secuencias repetidas puras en ausencia de elementos ATG o Kozak para investigar específicamente la traducción dependiente de RAN de proteínas de repetición de dipéptidos. Después de la hibridación de los oligonucleótidos repetidos sintéticos G4C2 o C4G2 como se describe en Métodos y en la Figura 4a-d, diseñamos plásmidos que expresan transcritos que contienen 15 o 38 repeticiones sentido ininterrumpidas (G4C2x15 o G4C2x38) y 15 o 39 repeticiones antisentido ininterrumpidas (G4C2x15 o C4G2x39) con codones de terminación en el extremo 3' en cada uno de los tres marcos (Figura 21a). Las longitudes de las repeticiones se confirmaron mediante secuenciación y electroforesis en gel de poliacrilamida. Las secuencias de nucleótidos se presentan en la Figura 4e complementaria. En los mamíferos, la exportación nuclear masiva de ARNm se acopla predominantemente al reclutamiento del complejo TREX durante el corte y empalme [47]. Tres transcritos de C9ORF72, cada una de los cuales contiene 4 o 10 intrones, se transcriben a partir del gen de C9ORF72. Las construcciones repetidas sintéticas se diseñaron sin elementos de corte y empalme o secuencias intrónicas para investigar el potencial de exportación nuclear de las secuencias de ARN repetidas de G4C2 y C4G2 en transcritos repetidos independientemente del acoplamiento funcional al corte y empalme del pre-ARNm. Las transfecciones de células N2A neuronales de ratón con construcciones repetidas sentido o antisentido condujeron a la formación de focos de ARN para todos los transcritos repetidos (Figura 21b, datos mostrados para la construcción G4C2x38) y a la producción de DPR específica de poli-Gly-Pro (expresado a partir de ambos transcritos sentido y antisentido) y poli-Gly-Ala (expresado a partir de transcritos sentido) para transcritos repetidos que llevan 38 repeticiones sentido o 39 repeticiones antisentido (Figura 21c). Curiosamente, la expresión de DPR se correlacionó con los efectos neurotóxicos en ensayos de proliferación celular MTT, mientras que el plásmido control o las construcciones que expresan 15 repeticiones sentido o antisentido pero ninguna DPR no exhibieron citotoxicidad significativa en las células neuronales (Figura 21d). Llegamos a la conclusión de que los mínimos transcritos de ARN repetidos de G4C2x38 sentido y C4G2x39 antisentido pueden exportarse al citoplasma independientemente del acoplamiento funcional al corte y empalme del pre-ARNm y son sustratos para la posterior traducción de DPR por RAN.

Ejemplo 7

El agotamiento de SRSF1 inhibe la producción de DPR sentido y antisentido en las células neuronales. Un plásmido de eliminación de SRSF1 de ratón que coexpresa un indicador de GFP y un casete de pre-miARN se diseñó de manera similar al SRSF1-ARNi humano descrito anteriormente (Figura 5). La transfección de células N2A neuronales de ratón con construcciones repetidas de G4C2x38 o C4G2x39 y el plásmido SRSF1-ARNi de ratón condujo a una reducción marcada en la traducción de DPR por RAN tanto sentido como antisentido (Figura 11a, paneles izquierdos y Figura 6b para la cuantificación). La inhibición dependiente del ARNi de SRSF1 de la producción de DPR no depende de la actividad de corte y empalme de SRSF1 ya que las construcciones de RAN están desprovistas de sitios de corte y empalme. Además, el agotamiento de SRSF1 es específico de las secuencias de repetición de hexanucleótidos, ya que la expresión de DPR sintéticas de 36 repeticiones de poli-Gly-Pro (GP36) o de 36 repeticiones de poli-Gly-Ala (GA36) mediante el uso de codones alternativos (secuencias disponibles en la Figura 27) no se altera tras el agotamiento de SRSF1 (Figura 11a, paneles de la derecha y Figura 6e-f para las cuantificaciones respectivas de GP36 y GA36). Estos resultados apoyan nuestros hallazgos previos de que el agotamiento de SRSF1 no mejoró los fenotipos de ojos rugosos (Figura 9e) o los déficits locomotores (Figura 9f) conferidos por la expresión de DPR de GR36 y PR36 independientes de G4C2 en Drosophila.

SRSF1 media la exportación nuclear del ARNm mediante la unión a NXF134,54. Anteriormente mostramos que cuatro residuos de arginina que se encuentran en el enlazador no estructurado entre los dos motivos de reconocimiento de ARN de SRSF1 (aminoácidos 11-196) son necesarios tanto para la exportación nuclear de ARN como para la

interacción con NXF1, mientras que las mutaciones de solo dos residuos de arginina conducen a una unión a NXF1 ligeramente reducida en células renales embrionarias humanas36. De manera similar, NXF1 endógena se inmunoprecipita específicamente en células neuronales N2A transfectadas con SRSF1 11-196 de tipo salvaje marcado con FLAG o doble mutante R117,118A (SRSF1-m2). Por el contrario, la coinmunoprecipitación de NXF1 está gravemente alterada por las mutaciones cuádruples R93,94,117,118A de SRSF1 (SRSF1-m4) (Figura 6g). La cotransfección del mutante dominante cuádruple SRSF1-m4 condujo además a una reducción marcada en la producción de DPR tanto sentido como antisentido, mientras que la secuencia de tipo salvaje o la variante que lleva dos mutaciones de arginina (SRSF1-m2), respectivamente, no tuvieron o tuvieron poco efecto (Figura 11c). Esto se evaluó estadísticamente tanto para poli-GP como para poli-GA DPR producidas por transcritos repetidos sentido (Figura 6c) y poli-GP DPR generadas a partir de transcritos repetidos antisentido (Figura 56b). Tomados en conjunto, nuestros datos demuestran que la expresión de los transcritos repetidos de C9ORF72 depende de un mecanismo que requiere la interacción de SRSF1 con el receptor de exportación nuclear NXF1. Por consiguiente, tanto el agotamiento de SRSF1 como la expresión del mutante dominante negativo SRSF1-m4 suprimen la neurotoxicidad mediada por la expresión de los transcritos repetidos de C9ORF72 en células neuronales N2A (Figura 11d). En apoyo a esto, el agotamiento mediado por el ARNi del homólogo de Drosophila NXF1, sbr, podría rescatar los déficits locomotores en larvas y moscas adultas que expresan G4C2x36 (Figura 28).

Ejemplo 8

El secuestro de SRSF1 desenhebra la exportación nuclear dependiente de NXF1 de transcritos repetidos de hexanucleótidos en células neuronales. Nuestro resultado muestra que la expresión de la proteína mutante SRSF1-m4 actúa como un mutante negativo dominante para la producción de DPR, lo que sugiere que la proteína SRSF1-m4 se secuestra en los transcritos repetidos de hexanucleótidos en lugar de la proteína SRSF1 endógena, lo que previene a su vez las interacciones de transcritos repetidos con NXF1 y la exportación nuclear. Mediante el uso de ensayos de reticulación UV in vitro, confirmamos que la proteína SRSF1-m4 etiquetada con hexahistidina recombinante purificada conserva la capacidad de interactuar directamente con G4C2x5 sintético de extremo 5'-terminal radiomarcado con 32P (Figura 8a) y el ARN repetido de C4G2x5 antisentido (Figura 8b). Estas interacciones son específicas ya que no se detectó unión del ARN en ausencia de irradiación UV o con la proteína control negativo MAGOH. A continuación, buscamos investigar si esto también era cierto en las células N2A vivas mediante el uso de ensayos de inmunoprecipitación de ARN (RIP). Se transfectaron células N2A con FLAG control, SRSF1 etiquetado con FLAG o SRSF1-m4 etiquetado con FLAG y diversas longitudes de construcciones de transcritos repetidos sentido o antisentido antes de la fijación de complejos de ribonucleoproteína. Después, los extractos celulares se sometieron a inmunoprecipitación anti-FLAG en las mismas condiciones usadas en la coinmunoprecipitación de NXF1 (Figura 6g) antes del análisis de qRT-PCR de moléculas de ARN reticuladas con SRSF1. En la validación de nuestro ensayo RIP, la inmunoprecipitación de SRSF1 o SRSF1-m4 condujo a la coprecipitación específica de SMN, un aglutinante de SRSF1 conocido55, pero no de JUN, un transcrito control sin intrones que no se espera que esté unido a SRSF156, en células N2A que expresan transcritos repetidos sentido (Figura 8c) o antisentido (Figura 8d). En agudo contraste, los niveles de transcritos repetidos de G4C2-sentido o C4G2-antisentido inmunoprecipitadas aumentaron significativamente con el número de repeticiones de hexanucleótidos (Figura 8c, d) que muestran el secuestro del ARN repetido dependiente de la longitud de SRSF1 y SRSF1-m4 en células neuronales.

Para evaluar los efectos del agotamiento de SRSF1 o la expresión de SRSF1-m4 en la exportación nuclear de los transcritos repetidos de C9ORF72, medimos los niveles totales y citoplasmáticos de los transcritos de G4C2x38 en presencia de Ctrl o SRSF1-ARNi y SRSF1 o SRSF1-m4 en células N2A transfectadas. La calidad del fraccionamiento celular se comprobó mediante inmunotransferencia mediante el uso de anticuerpos contra el factor SSRP1 de remodelación de la cromatina (Figura 8e) que muestra la ausencia de contaminación nuclear significativa en las fracciones citoplasmáticas. Los niveles totales de transcritos de G4C2x38 no se alteran significativamente con SRSF1-ARNi o la expresión de SRSF1-m4 mientras que los niveles citoplasmáticos de transcritos de G4C2x38 se reducen notablemente en ambas condiciones (Figura 29a). Los niveles de transcritos repetidos citoplasmáticos también se normalizaron a los niveles totales para evaluar específicamente el proceso de exportación nuclear como en nuestros estudios anteriores.[48, 49]. Las relaciones de niveles de ARNm citoplasmático/total se reducen notablemente tras la exposición a SRSF1-ARNi (Figura 8f, Ctrl-ARNi frente a SRSF1-ARNi). De manera similar, la cotransfección del mutante SRSF1-m4 que no interactúa con NXF1 condujo a una reducción marcada en los niveles de transcritos repetidos citoplasmáticos normalizados (Figura 8f, SRSF1 frente a SRSF1-m4). Para ampliar este análisis in vivo, las larvas de Drosophila que expresan G4C2x36 y Ctrl o SRSF1-ARNi se sometieron al mismo fraccionamiento (Figura 8g) y al análisis de transcritos. Al igual que con las células, los niveles totales de transcritos de G4C2x36 no se alteran significativamente en SRSF1-ARNi mientras que los niveles citoplasmáticos de transcritos de G4C2x36 se reducen notablemente (Figura 29b). En consecuencia, la relación del nivel de ARNm citoplasmático/total se reduce significativamente tras la exposición a SRSF1-ARNi (Figura 8h). Juntos, estos datos demuestran que el agotamiento de SRSF1 o la prevención de su secuestro de ARN repetido y la interacción con NXF1 inhiben específicamente la exportación nuclear de transcritos repetidos de hexanucleótidos in vitro e in vivo.

Ejemplo 9

El agotamiento de SRSF1 o la inhibición del secuestro de sus repeticiones de ARN y su interacción con NXF1 previenen la producción de DPR en neuronas primarias. A continuación, buscamos validar nuestros hallazgos en

neuronas primarias. Debido a la alta tinción de fondo obtenida con anticuerpos poli-GP y poli-GA, expresamos las etiquetas V5 en los tres marcos aguas abajo de la secuencia repetida G4C2x38 (Figura 30) para detectar simultáneamente todas las especies de DPR mediante el uso del anticuerpo anti-V5 más específico y sensible. Se transfectaron neuronas corticales de rata cultivadas con la construcción G4C2x38-3xV5 y los plásmidos de expresión Ctrl-ARNi, SRSF1-ARNi, SRSF1 o SRSF1-m4 antes de los estudios de inmunofluorescencia. La secuencia de nucleótidos dirigida por la horquilla 1 del miARN del ARNi de SRSF1 es idéntica en SRSF1 humana, de ratón y de rata (Métodos, Figura 5). En la Figura 22a se presentan ejemplos de imágenes microscópicas de neuronas DPR negativas y DPR positivas. La proporción de neuronas positivas para DPR en aproximadamente 100 neuronas transfectadas con éxito de dos experimentos independientes se cuantificó en cada grupo y todos los conteos se realizaron a ciegas. El agotamiento de SRSF1 condujo a una reducción significativa (25 %) en la proporción de neuronas con tinción con DPR traducida por RAN en comparación con las neuronas transfectadas con Ctrl-ARNi (Figura 22b). La inhibición del secuestro de SRSF1 endógena y la interacción con NXF1 mediante el reclutamiento del mutante dominante SRSF1-m4 también condujo a una reducción similar y significativa de neuronas que expresan DPR en comparación con las neuronas transfectadas con SRSF1 de tipo salvaje (Figura 22c). Sólo las células que mostraban ausencia de DPR se contaron como negativas para DPR. Es muy probable que los efectos de la expresión de SRSF1-ARNi o SRSF1-m4 se hayan subestimado ya que las neuronas que expresan cantidades reducidas de DPR todavía se habrían puntuado como DPR positivas. Concluimos que nuestros hallazgos en células N2A fueron corroborados en cultivos de neuronas primarias.

Ejemplo 10

El agotamiento de SRSF1 inhibe específicamente la exportación nuclear de transcritos patológicos que retienen expansiones de repetición de hexanucleótidos del intrón-1 en neuronas derivadas de pacientes con C9ORF72-ELA. Con el fin de investigar la exportación nuclear de los transcritos de C9ORF72 en el contexto de los genes C9ORF72 de tipo salvaje y expandidos repetidamente, diferenciamos las neuronas motoras de las células progenitoras neurales inducidas (iNPC) establecidas derivadas del control emparejado por sexo/edad y de fibroblastos de pacientes con C9ORF72-ELA52. Tanto las iNeuronas control como las inducidas por C9ORF72-ELA expresan el marcador de linaje neuronal Tuj1 y exhiben la propensión a formar ramificaciones complejas (Figura 31). El análisis de imágenes de alto contenido de la longitud axonal (Figura 23a) y el tamaño de las células del soma (Figura 23b) no mostró diferencias significativas entre las iNeuronas control y diferenciadas con C9ORF72-ELA en condiciones de cultivo basal, de acuerdo con informes anteriores[45,46]. Para probar un posible efecto neuroprotector de la depleción de SRSF1 en células relevantes para la enfermedad, a continuación se diferenciaron las iNPC en neuronas motoras (iMN), se transdujeron con un vector adenoviral que expresa RFP bajo el promotor Hb9 y se cultivaron en monocultivo o en cocultivo con iAstrocitos controles o derivados de ELA. No observamos un aumento de la muerte celular o anomalías morfológicas cuando los iMN de C9ORF72-ELA se cultivaron sin astrocitos (datos no mostrados). Sin embargo, sorprendentemente, la transducción de lentivirus SRSF1-ARNi en iMN antes de los cocultivos con iAstrocitos C9ORF72-ELA resultó en una supervivencia significativamente mayor de iMN contra la toxicidad derivada de los astrocíticos (Figura 23c), lo que indica un efecto neuroprotector de la depleción de SRSF1 en las neuronas motoras derivadas de pacientes con C9ORF72-ELA. De acuerdo con nuestros datos anteriores presentados en modelos de células neuronales y de Drosophila, también encontramos que el agotamiento de SRSF1 en las iMN derivadas de pacientes con C9ORF72-ELA conduce a una reducción específica en los niveles de expresión de poli-GP DPR (Figura 25b).

A continuación, cuantificamos los niveles totales, nucleares y citoplasmáticos de los transcritos de C9ORF72 empalmados con intrón-1 (mediante el uso de cebadores qPCR que se hibridan en el exón-1 y el exón-3)[50,51] y transcritos de C9ORF72 sin empalmar que retienen el intrón-1 (mediante el uso de cebadores qPCR que se hibridan en el exón-1 y en el intrón-1 aguas arriba de la expansión repetida de hexanucleótidos)[50, 51] para evaluar el impacto potencial de SRSF1-ARNi en el corte y empalme del intrón-1 y en la exportación nuclear de transcritos de C9ORF72 patológicos y de tipo salvaje. El agotamiento del factor de remodelación de la cromatina SSRP1 en las fracciones citoplasmáticas y de actina en las fracciones nucleares se usó para validar la calidad del fraccionamiento subcelular (Figura 23d). Los niveles de expresión relativos del ARNm de SRSF1 se regularon negativamente en aproximadamente un 80 % tras la transducción de iNeuronas con SRSF1-ARNi diferenciadas de dos líneas de pacientes controles o de dos líneas de pacientes C9ORF72-ELA emparejadas por sexo/edad (Figura 32). No se midieron cambios significativos en los niveles totales, nucleares o citoplasmáticos de transcritos empalmados con el intrón-1 entre el control y las iNeuronas C9ORF72-ELA transducidas o no con lentivirus SRSF1-ARNi (Figura 23e). Las relaciones del nivel de ARNm de SRSF1-ARNi (MOI5) sobre las no tratadas (MOI0) se graficaron adicionalmente para evaluar el impacto neto de SRSF1-ARNi en cada compartimento celular y control o en iNeuronas C9ORF72-ELA (Figura 23f). Estos datos muestran que la proporción de transcritos cortados y empalmados en exón1-exón3 no se altera en las neuronas C9ORF72-ELA y que la presencia de la expansión repetida de hexanucleótidos no afecta el corte y empalme del intrón-1, en total acuerdo con un estudio reciente[50]. Además, también mostramos aquí que el SRSF1-ARNi no tiene ningún efecto sobre la exportación nuclear o el corte y empalme del intrón-1 en los transcritos cortados y empalmados que conducen a la producción de la proteína C9ORF72 de tipo salvaje. Se llevó a cabo el mismo análisis experimental para los transcritos de C9ORF72 que retienen el intrón-1 (Figura 23g y 23h). No pudimos detectar niveles significativos (por encima de las reacciones de qRT-PCR control sin molde) de transcritos de C9ORF72 que retienen el intrón-1 en el citoplasma de las iNeuronas controles de acuerdo con la retención nuclear de los transcritos sin cortar y empalmar. En contraste llamativo, la presencia de la expansión de repetición de

hexanucleótidos en pacientes con C9ORF72-ELA desenhebra la exportación nuclear de transcritos repetidos de C9ORF72 que retienen el intrón-1 (Figura 23g), de acuerdo con nuestros datos anteriores que muestran que el secuestro de SRSF1 en expansiones de repetición de hexanucleótidos sintéticos promueve la exportación del ARNm nuclear a través de la interacción con NXF1 (Figura 8). Si bien el agotamiento de SRSF1 no afectó el nivel total y la biogénesis/estabilidad de los transcritos que retienen el intrón-1 en el control o en las iNeuronas C9ORF72-ELA, desenhebra específicamente una reducción citoplasmática y la acumulación nuclear de transcritos repetidos patológicos de C9ORF72 en iNeuronas C9ORF72-ELA (Figura 23g, h). Estos datos demuestran que el agotamiento de SRSF1 inhibe específicamente la exportación nuclear de transcritos repetidos de C9ORF72 expandidos. Tomados en conjunto, nuestros datos muestran que el agotamiento de SRSF1 no tiene ningún efecto sobre el nivel de expresión, el corte y empalme o la exportación nuclear de los transcritos de C9ORF72 en exón1-exón3 cortados y empalmados de tipo salvaje, mientras que inhibe específicamente la exportación nuclear de los transcritos patológicos de C9ORF72 que retienen las repeticiones de hexanucleótidos en el intrón-1.

Ejemplo 11

Las expansiones de microsatélites de 3-6 nucleótidos en regiones codificantes y no codificantes de genes causan neurodegeneración a través de mecanismos complejos que implican la pérdida de función de proteínas y mecanismos de ganancia de función tóxicos de proteínas/ARN[52]. La producción de proteínas repetidas poliméricas tóxicas por traducción por RAN ahora se ha caracterizado en múltiples trastornos neurodegenerativos causados por expansiones de microsatélites, incluida la ataxia espinocerebelosa tipo 8 (SCA8) 58, distrofia miotónica tipo 1 (DM1)[53], Síndrome de ataxia y temblor asociado al cromosoma X frágil (FXTAS) 59, C9ORF72-ELA6,13-16,27 y enfermedad de Huntington (HD) 60. Sin embargo, actualmente se desconocen los mecanismos implicados en la exportación nuclear de estos transcritos repetidos relacionados con la enfermedad.

Ejemplo 12

Anteriormente sugerimos que el secuestro de adaptadores de exportación nuclear en transcritos repetidos de C9ORF72 podría desenhebrar la exportación nuclear anormal de transcritos repetidos de C9ORF72 y la posterior traducción de DPR por RAN en el citoplasma12. En este estudio, identificamos por primera vez el mecanismo molecular que dirige la exportación nuclear de transcritos repetidos patológicos de C9ORF72. Investigamos si el agotamiento parcial de dos adaptadores de exportación nuclear conservados evolutivamente que interactúan ávidamente con los transcritos repetidos de hexanucleótidos12, ALYREF y SRSF1, mitigaría la neurotoxicidad mediada por DPR en un modelo establecido de Drosophila de C9ORF72-ELA16. Descubrimos que el agotamiento parcial de SRSF1 previene la neurodegeneración in vivo y suprime los déficits locomotores asociados, mientras que el agotamiento de ALYREF solo tuvo efectos marginales. El agotamiento de SRSF1 en neuronas motoras derivadas de pacientes C9ORF72-ELA también confirió neuroprotección de neuronas motoras en cocultivo con astrocitos C9ORF72-ELA. Además, también demostramos que esta intervención no afecta la morfología o el crecimiento de las neuronas motoras derivadas de pacientes controles y C9ORF72-ELA. Por otro lado, el agotamiento de SRSF1 en los astrocitos C9ORF72-ELA derivados de pacientes suprimió significativamente la muerte de las neuronas motoras en un sistema de cocultivo. Sin embargo, quedan por determinar los mecanismos para la supresión de la neurotoxicidad mediada por astrocitos. Ellos pueden implicar una modificación del ARN o la composición de proteínas en los exosomas extracelulares liberados por los astrocitos.

Ejemplo 13

Mediante el uso de células neuronales N2A, demostramos que la exportación nuclear de los transcritos repetidos de C9ORF72 y la posterior traducción por RAN depende de la interacción de SRSF1 con el receptor de exportación nuclear NXF1. La depleción de SRSF1 o la inhibición del secuestro de sus repeticiones de ARN endógeno y la interacción con NXF1 conduce a una inhibición marcada de la exportación nuclear de transcritos repetidos de C9ORF72 y la traducción de DPR sentido y antisentido por RAN para prevenir la neurotoxicidad mediada por repeticiones de C9ORF72. También demostramos que la inhibición dependiente de SRSF1 de la exportación nuclear de transcritos repetidos de C9ORF72 conduce a una producción alterada de DPR en Drosophila y en neuronas motoras derivadas de pacientes. Es importante destacar que el agotamiento de SRSF1 en las neuronas controles o en las derivadas de pacientes C9ORF72-ELA no afecta los niveles de expresión o la exportación nuclear de los transcritos cortados y empalmados con el intrón-1 requeridos para la traducción de la proteína C9ORF72 de tipo salvaje. Esto también indica que la exportación nuclear de transcritos de C9ORF72 no repetidos no depende del adaptador de exportación nuclear SRSF1. En marcado contraste con las neuronas controles, la presencia de la expansión de repetición de hexanucleótidos en el intrón-1 de los transcritos de C9ORF72 condujo a la exportación nuclear de ARNm dependiente de SRSF1, mientras que el agotamiento de SRSF1 inhibe específicamente la exportación nuclear pero no los niveles o el corte y empalme de los transcritos de C9ORF72 que retienen las repeticiones de hexanucleótidos expandidos en el intrón-1. Al tomar estos datos juntos, mostramos que el secuestro de SRSF1 en repeticiones de hexanucleótidos de C9ORF72 es capaz de autorizar la exportación nuclear dependiente de NXF1 de transcritos repetidos patológicos de C9ORF72 sin el acoplamiento funcional del proceso de exportación nuclear al corte y empalme del pre-ARNm. Esto explica a su vez por qué el agotamiento de SRSF1 no tiene ningún efecto sobre el nivel, el corte y empalme del intrón-1 o la exportación nuclear de los transcritos de C9ORF72 de tipo salvaje.

Ejemplo 14

En conclusión, hemos dilucidado por primera vez el mecanismo molecular que impulsa la exportación nuclear de transcritos repetidos patológicos de C9ORF72 que permiten la traducción por RAN de proteínas repetidas de dipéptidos en el citoplasma (Figura 24a). El agotamiento de SRSF1 inhibe específicamente la exportación nuclear de los transcritos de C9ORF72 patológicamente expandidos sin interferir con la biogénesis/procesamiento de los transcritos de C9ORF72 de tipo salvaje (Figura 24b). La expresión de la proteína SRSF1-m4 modificada, que conserva la capacidad específica de ser secuestrada en transcritos repetidos pero no interactúa eficazmente con NXF1, también inhibe la exportación nuclear de transcritos repetidos y la producción de DPR. (Figura 24c). Ambas intervenciones representan perspectivas prometedoras para el desarrollo de una estrategia neuroprotectora efectiva en la ELA relacionada con C9ORF72. Los efectos de antagonizar SRSF1 en el cerebro de vertebrados quedan por dilucidar en ratones de tipo salvaje, así como en modelos murinos C9ORF72-ELA. Curiosamente, recientemente se demostró que mientras SRSF1 se une directamente a miles de transcritos, el agotamiento de SRSF1 de forma aislada afecta la exportación nuclear de solo 225 transcritos (<1 % de los genes codificantes transcritos) debido a la presencia de 6 factores SRSF adicionales (SRSF2-7) que actúan como adaptadores de exportación nuclear redundantes dependientes de NXF1 46. Las vías celulares que causan la neurodegeneración mediada por repeticiones de C9ORF72 y el mecanismo o mecanismos precisos de neuroprotección conferidos por el direccionamiento de SRSF1 quedan por dilucidar en estudios futuros.

La inhibición de la exportación nuclear de transcritos repetidos también podría conferir neuroprotección en otros trastornos de expansión de microsatélites. Sin embargo, seguirá siendo esencial determinar las contribuciones fisiopatológicas entre la producción de proteínas con repeticiones poliméricas y la toxicidad mediada por ARN por la retención nuclear de transcritos y/o secuestro de factores de procesamiento del ARN en transcritos repetidos. Si bien la expresión de proteínas repetidas puede matar células in vitro, es difícil evaluar los niveles de traducción por RAN en pacientes y los umbrales requeridos para desenhebrar neurotoxicidad, que diferirán dependiendo de la naturaleza de las expansiones repetidas, la enfermedad en cuestión y los tipos celulares. Sin embargo, existe una evidencia creciente de un papel patógeno de la traducción por RAN y los datos presentados aquí apoyan completamente esto. Por ejemplo, inicialmente se pensó que FXTAS era causado por la retención intranuclear de transcritos y el secuestro de factores de corte y empalme[54,55]. Sin embargo, el descubrimiento de la traducción por RAN en el mismo modelo desafió este punto de vista[56]. De manera similar, en C9ORF72-ELA, un aumento de 10 veces en el número de focos de ARN intranuclear no altera significativamente la supervivencia o el procesamiento global del ARN, mientras que la expresión de DPR causó neurodegeneración[50] en total concordancia con los datos presentados aquí. El agotamiento parcial de los adaptadores de exportación nuclear individuales no parece ser perjudicial para el funcionamiento de las células eucariotas superiores. Por lo tanto, podrían constituir dianas terapéuticas viables para inhibir la exportación nuclear de transcritos repetidos y la producción de proteínas repetidas tóxicas, particularmente en enfermedades neurodegenerativas donde la traducción por RAN parece tener un papel patológico prominente.

Ejemplo 15

Expresión de péptidos SRSF1 antagonistas permeables a las células o codificados por AAV que interactúan con NXF1 aminoácidos 89-120 de SRSF1 (SEQ ID NO 15): PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE
Hemos usado el siguiente péptido y hemos mostrado una reducción en la producción de DPR expresados de una manera dependiente de RAN relevante para la enfermedad en células HEK humanas transfectadas con construcciones repetidas G4C2x38 que expresan etiquetas 3xV5 en todos los marcos (Figura 33).

SEQ ID NO 153

Secuencia SRSF1 (sitio de unión a TAP/NXF1) *Etiqueta V5* ***TAT PTD*** **PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE** GG *GKPIPNPLLGLDSTGG* ***YGRKKRRQRRR***

Ejemplo 16

Expresión de los aminoácidos 1-655 de SRPK1 (SEQ ID NO 19) para mantener el estado de fosforilación de SRSF1 para inhibir la interacción de SRSF1 con NXF1.

Ejemplo 17

Modelos de ratón de ELA/ FTD relacionados con C9ORF72

Los ratones de Ranum (Liu Y otros. Neuron 2016; 90: 521-34)

Modelo de ratón BAC de C9orf72 ELA/FTD que muestra disminución de la supervivencia, parálisis, denervación muscular, pérdida de neuronas motoras, comportamiento similar a la ansiedad y neurodegeneración cortical e hipocampal. Estos ratones expresan transcritos de C9ORF72 sentido y transcritos antisentido regulados positivamente. Este es nuestro primer modelo de elección para probar nuestros vectores de terapia génica.

Cleveland and Lagier-Tourenne's mice (Jiang y otros. Neuron 2016; 90:535-50)

Ratones que expresan el ARN de C9ORF72 con hasta 450 repeticiones de GGGGCC o con uno o ambos alelos de C9orf72 inactivados.

Virally-delivered models (Chew J. Science 2015; 348:1151-4)

Expresión de (G4C2) 66 en todo el sistema nervioso central murino mediante transgénesis cerebral somática mediada por virus adenoasociados.

Ejemplo 18 Vías de administración

Líquido cefalorraquídeo (LCR) mediante el uso de cisterna magna o suministro intracerebroventricular de vectores de VAA. El suministro de genes mediado por VAA se aplicará en 2 puntos de tiempo diferentes; antes del inicio e inicio de los síntomas. Los ratones transgénicos (Liu Y otros. Neuron 2016; 90:521-34) se dividirán en 2 grupos para ser tratados con un vector terapéutico como se describió anteriormente o un virus control [12 µL (2 µL de yodixanol); 1,2 x 1013/mL], vía cisterna magna o intracerebroventricular (ICV). Se sacrificarán 5 ratones por grupo 4 semanas después de la inyección y se recolectará tejido del SNC para evaluar la biodistribución viral (en la médula espinal) y los niveles de ARNm y proteínas de SRF1 (en la médula espinal, el cuerpo estriado, el tronco cerebral, el cerebelo y la corteza). Los 15 ratones restantes por grupo se sometieron a pruebas de comportamiento para evaluar los efectos sobre la progresión de la enfermedad, incluido el análisis semanal del rotarod después de tres días consecutivos de entrenamiento y análisis de la marcha. En todos los estudios, el inicio y la progresión de la enfermedad se evaluarán mediante puntuación neurológica 3 veces por semana a partir de los 60 días de edad y los ratones también se pesaron semanalmente. Los ratones se puntuarán por temblor, extensión de las extremidades traseras y déficit neurológico general mediante el uso de un sistema de puntuación previamente informado (Mead RJ y otros. PLoS One 2011; 6: e23244). Todos los ratones continuaron hasta la etapa terminal de la enfermedad y se registrará el tiempo para llegar a esta etapa. Todos los animales se perfundirán bajo anestesia terminal y se colectó tejido del SNC para evaluar la biodistribución viral, el ARNm de SRF1 y los niveles de proteína.

Suministro intravenoso

Los ratones se inyectarán en la vena facial o de la cola bajo anestesia con isoflurano con el genoma del vector 1-5x1011 de los vectores terapéuticos como se describió anteriormente o de los virus controles. A continuación, se permitió que los ratones se recuperaran, se enrollaron en el aserrín de su jaula original e inmediatamente se regresaron a su jaula. El análisis del comportamiento se llevará a cabo mediante el uso del diseño descrito anteriormente.

Referencias

1. M. DeJesus-Hernandez y otros. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 72, 245-256 (2011).
2. AE Renton y otros. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 72, 257-268 (2011).
3. J. Cooper-Knock y otros. C9ORF72 GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. PLoS ONE. 10, e0127376 (2015).
4. J. Cooper-Knock y otros. Antisense RNA foci in the motor neurons of. Acta Neuropathol. 130, 63-75 (2015).
5. K. Mori y otros. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol. 125, 413-423 (2013).
6. CJ Donnelly y otros. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron. 80, 415-428 (2013).
7. D. Sareen y otros. Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. Sci Transl Med. 5, 208ra149 (2013).
8. Y.-B. Lee y otros. Hexanucleotide Repeats in ALS/FTD Form Long-Depend RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. Cell Rep. 5, 1178-1186 (2013).
9. J. Cooper-Knock y otros. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain. 137, 2040-2051 (2014).
10. K. Mori y otros. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science. 339, 1335-1338 (2013).
11. K. Mori y otros. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol. 126, 881-893 (2013).
12. PEA Ash y otros. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron. 77, 639-646 (2013).
13. T. Zu y otros. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ELA and frontotemporal dementia. Proc. Natl. Acad. Sci. U.S.A. 110, E4968-77 (2013).
14. S. Mizielinska y otros. C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science. 345, 1192-1194 (2014).

15. S. Ciura y otros. Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. Ann. Neurol. 74:180-187 (2013), doi:10.1002/ana.23946.
16. JG O'Rourke y otros. C9orf72 is required for proper macrophage and microglial function in mice. Science. 351, 1324-1329 (2016).
17. MJ Walsh y otros. Invited review: decoding the pathophysiological mechanisms that underlie RNA dysregulation in neurodegenerative disorders: a review of the current state of the art. Neuropathol. Appl. Neurobiol. 41, 109-134 (2015).
18. JD Rohrer y otros. C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. Lancet Neurol. 14, 291-301 (2015).
19. TW Todd, L. Petrucelli, Insights into the pathogenic mechanisms of Chromosome 9 open reading frame 72 (C9orf72) repeat expansions. J Neurochem (2016), doi:10.1111/jnc.13623.
20. D. Edbauer, C. Haass, An amyloid-like cascade hypothesis for C9orf72 ELA/FTD. Curr. Opin. Neurobiol. 36, 99-106 (2016).
21. BD Freibaum y otros. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature. 525, 129-133 (2015).
22. AP Golovanov, GM Hautbergue, AM Tintaru, L.-Y. Lian, SA Wilson, The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. RNA. 12, 1933-1948 (2006).
23. AM Tintaru y otros. Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep. 8, 756-762 (2007).
24. GM Hautbergue, M.-L. Hung, AP Golovanov, L.-Y. Lian, SA Wilson, Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. Proc. Natl. Acad. Sci. U.S.A. 105, 5154-5159 (2008).
25. N. Viphakone y otros. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. Nat Commun. 3, 1006 (2012).
26. Y. Huang, R. Gattoni, J. Stevenin, JA Steitz, SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. Molecular Cell. 11, 837-843 (2003).
27. F. Stutz y otros. REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA. 6, 638-650 (2000).
28. G. Dietzl y otros. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature. 448, 151-156 (2007).
29. K. Meyer y otros. Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. Proc. Natl. Acad. Sci. U.S.A. 111, 829-832 (2014).
30. Y. Huang, TA Yario, JA Steitz, A molecular link between SR protein dephosphorylation and mRNA export. Proc. Natl. Acad. Sci. U.S.A. 101, 9666-9670 (2004).
31. R. Karni y otros. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat. Struct. Mol. Biol. 14, 185-193 (2007).
32. JC Greene y otros. Mitochondrial pathology and apoptotic muscle degeneration in Drosophila parkin mutants. Proc. Natl. Acad. Sci. U.S.A. 100, 4078-4083 (2003).
33. W. Sullivan, M. Ashburner, RS Hawley, Drosophila protocols. (Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 2000).
34. N. Deglon, JL Tseng, JC Bensadoun, Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease. Human Gene Ther. 11, 179-190 (2000).
35. H. Wichterle, I. Lieberam, JA Porter, TM Jessell, Directed differentiation of embryonic stem cells into motor neurons. Cell. 110, 385-397 (2002).
36. D. Ling, PM Salvaterra, Robust RT-qPCR data normalization: validation and selection of internal reference genes during post-experimental data analysis. PLoS ONE. 6, e17762 (2011).
37. Webster, C. P. y otros. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. EMBO J. 35, 1656-1676 (2016).
38. Jovičić, A. y otros. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat. Neurosci. 18, 1226-1229 (2015).
39. Golovanov, AP, Hautbergue, GM, Tintaru, AM, Lian, L.-Y. & Wilson, S.A. The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. RNA 12, 1933-1948 (2006).
40. Tintaru, A. M. y otros. Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep. 8, 756-762 (2007).
41. Meyer, K. y otros. Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. Proc. Natl. Acad. Sci. U.S.A. 111, 829-832 (2014).
42. Deglon, N., Tseng, J.L. & Bensadoun, J.C. Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease. Human Gene Ther. 11, 179-190 (2000).
43. Cooper-Knock, J. y otros. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain 137, 2040-2051 (2014).
44. Cooper-Knock, J. y otros. Antisense RNA foci in the motor neurons of. Acta Neuropathol 130, 63-75 (2015).
45. Mizielinska, S. y otros. C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345, 1192-1194 (2014).
46. Donnelly, C. J. y otros. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron 80, 415-428 (2013).

47. Sareen, D. y otros. Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. Sci Transl Med 5, 208ra149 (2013).
48. Masuda, S. y otros. Recruitment of the human TREX complex to mRNA during splicing. Genes & Development 19, 1512-1517 (2005).
49. Chang, C.-T. y otros. Chtop is a component of the dynamic TREX mRNA export complex. EMBO J. 32, 473-486 (2013).
50. Hautbergue, G. M. y otros. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. Curr. Biol. 19, 1918-1924 (2009).
51. Tran, H. y otros. Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ELA. Neuron 87, 1207-1214 (2015).
52. Niblock, M. y otros. Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALSALS/FTD. Acta Neuropathol Commun 4, 18 (2016).
53. Loureiro, J. R., Oliveira, C. L. & Silveira, I. Unstable repeat expansions in neurodegenerative diseases: nucleocytoplasmic transport emerges on the scene. Neurobiol. Aging 39, 174-183 (2016).
54. Zu, T. y otros. Non-ATG-initiated translation directed by microsatellite expansions. Proc. Natl. Acad. Sci. U.S.A. 108, 260-265 (2011).
55. Jin, P. y otros. Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron 55, 556-564 (2007).
56. Sofola, O. A. y otros. RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron 55, 565-571 (2007).
57. Todd, P. K. y otros. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78, 440-455 (2013).
58. Greene, J. C. y otros. Mitochondrial pathology and apoptotic muscle degeneration in Drosophila parkin mutants. Proc. Natl. Acad. Sci. U.S.A. 100, 4078-4083 (2003).
59. Sullivan, W., Ashburner, M. & Hawley, R. S. Drosophila protocols. (Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press., 2000).
60. Deglon, N., Tseng, J. L. & Bensadoun, R. S. Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease. Human Gene The. 11, 179-190 (2000).
61. Wichterle, H., Lieberam, I., Porter, J. A. & Jessell, T. M. Directed differentiation of embryonic stem cells into motor neurons. Cell 110, 385-397 (2002).
62. Ye, J. y otros. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC se sabe que 13, 134 (2012).
63. Ling, D. & Salvaterra, P. M. Robust RT-qPCR data normalization: validation and selection of internal reference genes during post-experimental data analysis. PLoS ONE 6, e17762 (2011).
64. Yang, D. y otros. FTD/ALS-associated poly (GR) protein impairs the Notch pathway and is recruited by poly (GA) into cytoplasmic inclusions. Acta Neuropathol (2015). doi: 10.1007/s00401-015-1448-6
65. Maranda, B., Fan, L., Soucy, J.-F., Simard, L. & Mitchell, G. A. Spinal muscular atrophy: clinical validation of a single-tube multiplex real time PCR assay for determination of SMN1 and SMN2 copy numbers. Clin. Biochem. 45, 88-91 (2012).

LISTADO DE SECUENCIAS

<110> Universidad de Sheffield

<120> TERAPIA GÉNICA

<130> 4589P/WO

<150> GB1609597.8
<151> 2016-06-01

<160> 158

<170> Versión PatentIn 3.5

<210> 1
<211> 747
<212> ADN
<213> homo sapiens

<400> 1

```
atgtcgggag gtggtgtgat tcgtggcccc gcagggaaca acgattgccg catctacgtg      60
ggtaacttac ctccagacat ccgaaccaag gacattgagg acgtgttcta caaatacggc     120
gctatccgcg acatcgacct caagaatcgc cgcgggggac cgcccttcgc cttcgttgag     180
ttcgaggacc cgcgagacgc ggaagacgcg gtgtatggtc gcgacggcta tgattacgat     240
gggtaccgtc tgcgggtgga gtttcctcga agcggccgtg gaacaggccg aggcggcggc     300
gggggtggag gtggcggagc tccccgaggt cgctatggcc ccccatccag gcggtctgaa     360
aacagagtgg ttgtctctgg actgcctcca agtggaagtt ggcaggattt aaaggatcac     420
atgcgtgaag caggtgatgt atgttatgct gatgtttacc gagatggcac tggtgtcgtg     480
gagtttgtac ggaaagaaga tatgacctat gcagttcgaa aactggataa cactaagttt     540
agatctcatg agggagaaac tgcctacatc cgggttaaag ttgatgggcc cagaagtcca     600
agttatggaa gatctcgatc tcgaagccgt agtcgtagca gaagccgtag cagaagcaac     660
agcaggagtc gcagttactc cccaaggaga agcagaggat caccacgcta ttctccccgt     720
catagcagat ctcgctctcg tacataa                                        747
```

<210> 2
<211> 248
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                  10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
        35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

Gly Tyr Arg Leu Arg Val Glu Phe Pro Arg Ser Gly Arg Gly Thr Gly
                85                  90                  95

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145                 150                 155                 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
                165                 170                 175

Asn Thr Lys Phe Arg Ser His Glu Gly Glu Thr Ala Tyr Ile Arg Val
            180                 185                 190

Lys Val Asp Gly Pro Arg Ser Pro Ser Tyr Gly Arg Ser Arg Ser Arg
        195                 200                 205

Ser Arg Ser Arg Ser Arg Ser Arg Ser Arg Ser Asn Ser Arg Ser Arg
    210                 215                 220

Ser Tyr Ser Pro Arg Arg Ser Arg Gly Ser Pro Arg Tyr Ser Pro Arg
225                 230                 235                 240

His Ser Arg Ser Arg Ser Arg Thr
                245
```

<210> 3
<211> 606
<212> ADN
<213> Homo sapiens

<400> 3

```
atgtcgggag gtggtgtgat tcgtggcccc gcagggaaca acgattgccg catctacgtg       60

ggtaacttac ctccagacat ccgaaccaag gacattgagg acgtgttcta caaatacggc      120

gctatccgcg acatcgacct caagaatcgc cgcgggggac cgcccttcgc cttcgttgag      180

ttcgaggacc cgcgagacgc ggaagacgcg gtgtatggtc gcgacggcta tgattacgat      240

gggtaccgtc tgcgggtgga gtttcctcga agcggccgtg gaacaggccg aggcggcggc      300

gggggtggag gtggcggagc tccccgaggt cgctatggcc ccccatccag gcggtctgaa      360

aacagagtgg ttgtctctgg actgcctcca agtggaagtt ggcaggattt aaaggatcac      420

atgcgtgaag caggtgatgt atgttatgct gatgtttacc gagatggcac tggtgtcgtg      480

gagtttgtac ggaaagaaga tatgacctat gcagttcgaa aactggataa cactaagttt      540

agatctcatg aggtaggtta tacacgtatt cttttctttg accagaattg gatacagtgg      600

tcttaa                                                                 606
```

<210> 4
<211> 201
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                  10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
        35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

Gly Tyr Arg Leu Arg Val Glu Phe Pro Arg Ser Gly Arg Gly Thr Gly
                85                  90                  95

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140
```

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145 150 155 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
165 170 175

Asn Thr Lys Phe Arg Ser His Glu Val Gly Tyr Thr Arg Ile Leu Phe
180 185 190

Phe Asp Gln Asn Trp Ile Gln Trp Ser
195 200

<210> 5
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 5
aactgcctac atccgggtta a 21

<210> 6
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 6
ttaaagttga tgggcccaga a 21

<210> 7
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 7
gggcccagaa gtccaagtta t 21

<210> 8
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 8
agcagaggat caccacgcta t 21

<210> 9
<211> 21
<212> ADN
<213> secuencia artificial

<220>

<223> ARN inhibidor

<400> 9
tcaataatgg aggcaatggt a 21

<210> 10
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 10
aatggtatga ctccaagtgc t 21

<210> 11
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 11
gctaatttgt cacagtgctt a 21

<210> 12
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 12
gttaatgtgt gacctgctgt t 21

<210> 13
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 13
actgctaaat ctgcatgtcc t 21

<210> 14
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 14
tgatagagcg ttgctatttc a 21

<210> 15
<211> 32
<212> PRT
<213> secuencia artificial

<220>

<223> agente antagonista

<400> 15

```
Pro Arg Ser Gly Arg Gly Thr Gly Arg Gly Gly Gly Gly Gly Gly Gly
1               5                  10                  15

Gly Gly Ala Pro Arg Gly Arg Tyr Gly Pro Pro Ser Arg Arg Ser Glu
            20                  25                  30
```

<210> 16
<211> 196
<212> PRT
<213> secuencia artificial

<220>
<223> proteína modificada

<400> 16

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                  10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
        35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

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

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145                 150                 155                 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
                165                 170                 175

Asn Thr Lys Phe Arg Ser His Glu Gly Glu Thr Ala Tyr Ile Arg Val
            180                 185                 190

Lys Val Asp Gly
        195
```

<210> 17
<211> 187
<212> PRT
<213> secuencia artificial

<220>
<223> proteína modificada

<400> 17

```
Met Ala Gly Asn Asn Asp Cys Arg Ile Tyr Val Gly Asn Leu Pro Pro
1               5                   10                  15

Asp Ile Arg Thr Lys Asp Ile Glu Asp Val Phe Tyr Lys Tyr Gly Ala
            20                  25                  30

Ile Arg Asp Ile Asp Leu Lys Asn Arg Arg Gly Gly Pro Pro Phe Ala
        35                  40                  45

Phe Val Glu Phe Glu Asp Pro Arg Asp Ala Glu Asp Ala Val Tyr Gly
    50                  55                  60

Arg Asp Gly Tyr Asp Tyr Asp Gly Tyr Arg Leu Arg Val Glu Phe Pro
65                  70                  75                  80

Ala Ser Gly Ala Gly Thr Gly Arg Gly Gly Gly Gly Gly Gly Gly Gly
                85                  90                  95

Met Ala Gly Asn Asn Asp Cys Arg Ile Tyr Val Gly Asn Leu Pro Pro
1               5                   10                  15

Asp Ile Arg Thr Lys Asp Ile Glu Asp Val Phe Tyr Lys Tyr Gly Ala
            20                  25                  30

Ile Arg Asp Ile Asp Leu Lys Asn Arg Arg Gly Gly Pro Pro Phe Ala
        35                  40                  45

Phe Val Glu Phe Glu Asp Pro Arg Asp Ala Glu Asp Ala Val Tyr Gly
    50                  55                  60

Arg Asp Gly Tyr Asp Tyr Asp Gly Tyr Arg Leu Arg Val Glu Phe Pro
65                  70                  75                  80

Ala Ser Gly Ala Gly Thr Gly Arg Gly Gly Gly Gly Gly Gly Gly Gly
                85                  90                  95
```

```
Gly Ala Pro Arg Gly Arg Tyr Gly Pro Pro Ser Ala Ala Ser Glu Asn
            100                 105                 110

Arg Val Val Val Ser Gly Leu Pro Pro Ser Gly Ser Trp Gln Asp Leu
        115                 120                 125

Lys Asp His Met Arg Glu Ala Gly Asp Val Cys Tyr Ala Asp Val Tyr
    130                 135                 140

Arg Asp Gly Thr Gly Val Val Glu Phe Val Arg Lys Glu Asp Met Thr
145                 150                 155                 160

Tyr Ala Val Arg Lys Leu Asp Asn Thr Lys Phe Arg Ser His Glu Gly
                165                 170                 175

Glu Thr Ala Tyr Ile Arg Val Lys Val Asp Gly
            180                 185
```

<210> 18
<211> 1968
<212> ADN
<213> homo sapiens

<400> 18

```
atggagcgga aagtgcttgc gctccaggcc cgaaagaaaa ggaccaaggc caagaaggac       60

aaagcccaaa ggaaatctga aactcagcac cgaggctctg ctccccactc tgagagtgat      120

ctaccagagc aggaagagga gattctggga tctgatgatg atgagcaaga agatcctaat      180

gattattgta aaggaggtta tcatcttgtg aaaattggag atctattcaa tgggagatac      240

catgtgatcc gaaagttagg ctggggacac ttttcaacag tatggttatc atgggatatt      300

caggggaaga aatttgtggc aatgaaagta gttaaaagtg ctgaacatta cactgaaaca      360

gcactagatg aaatccggtt gctgaagtca gttcgcaatt cagaccctaa tgatccaaat      420

agagaaatgg ttgttcaact actagatgac tttaaaatat caggagttaa tggaacacat      480

atctgcatgg tatttgaagt tttggggcat catctgctca agtggatcat caaatccaat      540

tatcaggggc ttccactgcc ttgtgtcaaa aaaattattc agcaagtgtt acagggtctt      600

gattatttac ataccaagtg ccgtatcatc cacactgaca ttaaaccaga gaacatctta      660

ttgtcagtga atgagcagta cattcggagg ctggctgcag aagcaacaga atggcagcga      720

tctggagctc ctccgccttc cggatctgca gtcagtactg ctccccagcc taaaccagct      780

gacaaaatgt caaagaataa gaagaagaaa ttgaagaaga agcagaagcg ccaggcagaa      840

ttactagaga agcgaatgca ggaaattgag gaaatggaga aagagtcggg ccctgggcaa      900

aaaagaccaa acaagcaaga agaatcagag agtcctgttg aaagaccctt gaaagagaac      960

ccacctaata aaatgaccca agaaaaactt gaagagtcaa gtaccattgg ccaggatcaa     1020
```

acgcttatgg aacgtgatac agagggtggt gcagcagaaa ttaattgcaa tggagtgatt 1080

gaagtcatta attatactca gaacagtaat aatgaaacat tgagacataa agaggatcta 1140

cataatgcta atgactgtga tgtccaaaat ttgaatcagg aatctagttt cctaagctcc 1200

caaaatggag acagcagcac atctcaagaa acagactctt gtacacctat aacatctgag 1260

gtgtcagaca ccatggtgtg ccagtcttcc tcaactgtag gtcagtcatt cagtgaacaa 1320

cacattagcc aacttcaaga aagcattcgg gcagagatac cctgtgaaga tgaacaagag 1380

caagaacata acggaccact ggacaacaaa ggaaaatcca cggctggaaa ttttcttgtt 1440

aatccccttg agccaaaaaa tgcagaaaag ctcaaggtga agattgctga ccttggaaat 1500

gcttgttggg tgcacaaaca tttcactgaa gatattcaaa caaggcaata tcgttccttg 1560

gaagttctaa tcggatctgg ctataatacc cctgctgaca tttggagcac ggcatgcatg 1620

gcctttgaac tggccacagg tgactatttg tttgaacctc attcagggga agagtacact 1680

cgagatgaag atcacattgc attgatcata gaacttctgg ggaaggtgcc tcgcaagctc 1740

attgtggcag gaaaatattc caaggaattt ttcaccaaaa aaggtgacct gaaacatatc 1800

acgaagctga aaccttgggg ccttttttgag gttctagtgg agaagtatga gtggtcgcag 1860

gaagaggcag ctggcttcac agatttctta ctgcccatgt tggagctgat ccctgagaag 1920

agagccactg ccgccgagtg tctccggcac ccttggctta actcctaa 1968

<210> 19
<211> 655
<212> PRT
<213> homo sapiens

<400> 19

Met Glu Arg Lys Val Leu Ala Leu Gln Ala Arg Lys Lys Arg Thr Lys
1 5 10 15

Ala Lys Lys Asp Lys Ala Gln Arg Lys Ser Glu Thr Gln His Arg Gly
20 25 30

Ser Ala Pro His Ser Glu Ser Asp Leu Pro Glu Gln Glu Glu Glu Ile
35 40 45

Leu Gly Ser Asp Asp Asp Glu Gln Glu Asp Pro Asn Asp Tyr Cys Lys
50 55 60

Gly Gly Tyr His Leu Val Lys Ile Gly Asp Leu Phe Asn Gly Arg Tyr
65 70 75 80

His Val Ile Arg Lys Leu Gly Trp Gly His Phe Ser Thr Val Trp Leu
85 90 95

```
Ser Trp Asp Ile Gln Gly Lys Lys Phe Val Ala Met Lys Val Val Lys
            100                 105                 110

Ser Ala Glu His Tyr Thr Glu Thr Ala Leu Asp Glu Ile Arg Leu Leu
        115                 120                 125

Lys Ser Val Arg Asn Ser Asp Pro Asn Asp Pro Asn Arg Glu Met Val
    130                 135                 140

Val Gln Leu Leu Asp Asp Phe Lys Ile Ser Gly Val Asn Gly Thr His
145                 150                 155                 160

Ile Cys Met Val Phe Glu Val Leu Gly His His Leu Leu Lys Trp Ile
                165                 170                 175

Ile Lys Ser Asn Tyr Gln Gly Leu Pro Leu Pro Cys Val Lys Lys Ile
            180                 185                 190

Ile Gln Gln Val Leu Gln Gly Leu Asp Tyr Leu His Thr Lys Cys Arg
        195                 200                 205

Ile Ile His Thr Asp Ile Lys Pro Glu Asn Ile Leu Leu Ser Val Asn
    210                 215                 220

Glu Gln Tyr Ile Arg Arg Leu Ala Ala Glu Ala Thr Glu Trp Gln Arg
225                 230                 235                 240

Ser Gly Ala Pro Pro Pro Ser Gly Ser Ala Val Ser Thr Ala Pro Gln
                245                 250                 255

Pro Lys Pro Ala Asp Lys Met Ser Lys Asn Lys Lys Lys Lys Leu Lys
            260                 265                 270

Lys Lys Gln Lys Arg Gln Ala Glu Leu Leu Glu Lys Arg Met Gln Glu
        275                 280                 285

Ile Glu Glu Met Glu Lys Glu Ser Gly Pro Gly Gln Lys Arg Pro Asn
    290                 295                 300

Lys Gln Glu Glu Ser Glu Ser Pro Val Glu Arg Pro Leu Lys Glu Asn
305                 310                 315                 320

Pro Pro Asn Lys Met Thr Gln Glu Lys Leu Glu Glu Ser Ser Thr Ile
                325                 330                 335

Gly Gln Asp Gln Thr Leu Met Glu Arg Asp Thr Glu Gly Gly Ala Ala
```

```
                340                 345                 350

Glu Ile Asn Cys Asn Gly Val Ile Glu Val Ile Asn Tyr Thr Gln Asn
        355                 360                 365

Ser Asn Asn Glu Thr Leu Arg His Lys Glu Asp Leu His Asn Ala Asn
    370                 375                 380

Asp Cys Asp Val Gln Asn Leu Asn Gln Glu Ser Ser Phe Leu Ser Ser
385                 390                 395                 400

Gln Asn Gly Asp Ser Ser Thr Ser Gln Glu Thr Asp Ser Cys Thr Pro
                405                 410                 415

Ile Thr Ser Glu Val Ser Asp Thr Met Val Cys Gln Ser Ser Ser Thr
            420                 425                 430

Val Gly Gln Ser Phe Ser Glu Gln His Ile Ser Gln Leu Gln Glu Ser
        435                 440                 445

Ile Arg Ala Glu Ile Pro Cys Glu Asp Glu Gln Glu Gln Glu His Asn
    450                 455                 460

Gly Pro Leu Asp Asn Lys Gly Lys Ser Thr Ala Gly Asn Phe Leu Val
465                 470                 475                 480

Asn Pro Leu Glu Pro Lys Asn Ala Glu Lys Leu Lys Val Lys Ile Ala
                485                 490                 495

Asp Leu Gly Asn Ala Cys Trp Val His Lys His Phe Thr Glu Asp Ile
            500                 505                 510

Gln Thr Arg Gln Tyr Arg Ser Leu Glu Val Leu Ile Gly Ser Gly Tyr
        515                 520                 525

Asn Thr Pro Ala Asp Ile Trp Ser Thr Ala Cys Met Ala Phe Glu Leu
    530                 535                 540

Ala Thr Gly Asp Tyr Leu Phe Glu Pro His Ser Gly Glu Glu Tyr Thr
545                 550                 555                 560

Arg Asp Glu Asp His Ile Ala Leu Ile Ile Glu Leu Leu Gly Lys Val
                565                 570                 575

Pro Arg Lys Leu Ile Val Ala Gly Lys Tyr Ser Lys Glu Phe Phe Thr
            580                 585                 590
```

Lys Lys Gly Asp Leu Lys His Ile Thr Lys Leu Lys Pro Trp Gly Leu
595 600 605

Phe Glu Val Leu Val Glu Lys Tyr Glu Trp Ser Gln Glu Glu Ala Ala
610 615 620

Gly Phe Thr Asp Phe Leu Leu Pro Met Leu Glu Leu Ile Pro Glu Lys
625 630 635 640

Arg Ala Thr Ala Ala Glu Cys Leu Arg His Pro Trp Leu Asn Ser
645 650 655

<210> 20
<211> 1933
<212> ADN
<213> Homo sapiens

<400> 20

gtgcatgacc cgccccgcgg cggagacgcg ctcgctgcgt catcagtgtt ttcgagacga 60

gtctcgacgc agcagctgtc agctccattt tgttgttggt gcgcgacgca gtcagctgcg 120

tgattcccgt gattgcgtta caagctttgt ctccttcgac ttggagtctt tgtccaggac 180

gatgagacac tcaaagagaa cttactgtcc tgattgggat gacaaggatt gggattatgg 240

aaaatggagg agcagcagca gtcataaaag aaggaagaga tcacatagca gtgcccagga 300

gaacaagcgc tgcaaataca atcactctaa aatgtgtgat agccattatt tggaaagcag 360

gtctataaat gagaaagatt atcatagtcg acgctacatt gatgagtaca gaaatgacta 420

cactcaagga tgtgaacctg gacatcgcca aagagaccat gaaagccggt atcagaacca 480

tagtagcaag tcttctggta gaagtggaag aagtagttat aaaagcaaac acaggattca 540

ccacagtact tcacatcgtc gttcacatgg gaagagtcac cgaaggaaaa gaaccaggag 600

tgtagaggat gatgaggagg gtcacctgat ctgtcagagt ggagacgtac taagtgcaag 660

atatgaaatt gttgatactt taggtgaagg agcttttgga aaagttgtgg agtgcatcga 720

tcataaagcg ggaggtagac atgtagcagt aaaaatagtt aaaaatgtgg atagatactg 780

tgaagctgct cgctcagaaa tacaagttct ggaacatctg aatacaacag accccaacag 840

tactttccgc tgtgtccaga tgttggaatg gtttgagcat catggtcaca tttgcattgt 900

ttttgaacta ttgggactta gtacttacga cttcattaaa gaaaatggtt ttctaccatt 960

tcgactggat catatcagaa agatggcata tcagatatgc aagtctgtga attttttgca 1020

cagtaataag ttgactcaca cagacttaaa gcctgaaaac atcttatttg tgcagtctga 1080

ctacacagag gcgtataatc ccaaaataaa acgtgatgaa cgcaccttaa taaatccaga 1140

tattaaagtt gtagactttg gtagtgcaac atatgatgac gaacatcaca gtacattggt 1200

atctacaaga cattatagag cacctgaagt tattttagcc ctagggtggt cccaaccatg 1260

```
tgatgtctgg agcataggat gcattcttat tgaatactat cttgggttta ccgtatttcc     1320

aacacacgat agtaaggagc atttagcaat gatggaaagg attcttggac ctctaccaaa     1380

acatatgata cagaaaacca ggaaacgtaa atattttcac cacgatcgat tagactggga     1440

tgaacacagt tctgccggca gatatgtttc aagacgctgt aaacctctga aggaatttat     1500

gctttctcaa gatgttgaac atgagcgtct ctttgacctc attcagaaaa tgttggagta     1560

tgatccagcc aaaagaatta ctctcagaga agccttaaag catcctttct ttgaccttct     1620

gaagaaaagt atatagatct gtaattggac agctctctcg aagagatctt acagactgta     1680

tcagtctaat ttttaaattt taagttattt tgtacagctt tgtaaattct taacattttt     1740

atattgccat gtttattttg tttgggtaat ttggttcatt aagtacatag ctaaggtaat     1800

gaacatcttt ttcagtaatt gtaaagtgat ttattcagaa taaatttttt gtgcttatga     1860

agttgatatg tatctgaaca gtttgttcta agtaccattt ttcttcctac ttctattaaa     1920

gaatggacat aga                                                       1933
```

<210> 21
<211> 484
<212> PRT
<213> homo sapiens

<400> 21

```
Met Arg His Ser Lys Arg Thr Tyr Cys Pro Asp Trp Asp Asp Lys Asp
1               5                   10                  15

Trp Asp Tyr Gly Lys Trp Arg Ser Ser Ser Ser His Lys Arg Arg Lys
            20                  25                  30

Arg Ser His Ser Ser Ala Gln Glu Asn Lys Arg Cys Lys Tyr Asn His
        35                  40                  45

Ser Lys Met Cys Asp Ser His Tyr Leu Glu Ser Arg Ser Ile Asn Glu
    50                  55                  60

Lys Asp Tyr His Ser Arg Arg Tyr Ile Asp Glu Tyr Arg Asn Asp Tyr
65                  70                  75                  80

Thr Gln Gly Cys Glu Pro Gly His Arg Gln Arg Asp His Glu Ser Arg
                85                  90                  95

Tyr Gln Asn His Ser Ser Lys Ser Ser Gly Arg Ser Gly Arg Ser Ser
            100                 105                 110

Tyr Lys Ser Lys His Arg Ile His His Ser Thr Ser His Arg Arg Ser
        115                 120                 125
```

```
His Gly Lys Ser His Arg Arg Lys Arg Thr Arg Ser Val Glu Asp Asp
    130                 135                 140

Glu Glu Gly His Leu Ile Cys Gln Ser Gly Asp Val Leu Ser Ala Arg
145                 150                 155                 160

Tyr Glu Ile Val Asp Thr Leu Gly Glu Gly Ala Phe Gly Lys Val Val
                165                 170                 175

Glu Cys Ile Asp His Lys Ala Gly Gly Arg His Val Ala Val Lys Ile
            180                 185                 190

Val Lys Asn Val Asp Arg Tyr Cys Glu Ala Ala Arg Ser Glu Ile Gln
        195                 200                 205

Val Leu Glu His Leu Asn Thr Thr Asp Pro Asn Ser Thr Phe Arg Cys
    210                 215                 220

Val Gln Met Leu Glu Trp Phe Glu His His Gly His Ile Cys Ile Val
225                 230                 235                 240

Phe Glu Leu Leu Gly Leu Ser Thr Tyr Asp Phe Ile Lys Glu Asn Gly
                245                 250                 255

Phe Leu Pro Phe Arg Leu Asp His Ile Arg Lys Met Ala Tyr Gln Ile
            260                 265                 270

Cys Lys Ser Val Asn Phe Leu His Ser Asn Lys Leu Thr His Thr Asp
        275                 280                 285

Leu Lys Pro Glu Asn Ile Leu Phe Val Gln Ser Asp Tyr Thr Glu Ala
    290                 295                 300

Tyr Asn Pro Lys Ile Lys Arg Asp Glu Arg Thr Leu Ile Asn Pro Asp
305                 310                 315                 320

Ile Lys Val Val Asp Phe Gly Ser Ala Thr Tyr Asp Asp Glu His His
                325                 330                 335

Ser Thr Leu Val Ser Thr Arg His Tyr Arg Ala Pro Glu Val Ile Leu
            340                 345                 350

Ala Leu Gly Trp Ser Gln Pro Cys Asp Val Trp Ser Ile Gly Cys Ile
        355                 360                 365

Leu Ile Glu Tyr Tyr Leu Gly Phe Thr Val Phe Pro Thr His Asp Ser
```

370 375 380

Lys Glu His Leu Ala Met Met Glu Arg Ile Leu Gly Pro Leu Pro Lys
385 390 395 400

His Met Ile Gln Lys Thr Arg Lys Arg Lys Tyr Phe His His Asp Arg
405 410 415

Leu Asp Trp Asp Glu His Ser Ser Ala Gly Arg Tyr Val Ser Arg Arg
420 425 430

Cys Lys Pro Leu Lys Glu Phe Met Leu Ser Gln Asp Val Glu His Glu
435 440 445

Arg Leu Phe Asp Leu Ile Gln Lys Met Leu Glu Tyr Asp Pro Ala Lys
450 455 460

Arg Ile Thr Leu Arg Glu Ala Leu Lys His Pro Phe Phe Asp Leu Leu
465 470 475 480

Lys Lys Ser Ile

<210> 22
<211> 1367
<212> ADN
<213> Homo sapiens

<400> 22

gggcaaggag ctgctggctg gacggcggca tgtccgacag cgagaagctc aacctggact 60

cgatcatcgg gcgcctgctg gaagtgcagg gctcgcggcc tggcaagaat gtacagctga 120

cagagaacga gatccgcggt ctgtgcctga aatcccggga gatttttctg agccagccca 180

ttcttctgga gctggaggca cccctcaaga tctgcggtga catacacggc cagtactacg 240

accttctgcg actatttgag tatggcggtt tccctcccga gagcaactac ctctttctgg 300

gggactatgt ggacaggggc aagcagtcct tggagaccat ctgcctgctg ctggcctata 360

agatcaagta ccccgagaac ttcttcctgc tccgtgggaa ccacgagtgt gccagcatca 420

accgcatcta tggtttctac gatgagtgca agagacgcta caacatcaaa ctgtggaaaa 480

ccttcactga ctgcttcaac tgcctgccca tcgcggccat agtggacgaa aagatcttct 540

gctgccacgg aggcctgtcc ccggacctgc agtctatgga gcagattcgg cggatcatgc 600

ggcccacaga tgtgcctgac cagggcctgc tgtgtgacct gctgtggtct gaccctgaca 660

aggacgtgca gggctggggc gagaacgacc gtggcgtctc ttttaccttt ggagccgagg 720

tggtggccaa gttcctccac aagcacgact tggacctcat ctgccgagca caccaggtgg 780

tagaagacgg ctatgagttc tttgccaagc ggcagctggt gacactttc tcagctccca 840

actactgtgg cgagtttgac aatgctggcg ccatgatgag tgtggacgag accctcatgt 900

gctctttcca gatcctcaag cccgccgaca agaacaaggg gaagtacggg cagttcagtg 960

gcctgaaccc tggaggccga cccatcaccc caccccgcaa ttccgccaaa gccaagaaat 1020

agccccccgca caccaccctg tgccccagat gatggattga ttgtacagaa atcatgctgc 1080

catgctgggg gggggtcacc ccgaccccta aggcccacct gtcacgggga acatggagcc 1140

ttggtgtatt tttcttttct ttttttaatg aatcaatagc agcgtccagt cccccagggc 1200

tgcttcctgc ctgcacctgc ggtactgtga gcaggatcct ggggccgagg ctgcagctca 1260

gggcaacggc aggccaggtc gtgggtctcc agccgtgctt ggcctcaggc tggcagcccg 1320

gatcctgggg caacccatct ggtctcttga ataaaggtca aagctgg 1367

<210> 23
<211> 4747
<212> ADN
<213> Homo sapiens

<400> 23

gttgcgctcc cgcgcgtgcg tgttgggatc gaatcgctgt ttccttccgc ttctcttcct 60

ctgtctcccc cccatatccg tgcgccgagc tgataaaggc gccattttgg aggggccgcg 120

ggagacgtgg tgccgctgcg ggctcgctct gccgtgcgct aggcttggtg ggaaggcctg 180

ttctcgagtc cgcgcttttc gtcaccgcca tgtcgggagg tggtgtgatt cgtggccccg 240

cagggaacaa cgattgccgc atctacgtgg gtaacttacc tccagacatc cgaaccaagg 300

acattgagga cgtgttctac aaatacggcg ctatccgcga catcgacctc aagaatcgcc 360

gcgggggacc gcccttcgcc ttcgttgagt tcgaggaccc gcgagacgcg gaagacgcgg 420

tgtatggtcg cgacggctat gattacgatg ggtaccgtct gcgggtggag tttcctcgaa 480

gcggccgtgg aacaggccga ggcggcggcg ggggtggagg tggcggagct ccccgaggtc 540

gctatggccc cccatccagg cggtctgaaa acagagtggt tgtctctgga ctgcctccaa 600

gtggaagttg gcaggattta aaggatcaca tgcgtgaagc aggtgatgta tgttatgctg 660

atgtttaccg agatggcact ggtgtcgtgg agtttgtacg gaaagaagat atgacctatg 720

cagttcgaaa actggataac actaagttta gatctcatga ggtaggttat acacgtattc 780

ttttctttga ccagaattgg atacagtggt cttaacagtg gaatttcaag gtaaggattc 840

aggcaaggtt gtccaagtaa attgccagat ttctggtttt agttacattg tattcattca 900

gcatgtctga agatagatga aagcttagat ctttcaatgg aaagttctgt ctatccaata 960

gggagaaact gcctacatcc gggttaaagt tgatgggccc agaagtccaa gttatggaag 1020

atctcgatct cgaagccgta gtcgtagcag aagccgtagc agaagcaaca gcaggagtcg 1080

```
cagttactcc ccaaggagaa gcagaggatc accacgctat tctccccgtc atagcagatc     1140
tcgctctcgt acataagatg attggtgaca ctttttgtag aacccatgtt gtatacagtt     1200
ttcctttatt cagtacaatc ttttcatttt ttaattcaaa ctgttttgtt cagaatgggc     1260
taaagtgttg aattgcattc ttgtaatatc cccttgctcc taacatctac attcccttcg     1320
tgtctttgat aaattgtatt ttaagtgatg tcatagacag gattgtttaa atttagttaa     1380
ctccatactc ttcagactgt gatattgtgt aaatgtctat ctgccctggt ttgtgtgaac     1440
tgggatgttg ggggtgtttg tggttatctt acctggggaa gttcttatgt ttatcttgct     1500
tttcatgtgt ctttctgtag acatatctga agagatggat taagaatgct ttggattaag     1560
gattgtggag cacatttcaa tcattttagg attgtcaaaa ggaggattga ggaggatcag     1620
atcaataatg gaggcaatgg tttggattgg agagggctca ctggatccca atccttggag     1680
ctggatcatt ggattcaaat cataatgtgg ataggatagg gaggatgaat taccaggatt     1740
catggagcgg gatcagatta ccaggaacat aggagtggat tcctgcccca accaaaccgc     1800
attcgtgtgg attttttttat tcaacttaat tggctattcc aaagattttt tttttcctat     1860
ttttgacgat tggagcccct aagatgcacg atggaattgt gttttgcgtt ttttggtaaa     1920
aggagcaaag cgaggacctg gagataaacg ctggagcaat ctccttggaa ggattcagca     1980
cgagtagatg gtaaacattt aaaggggaaa gggggggttt gtttaaaata gtaaatcagt     2040
aagtcacttc taaatttaaa gaaaacaaaa ttggagttga agaataagta ggtttccaat     2100
tggctattgc cgttttcttt gaaaaaataa acatttttta aaaaactatg catggttgtc     2160
ctttttcctc ttcatgtaag attctaactg ggtctatcag ttaatcttta aattgttaag     2220
taagataaga ttttgactct tgtgttaatg tgttagcaaa ttaaaagttc ttaaaaggca     2280
atctaatggt attagccatc ttttattgtt aattgtaaaa gtcttcaggg gaaagcaaaa     2340
ggggagaata aggcatttgt gtatgtaact tggtaaatga cggtggggga tggatctagc     2400
atctgaaaga taagcttctc tactttgtta taaagtggtt aaaaaactat agatgctgct     2460
tattttctgg tggtcataga caacataggc ttttgtgcaa aattggttga tggctactaa     2520
tgttcacttg gagatagctt ttgatattct caatgaaact catctcaaaa aaaggtaagt     2580
attaaatgtt aacatcagca cagatgtatt agaactgttt tttgtttttg agacagagtc     2640
tcgctctgtt ctccagactg gagggcagtg gtgtgatcta ggctcactgc aacctccacc     2700
cctggatttg agtgattctc gtgcctcagt ctcccaagta gctgagacta caagtgtgtg     2760
ccacccttgc ccggctaatt ttgtattttt agtagagatg tggtttctct gtgttaccca     2820
ggctggtccc aaaactcctg gcctcaagtg atctgcctgc cttggcctcc caaagtgtta     2880
ggattacagg tgtgagccac catgctcagc ctgtagaact tttaacccaa gtctcatttc     2940
ttttttgaaa gggaagagtg cacaagatta actgcttctt tggatgaatc attgttaata     3000
```

aaaagctggg catttagaat tttgccttat aagcccttct ccaaccataa gattattttg 3060

taccaaaaac tttggtgttc tctaccaaag cagttaaaaa cttttagcct gctacttctt 3120

gtatttgtct actgacagcc ccttggtact atttaggttg ggggagggga cctaaaataa 3180

atagacttta acatttccct tgggtgctaa tcatagttgg aagttgaatt taaggtgatt 3240

atttgggtga caattaaaaa cctaaggaaa accagaaatc ttggtagtgg aagaaatgtg 3300

taaggtcacc ccaatcggta gattttaatg aacgttgtgg aatgttggga agaggggatg 3360

ttaagttgaa tgcagaattt cactaagtac ttagtgtaag tttaaggatg tagctctttt 3420

tatctaagaa ttcaatgtaa tggccaaaag gcagatttac tgtttaaaaa tttgaataat 3480

tttacatgac attcttgaaa ttctaagaag ttttatgtgt agaacatttt aaaaattcat 3540

cagattatta aagggaaaat aaatgattaa tgataatttt ggaaggttaa tgtgagctag 3600

acttaagtaa actttggttc atttgtgttc attgaatgtt ttggaaatga ccaaaaaatg 3660

taaatggcct tcactcaagt ttgagtgttt aaagttgaaa gatgtgctct actaaaagtt 3720

atagtaattc taacctcaca ttgaaatgag acagtattcc ttgttataca ggctgaattt 3780

gaagattaga gaggatctaa tgtttactta ggtaaagggg cataggtttt gtagttaaga 3840

tgaccagaca gctaaaagct gtgatgggaa gtatggactg ctcctattta tagtctcaga 3900

aaatggacct ttaggtctct atccgtattg gcaattatta gaagaaagtt acacccttttt 3960

gaaactacaa aagctgtctt ggaatttccc cctcttctcc ctatttatgt ccccttagaa 4020

tattttaggg agcctataat tattttctaa ccaaggaaaa acttaagtct ctttaagaag 4080

caattacttt tcataacatc agattgaata acccaccttg ctgttcagcc cacatcctac 4140

tggaaacaaa aggtaagaaa cccattttct ggttcttgat tgtttgggtc tgaattttgt 4200

ttttaaaact aagctaagtt taatgttttt taaaatgctg tttggaatat gaatagattc 4260

cccgtaaaat gatttttcct aagttttatg ctttagtaaa tattcagtgc tcacgtctgt 4320

gcatcatagt gcttgcgttt aatatgattt attgtagaat ctcaactttt cttggtgttt 4380

gttgtctttg aaacattgtc ttggtcatta gggctggtgt tttcacattt ctgtggtcaa 4440

ggtggatttc ttatgtgtgc ctttttgctt actttgtata tgaattttgt aatttaaatt 4500

gcaagtaagt tatatatatg tatttaccat aaatagtatt aaaagatgag aaactgttag 4560

actgaagttc tgttgtaaca taaccattat ttccatcaca gtatgaagac tgcaaacgca 4620

gaaaacagat tacagtctct tatccatttt ttgaaatcca aaaactacga aaacaaaaga 4680

ttttctgttg ttgagctaat taaatgtgaa ccctgaccag aaaaaaaaaa aaaaaaaaaa 4740

aaaaaaa 4747

<210> 24
<211> 747
<212> ADN
<213> Mus musulus

<400> 24

```
atgtcgggag gtggtgtgat ccgtggcccg gcggggaaca acgactgccg catctacgtg       60
ggtaacctac ctccggatat ccgaaccaag gacatcgagg acgtgtttta caaatacggc      120
gccatccgcg acatcgacct gaagaaccgc cgcgggggac cgcccttcgc cttcgttgag      180
ttcgaggacc cgcgagacgc ggaagatgcg gtgtacggtc gcgacggcta cgactacgac      240
ggctaccggc tgcgggtaga gtttccccga agcggccgcg ggaccggccg aggcggcggc      300
gggggtggag gcggcggcgc cccgagaggc cgctatggcc cgccgtccag gcggtccgag      360
aacagagtgg ttgtctctgg actgcctccg agtggaagct ggcaggactt aaaggatcac      420
atgcgtgagg caggtgatgt atgttacgct gatgtttacc gagatggcac tggtgtcgtg      480
gagtttgtac ggaaagaaga tatgacgtat gcagttcgaa aactggataa cactaagttt      540
agatctcacg agggagaaac tgcctacatc cgggttaaag ttgatgggcc cagaagtcca      600
agttatggaa gatctcgatc tcgaagccgt agtcgtagca gaagccgtag cagaagcaac      660
agcaggagtc gcagttactc cccaaggaga agcagaggat caccacgcta ttctccccgt      720
catagcagat ctcgctctcg tacataa                                         747
```

<210> 25
<211> 248
<212> PRT
<213> Mus musulus

<400> 25

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                  10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
        35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

Gly Tyr Arg Leu Arg Val Glu Phe Pro Arg Ser Gly Arg Gly Thr Gly
                85                  90                  95
```

```
Arg Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145                 150                 155                 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
                165                 170                 175

Asn Thr Lys Phe Arg Ser His Glu Gly Glu Thr Ala Tyr Ile Arg Val
            180                 185                 190

Lys Val Asp Gly Pro Arg Ser Pro Ser Tyr Gly Arg Ser Arg Ser Arg
        195                 200                 205

Ser Arg Ser Arg Ser Arg Ser Arg Ser Arg Ser Asn Ser Arg Ser Arg
    210                 215                 220

Ser Tyr Ser Pro Arg Arg Ser Arg Gly Ser Pro Arg Tyr Ser Pro Arg
225                 230                 235                 240

His Ser Arg Ser Arg Ser Arg Thr
                245
```

<210> 26
<211> 747
<212> ADN
<213> Mus musculus

<400> 26

```
atgtcgggag gtggtgtgat ccgtggcccg gcggggaaca acgactgccg catctacgtg       60
ggtaacctac ctccggatat ccgaaccaag gacatcgagg acgtgtttta caaatacggc      120
gccatccgcg acatcgacct gaagaaccgc cgcgggggac cgcccttcgc cttcgttgag      180
ttcgaggacc cgcgagacgc ggaagatgcg gtgtacggtc gcgacggcta cgactacgac      240
ggctaccggc tgcgggtaga gtttccccga agcggccgcg ggaccggccg aggcggcggc      300
gggggtggag gcggcggcgc cccgagaggc cgctatggcc cgccgtccag gcggtccgag      360
aacagagtgg ttgtctctgg actgcctccg agtggaagct ggcaggactt aaaggatcac      420
atgcgtgagg caggtgatgt atgttacgct gatgtttacc gagatggcac tggtgtcgtg      480
gagtttgtac ggaaagaaga tatgacgtat gcagttcgaa aactggataa cactaagttt      540
```

agatctcacg agggagaaac tgcctacatc cgggttaaag ttgatgggcc cagaagtcca 600

agttatggaa gatctcgatc tcgaagccgt agtcgtagca gaagccgtag cagaagcaac 660

agcaggagtc gcagttactc cccaaggaga agcagaggat caccacgcta ttctccccgt 720

catagcagat ctcgctctcg tacataa 747

<210> 27
<211> 248
<212> PRT
<213> Mus musulus

<400> 27

Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1 5 10 15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
20 25 30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
35 40 45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
50 55 60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65 70 75 80

Gly Tyr Arg Leu Arg Val Glu Phe Pro Arg Ser Gly Arg Gly Thr Gly
85 90 95

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
100 105 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
130 135 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145 150 155 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
165 170 175

Asn Thr Lys Phe Arg Ser His Glu Gly Glu Thr Ala Tyr Ile Arg Val
180 185 190

```
Lys Val Asp Gly Pro Arg Ser Pro Ser Tyr Gly Arg Ser Arg Ser Arg
        195                 200                 205

Ser Arg Ser Arg Ser Arg Ser Arg Ser Arg Ser Asn Ser Arg Ser Arg
    210                 215                 220

Ser Tyr Ser Pro Arg Arg Ser Arg Gly Ser Pro Arg Tyr Ser Pro Arg
225                 230                 235                 240

His Ser Arg Ser Arg Ser Arg Thr
                245
```

<210> 28
<211> 606
<212> ADN
<213> Mus musulus

<400> 28

```
atgtcgggag gtggtgtgat ccgtggcccg gcggggaaca acgactgccg catctacgtg       60
ggtaacctac ctccggatat ccgaaccaag gacatcgagg acgtgtttta caaatacggc      120
gccatccgcg acatcgacct gaagaaccgc cgcgggggac cgcccttcgc cttcgttgag      180
ttcgaggacc cgcgagacgc ggaagatgcg gtgtacggtc gcgacggcta cgactacgac      240
ggctaccggc tgcgggtaga gtttccccga agcggccgcg ggaccggccg aggcggcggc      300
gggggtggag gcggcggcgc cccgagaggc cgctatggcc cgccgtccag gcggtccgag      360
aacagagtgg ttgtctctgg actgcctccg agtggaagct ggcaggactt aaaggatcac      420
atgcgtgagg caggtgatgt atgttacgct gatgtttacc gagatggcac tggtgtcgtg      480
gagtttgtac ggaaagaaga tatgacgtat gcagttcgaa aactggataa cactaagttt      540
agatctcacg aggtaggtta tacacttatt ctttttttg gccagaattg gatacagttt       600
tcttaa                                                                 606
```

<210> 29
<211> 201
<212> PRT
<213> Mus musulus

<400> 29

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                   10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
```

```
                35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

Gly Tyr Arg Leu Arg Val Glu Phe Pro Arg Ser Gly Arg Gly Thr Gly
                85                  90                  95

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145                 150                 155                 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
                165                 170                 175

Asn Thr Lys Phe Arg Ser His Glu Val Gly Tyr Thr Leu Ile Leu Phe
            180                 185                 190

Phe Gly Gln Asn Trp Ile Gln Phe Ser
        195                 200
```

<210> 30
<211> 330
<212> PRT
<213> Homo sapiens

<400> 30

```
  Met Ser Asp Ser Glu Lys Leu Asn Leu Asp Ser Ile Ile Gly Arg Leu
  1               5                  10                  15

  Leu Glu Val Gln Gly Ser Arg Pro Gly Lys Asn Val Gln Leu Thr Glu
              20                  25                  30

  Asn Glu Ile Arg Gly Leu Cys Leu Lys Ser Arg Glu Ile Phe Leu Ser
          35                  40                  45

  Gln Pro Ile Leu Leu Glu Leu Glu Ala Pro Leu Lys Ile Cys Gly Asp
```

```
    50                  55                  60

Ile His Gly Gln Tyr Tyr Asp Leu Leu Arg Leu Phe Glu Tyr Gly Gly
65                  70                  75                  80

Phe Pro Pro Glu Ser Asn Tyr Leu Phe Leu Gly Asp Tyr Val Asp Arg
                85                  90                  95

Gly Lys Gln Ser Leu Glu Thr Ile Cys Leu Leu Leu Ala Tyr Lys Ile
            100                 105                 110

Lys Tyr Pro Glu Asn Phe Phe Leu Leu Arg Gly Asn His Glu Cys Ala
        115                 120                 125

Ser Ile Asn Arg Ile Tyr Gly Phe Tyr Asp Glu Cys Lys Arg Arg Tyr
    130                 135                 140

Asn Ile Lys Leu Trp Lys Thr Phe Thr Asp Cys Phe Asn Cys Leu Pro
145                 150                 155                 160

Ile Ala Ala Ile Val Asp Glu Lys Ile Phe Cys Cys His Gly Gly Leu
                165                 170                 175

Ser Pro Asp Leu Gln Ser Met Glu Gln Ile Arg Arg Ile Met Arg Pro
            180                 185                 190

Thr Asp Val Pro Asp Gln Gly Leu Leu Cys Asp Leu Leu Trp Ser Asp
        195                 200                 205

Pro Asp Lys Asp Val Gln Gly Trp Gly Glu Asn Asp Arg Gly Val Ser
    210                 215                 220

Phe Thr Phe Gly Ala Glu Val Val Ala Lys Phe Leu His Lys His Asp
225                 230                 235                 240

Leu Asp Leu Ile Cys Arg Ala His Gln Val Val Glu Asp Gly Tyr Glu
                245                 250                 255

Phe Phe Ala Lys Arg Gln Leu Val Thr Leu Phe Ser Ala Pro Asn Tyr
            260                 265                 270

Cys Gly Glu Phe Asp Asn Ala Gly Ala Met Met Ser Val Asp Glu Thr
        275                 280                 285

Leu Met Cys Ser Phe Gln Ile Leu Lys Pro Ala Asp Lys Asn Lys Gly
    290                 295                 300

Lys Tyr Gly Gln Phe Ser Gly Leu Asn Pro Gly Gly Arg Pro Ile Thr
305                 310                 315                 320

Pro Pro Arg Asn Ser Ala Lys Ala Lys Lys
                325                 330
```

<210> 31
<211> 659
<212> ADN
<213> Homo sapiens

<400> 31

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60

catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 120

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggа 180

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 300

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 360

tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct 420

ccccccccctc cccaccccca attttgtatt tatttatttt ttaattattt tgtgcagcga 480

tgggggcggg gggggggggg gggccccccc caggcggggc ggggcggggc gaggggcggg 540

gcggggcgag gcggaaaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc 600

cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcg 659

<210> 32
<211> 724
<212> ADN
<213> Homo sapiens

<400> 32

acgcgttgac attgattatt gactagttat taatagtaat caattacggg gtcattagtt 60

catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga 120

ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca 180

atagggactt tccattgacg tcaatgggtg gactatttac ggtaaactgc ccacttggca 240

gtacatcaag tgtatcatat gccaagtacg ccccctattg acgtcaatga cggtaaatgg 300

cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc 360

tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt 420

ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt 480

ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg 540

acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata taagcagagc tctctggcta 600

actagagaac ccactgctta ctggcttatc gaaattaata cgactcacta tagggagacc 660

caagctggct agcgtttaaa cttaagcttg gtaccgagct cggatccact agtccagtgt 720

ggtg 724

<210> 33
<211> 278
<212> ADN
<213> Homo sapiens

<400> 33

tcgaggtgag ccccacgttc tgcttcactc tccccatctc ccccccctcc ccacccccaa 60

ttttgtattt atttattttt taattatttt gtgcagcgat gggggcgggg gggggggggg 120

ggcccccccc aggcgggggcg gggcggggcg aggggcgggg cggggcgagg cggaaaggtg 180

cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc 240

ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcg 278

<210> 34
<211> 401
<212> ADN
<213> Homo sapiens

<400> 34

ctgcagaggg ccctgcgtat gagtgcaagt gggttttagg accaggatga ggcggggtgg 60

gggtgcctac ctgacgaccg accccgaccc actggacaag cacccaaccc ccattcccca 120

aattgcgcat cccctatcag agagggggag gggaaacagg atgcggcgag gcgcgtgcgc 180

actgccagct tcagcaccgc ggacagtgcc ttcgcccccg cctggcggcg cgcgccaccg 240

ccgcctcagc actgaaggcg cgctgacgtc actcgccggt cccccgcaaa ctccccttcc 300

cggccacctt ggtcgcgtcc gcgccgccgc cggcccagcc ggaccgcacc acgcgaggcg 360

cgagataggg gggcacgggc gcgaccatct gcgctgcggc g 401

<210> 35
<211> 1675
<212> ADN
<213> Homo sapiens

<400> 35

gtctgcaagc agacctggca gcattgggct ggccgccccc cagggcctcc tcttcatgcc 60

cagtgaatga ctcaccttgg cacagacaca atgttcgggg tgggcacagt gcctgcttcc 120

cgccgcaccc cagccccccc caaatgcctt ccgagaagcc cattgagtag ggggcttgca 180

ttgcacccca gcctgacagc ctggcatctt gggataaaag cagcacagcc ccctaggggc 240

tgcccttgct gtgtggcgcc accggcggtg gagaacaagg ctctattcag cctgtgccca 300

ggaaagggga tcaggggatg cccaggcatg gacagtgggt ggcaggggag gagaggaggg 360

ctgtctgctt cccagaagtc caaggacaca aatgggtgag gggactgggc agggttctga 420

ccctgtggga ccagagtgga gggcgtagat ggacctgaag tctccaggga caacagggcc 480

caggtctcag gctcctagtt gggcccagtg gctccagcgt ttccaaaccc atccatcccc 540

agaggttctt cccatctctc caggctgatg tgtgggaact cgaggaaata aatctccagt 600

gggagacgga ggggtggcca gggaaacggg gcgctgcagg aataaagacg agccagcaca 660

gccagctcat gcgtaacggc tttgtggagc tgtcaaggcc tggtctctgg gagagaggca 720

cagggaggcc agacaaggaa ggggtgacct ggagggacag atccaggggc taaagtcctg 780

ataaggcaag agagtgccgg ccccctcttg ccctatcagg acctccactg ccacatagag 840

gccatgattg acccttagac aaagggctgg tgtccaatcc cagcccccag ccccagaact 900

ccagggaatg aatgggcaga gagcaggaat gtgggacatc tgtgttcaag ggaaggactc 960

caggagtctg ctgggaatga ggcctagtag gaaatgaggt ggcccttgag ggtacagaac 1020

aggttcattc ttcgccaaat tcccagcacc ttgcaggcac ttacagctga gtgagataat 1080

gcctgggtta tgaaatcaaa aagttggaaa gcaggtcaga ggtcatctgg tacagccctt 1140

ccttcccttt ttttttttct ttttttttg tgagacaagg tctctctctg ttgcccaggc 1200

tggagtggcg caaacacagc tcactgcagc ctcaacctac tgggctcaag caatcctcca 1260

gcctcagcct cccaaagtgc tgggattaca agcatgagcc accccactca gcccttcct 1320

tccttttaa ttgatgcata ataattgtaa gtattcatca tggtccaacc aacccttct 1380

tgacccacct tcctagagag agggtcctct tgattcagcg gtcagggccc cagacccatg 1440

gtctggctcc aggtaccacc tgcctcatgc aggagttggc gtgcccagga agctctgcct 1500

ctgggcacag tgacctcagt ggggtgaggg gagctctccc catagctggg ctgcggccca 1560

accccacccc ctcaggctat gccaggggt gttgccaggg gcacccgggc atcgccagtc 1620

tagcccactc cttcataaag ccctcgcatc ccaggagcga gcagagccag agcat 1675

<210> 36
<211> 260
<212> ADN
<213> Homo sapiens

<400> 36

gaatcctatg cttcgaacgc tgacgtcatc aacccgctcc aaggaatcgc gggcccagtg 60

tcactaggcg ggaacaccca gcgcgcgtgc gccctggcag gaagatggct gtgagggaca 120

ggggagtggc gccctgcaat atttgcatgt cgctatgtgt tctgggaaat caccataaac 180

gtgaaatgtc tttggatttg ggagtcttat aagttctgta tgagaccact ctttcccatg 240

gtcttcatcc tatctagaca 260

<210> 37
<211> 261
<212> ADN
<213> Homo sapiens

<400> 37

gccgccccct tcaccgaggg cctatttccc atgattcctt catatttgca tatacgatac 60

aaggctgtta gagagataat tggaattaat ttgactgtaa acacaaagat attagtacaa 120

aatacgtgac gtagaaagta ataatttctt gggtagtttg cagttttaaa attatgtttt 180

aaaatggact atcatatgct taccgtaact tgaaagtatt tcgatttctt ggctttatat 240

atcttgtgga aaggacgaaa c 261

<210> 38
<211> 4661
<212> ADN
<213> Adenovirus

<400> 38

aaagcttccc ggggggatct gggccactcc ctctctgcgc gctcgctcgc tcactgaggc 60

cgggcgacca aaggtcgccc gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg 120

agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctggaggg gtggagtcgt 180

gacccctaaa atgggcaaac attgcgctag cgcggccgcg gtacctaccg atgtacgggc 240

cagatatacg cgtggagcta gttattaata gtaatcaatt acggggtcat tagttcatag 300

cccatatatg gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc 360

caacgacccc cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg 420

gactttccat tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca 480

tcaagtgtat catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc 540

ctggcattat gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt 600

attagtcatc gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata 660

gcggtttgac tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt 720

ttggcaccaa aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca 780

aatgggcggt aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg 840

tcagatcgcc tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg 900

atccagcctg aattctgcag tcgacggtac cgcgggcccg ggatccaccg gtcgccacca 960

tggtgagcaa gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg 1020

gcgacgtaaa cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg 1080

gcaagctgac cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc 1140

tcgtgaccac cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc 1200

agcacgactt cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct 1260

tcaaggacga cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg 1320

tgaaccgcat cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca 1380

agctggagta caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg 1440

gcatcaaggt gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg 1500

accactacca gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact 1560

acctgagcac ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc 1620

tgctggagtt cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaaa 1680

gcggccgcga ctctagagtc gacctgcagg gaagaccaag ctgacctgac tcgatgcttt 1740

atttgtgaaa tttgtgatgc tattgcttta tttgtaacca ttataagctg caataaacaa 1800

gttaacaaca acaattgcat tcattttatg tttcaggttc agggggaggt gtgggaggtt 1860

ttttaaacta gtccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa 1920

aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgcagag 1980

agggacagat ccgggcccgc atgcgtcgac aattcactgg ccgtcgtttt acaacgtcgt 2040

gactgggaaa accctggcgt tacccaactt aatcgccttg cagcacatcc cctttcgcc 2100

agctggcgta atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg 2160

aatggcgaat ggcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac 2220

cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga 2280

cacccgccaa cacccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac 2340

agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg 2400

aaacgcgcga gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata 2460

ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt 2520

tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 2580

atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 2640

attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 2700

gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 2760

agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 2820

aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 2880

cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 2940

cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 3000

actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 3060

cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 3120

ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 3180

ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 3240

gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 3300

gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat 3360

ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 3420

cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 3480

caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 3540

taggtgaaga tccttttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 3600

cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg 3660

cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 3720

gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 3780

aatactgttc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 3840

cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 3900

tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 3960

acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 4020

ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 4080

ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 4140

tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 4200

tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 4260

ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg 4320

gataaccgta ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag 4380

cgcagcgagt cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc 4440

gcgcgttggc cgattcatta atgcagctgg cacgacaggt ttcccgactg gaaagcgggc 4500

agtgagcgca acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac 4560

tttatgcttc cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga 4620

aacagctatg accatgatta cgccaagctc tcgagatcta g 4661

<210> 39
<211> 7390
<212> ADN
<213> Adenovirus

<400> 39

atgccggggt tttacgagat tgtgattaag gtccccagcg accttgacgg gcatctgccc 60

ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat 120

tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180

cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg 240

caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac caccggggtg 300

aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt 360

taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc 420

gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctccccaaa 480

acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag cgcctgtttg 540

aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag 600

gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag atcaaaaact 660

tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac ctcggagaag 720

cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcgg 780

tcccaaatca aggtgccttg gacaatgcgg gaaagattat gagcctgact aaaaccgccc 840

ccgactacct ggtgggccag cagcccgtgg aggacatttc cagcaatcgg atttataaaa 900

ttttggaact aaacgggtac gatccccaat atgcggcttc cgtctttctg ggatgggcca 960

cgaaaaagtt cggcaagagg aacaccatct ggctgtttgg gcctgcaact accgggaaga 1020

ccaacatcgc ggaggccata gcccacactg tgcccttcta cgggtgcgta aactggacca 1080

atgagaactt tcccttcaac gactgtgtcg acaagatggt gatctggtgg gaggagggga 1140

agatgaccgc caaggtcgtg gagtcggcca aagccattct cggaggaagc aaggtgcgcg 1200

tggaccagaa atgcaagtcc tcggcccaga tagacccgac tcccgtgatc gtcacctcca 1260

acaccaacat gtgcgccgtg attgacggga actcaacgac cttcgaacac cagcagccgt 1320

tgcaagaccg gatgttcaaa tttgaactca cccgccgtct ggatcatgac tttgggaagg 1380

tcaccaagca ggaagtcaaa gacttttttcc ggtgggcaaa ggatcacgtg gttgaggtgg 1440

agcatgaatt ctacgtcaaa aagggtggag ccaagaaaag acccgccccc agtgacgcag 1500

atataagtga gcccaaacgg gtgcgcgagt cagttgcgca gccatcgacg tcagacgcgg 1560

aagcttcgat caactacgcg gacaggtacc aaaacaaatg ttctcgtcac gtgggcatga 1620

atctgatgct gtttccctgc agacaatgcg agagactgaa tcagaattca aatatctgct 1680

tcactcacgg tgtcaaagac tgtttagagt gctttcccgt gtcagaatct caacccgttt 1740

ctgtcgtcaa aaaggcgtat cagaaactgt gctacattca tcacatcatg ggaaaggtgc 1800

cagacgcttg cactgcttgc gacctggtca atgtggactt ggatgactgt gtttctgaac 1860

aataaatgac ttaaaccagg tatggctgcc gatggttatc ttccagattg gctcgaggac 1920

aaccttagtg aaggaattcg cgagtggtgg gctttgaaac ctggagcccc tcaacccaag 1980

gcaaatcaac aacatcaaga caacgctcga ggtcttgtgc ttccgggtta caaatacctt 2040

ggacccggca acggactcga caagggggag ccggtcaacg cagcagacgc ggcggccctc 2100

gagcacgaca aggcctacga ccagcagctc aaggccggag acaacccgta cctcaagtac 2160

aaccacgccg acgccgagtt ccaggagcgg ctcaaagaag atacgtcttt tgggggcaac 2220

ctcgggcgag cagtcttcca ggccaaaaag aggcttcttg aacctcttgg tctggttgag 2280

gaagcggcta agacggctcc tggaaagaag aggcctgtag agcagtctcc tcaggaaccg 2340

gactcctccg cgggtattgg caaatcgggt gcacagcccg ctaaaaagag actcaatttc 2400

ggtcagactg gcgacacaga gtcagtccca gaccctcaac caatcggaga acctcccgca 2460

gccccctcag gtgtgggatc tcttacaatg gcttcaggtg gtggcgcacc agtggcagac 2520

aataacgaag gtgccgatgg agtgggtagt tcctcgggaa attggcattg cgattcccaa 2580

tggctggggg acagagtcat caccaccagc acccgaacct gggccctgcc cacctacaac 2640

aatcacctct acaagcaaat ctccaacagc acatctggag gatcttcaaa tgacaacgcc 2700

tacttcggct acagcacccc ctgggggtat tttgacttca acagattcca ctgccacttc 2760

tcaccacgtg actggcagcg actcatcaac aacaactggg gattccggcc taagcgactc 2820

aacttcaagc tcttcaacat tcaggtcaaa gaggttacgg acaacaatgg agtcaagacc 2880

atcgccaata accttaccag cacggtccag gtcttcacgg actcagacta tcagctcccg 2940

tacgtgctcg ggtcggctca cgagggctgc ctcccgccgt tcccagcgga cgttttcatg 3000

attcctcagt acgggtatct gacgcttaat gatggaagcc aggccgtggg tcgttcgtcc 3060

ttttactgcc tggaatattt cccgtcgcaa atgctaagaa cgggtaacaa cttccagttc 3120

agctacgagt ttgagaacgt acctttccat agcagctacg ctcacagcca aagcctggac 3180

cgactaatga atccactcat cgaccaatac ttgtactatc tctcaaagac tattaacggt 3240

tctggacaga atcaacaaac gctaaaattc agtgtggccg gacccagcaa catggctgtc 3300

cagggaagaa actacatacc tggacccagc taccgacaac aacgtgtctc aaccactgtg 3360

actcaaaaca acaacagcga atttgcttgg cctggagctt cttcttgggc tctcaatgga 3420

cgtaatagct tgatgaatcc tggacctgct atggccagcc acaaagaagg agaggaccgt 3480

ttctttcctt tgtctggatc tttaattttt ggcaaacaag gaactggaag agacaacgtg 3540

gatgcggaca aagtcatgat aaccaacgaa gaagaaatta aaactactaa cccggtagca 3600

acggagtcct atggacaagt ggccacaaac caccagagtg cccaagcaca ggcgcagacc 3660

ggctgggttc aaaaccaagg aatacttccg ggtatggttt ggcaggacag agatgtgtac 3720

ctgcaaggac ccatttgggc caaaattcct cacacggacg gcaactttca cccttctccg 3780

ctgatgggag ggtttggaat gaagcacccg cctcctcaga tcctcatcaa aaacacacct 3840

gtacctgcgg atcctccaac ggccttcaac aaggacaagc tgaactcttt catcacccag 3900

tattctactg gccaagtcag cgtggagatc gagtgggagc tgcagaagga aaacagcaag 3960

cgctggaacc cggagatcca gtacacttcc aactattaca agtctaataa tgttgaattt 4020

gctgttaata ctgaaggtgt atatagtgaa ccccgcccca ttggcaccag atacctgact 4080

cgtaatctgt aattgcttgt taatcaataa accgtttaat tcgtttcagt tgaactttgg 4140

tctctgcgaa gggcgaattc gtttaaacct gcaggactag aggtcctgta ttagaggtca 4200

cgtgagtgtt ttgcgacatt ttgcgacacc atgtggtcac gctgggtatt taagcccgag 4260

tgagcacgca gggtctccat tttgaagcgg gaggtttgaa cgcgcagccg ccaagccgaa 4320

ttctgcagat atccatcaca ctggcggccg ctcgactaga gcggccgcca ccgcggtgga 4380

gctccagctt ttgttccctt tagtgagggt taattgcgcg cttggcgtaa tcatggtcat 4440

agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 4500

gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 4560

gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 4620

aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact 4680

cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 4740

ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 4800

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg 4860

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 4920

gataccaggc gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 4980

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 5040

gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 5100

cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 5160

taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 5220

atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 5280

cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 5340

cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 5400

ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 5460

ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 5520

tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 5580

aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 5640

tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg 5700

gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag 5760

atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt 5820

tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag 5880

ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt 5940

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 6000

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 6060

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 6120

ccgtaagatg ctttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 6180

gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag 6240

aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 6300

accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 6360

ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 6420

gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg 6480

aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6540

taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctaaat tgtaagcgtt 6600

aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 6660

gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 6720

gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 6780

aaaaccgtct atcagggcga tggcccacta cgtgaaccat caccctaatc aagtttttg 6840

gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagcccccg atttagagct 6900

tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc 6960

gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt 7020

aatgcgccgc tacagggcgc gtcccattcg ccattcaggc tgcgcaactg ttgggaaggg 7080

cgatcggtgc gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg 7140

cgattaagtt gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt 7200

gagcgcgcgt aatacgactc actatagggc gaattgggta ccgggccccc cctcgatcga 7260

ggtcgacggt atcgggggag ctcggatcca ctagtaacgg ccgccagtgt gctggattcg 7320

gctttattta agcccgagtg agcacgcagg gtctccattt tgaagcggga ggtttgaacg 7380

cgcagccgcc 7390

<210> 40
<211> 479
<212> ADN
<213> secuencia artificial

<220>
<223> repeticiones de RAN-G4C2x38-sentido

<400> 40

cttatcgaaa ttaatacgac tcactatagg gagacccaag ctggctagtt aagcttggta 60

ccgagctcgg atccactagt ccagtgtggt ggaattgggg cccggggccg gggccggggc 120

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 180

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 240

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 300

cggggccggg gccggggccg gggccggggc cggggaattc tgcagatatc cagcacagtg 360

gcggccgctc gagtctagag ggcccttcga acaaaaactc atctcagaag aggatctgaa 420

tatgcatacc ggtcatcatc accatcacca ttgagtttaa acccgctgat cagcctcga 479

<210> 41
<211> 478
<212> ADN
<213> secuencia artificial

<220>
<223> repeticiones de RAN-C4G2x39-antisentido:

<400> 41

cttatcgaaa ttaatacgac tcactatagg gagacccaag ctggctagtt aagcttggta 60

ccgagctcgg atccactagt ccagtgtggt ggaattcccc ggccccggcc ccggccccgg 120

ccccggcccc ggccccggcc ccggccccgg ccccggcccc ggccccggcc ccggccccgg 180

ccccggcccc ggccccggcc ccggccccgg ccccggcccc ggccccggcc ccggccccgg 240

ccccggcccc ggccccggcc ccggccccgg ccccggcccc ggccccggcc ccggccccgg 300

ccccggcccc ggccccggcc ccggccccgg ccccaattct gcagatatcc agcacagtgg 360

cggccgctcg agtctagagg gcccttcgaa caaaaactca tctcagaaga ggatctgaat 420

atgcataccg gtcatcatca ccatcaccat tgagtttaaa cccgctgatc agcctcga 478

<210> 42
<211> 285
<212> ADN
<213> secuencia artificial

<220>
<223> DPR poli-Gly-Pro x36 independientes de repeticiones de G4C2

<400> 42

atgggcaaac cgattccgaa cccgctgctg ggcctggata gcaccctcga gaatgatccc 60

accatgggcc ctggccctgg accaggacct ggccccggac ccggtccagg tcccggccca 120

ggccccggtc ccggccctgg accaggccca ggaccaggac caggcccagg tcccggacca 180

ggacccggac ctggcccagg ccctggccct ggccctggcc ccggaccagg ccctggaccc 240

ggccctggtc ccggcccagg acccggacca ggacctggcc cttaa 285

<210> 43
<211> 285
<212> ADN
<213> secuencia artificial

<220>
<223> poli-Gly-Ala x36 DPR independientes de repeticiones de G4C2

<400> 43

atgggcaaac cgattccgaa cccgctgctg ggcctggata gcaccctcga gaatgatccc 60

accatgggag ctggtgctgg tgcaggcgct ggcgcagggg caggcgctgg tgctggggct 120

ggtgccgggg ctggggcagg cgcaggggct ggtgccggtg caggcgcagg ggctggggct 180

ggcgctggtg ccggcgcagg cgcgggtgcc ggcgcagggg ctggtgcagg ggccggtgct 240

ggcgcgggtg caggggccgg tgcaggggca ggcgcaggcg cttaa 285

<210> 44
<211> 634
<212> ADN
<213> secuencia artificial

<220>
<223> repeticiones de RAN-G4C2x38 sentido con etiquetas 3x V5

<400> 44

cttatcgaaa ttaatacgac tcactatagg gagacccaag ctggctagtt aagcttggta 60

ccgagctcgg atccactagt ccagtgtggt ggaattgggg cccggggccg gggccggggc 120

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 180

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 240

cggggccggg gccggggccg gggccggggc cggggccggg gccggggccg gggccggggc 300

cggggccggg gccggggccg gggccggggc cggggaattc tgcagatatc cagcacagtg 360

gcggccgctg ggcaagccca tccccaaccc cctgctcggt ctggacagca ccggctaacg 420

gcaagcccat ccccaacccc ctgctcggtc tggacagcac cggctaacgg caagcccatc 480

cccaaccccc tgctcggtct ggacagcacc ggctaactcg agtctagatc tagagggccc 540

ttcgaacaaa aactcatctc agaagaggat ctgaatatgc ataccggtca tcatcaccat 600

caccattgag tttaaacccg ctgatcagcc tcga 634

<210> 45
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 45
ttctgggccc atcaacttta a 21

<210> 46
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 46
agcacttgga gtcataccat t 21

<210> 47
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 47
gggatcgaat cgctgtttcc t 21

<210> 48
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 48
gctgtttcct tccgcttctc t 21

<210> 49
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 49
gggcccagaa gtccaagtta t 21

<210> 50
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 50
gctgatgttt accgagatgg c 21

<210> 51
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 51
aatgtctatt ctgctctggt t 21

<210> 52
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 52
aaattgcaga tgggagcaat a 21

<210> 53
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 53
ttgcagatgg gagcaatagt t 21

<210> 54
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 54
tgggagcaat agtttaggtt t 21

<210> 55
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 55
ttaggtttag gtgggtagta a 21

<210> 56
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 56
taggtttagg tgggtagtaa t 21

<210> 57
<211> 359
<212> ADN
<213> secuencia artificial

<220>
<223> secuencia de repetición invertida

<400> 57

atgccgacga tgcggtgaag gcgcgcgacg gctacgacta cgatgggtat cgtctgcgcg 60

tggagttccc gcggggcggt ggtcctggaa gctaccgcgg cggcaaccgc aatgaccgaa 120

gccgcgacgg tgggggacgg atgggcggac gcggaccgcc agccaagcgc tcgcagtacc 180

gcgtcatggt tactggactg cccgcctccg gatcgtggca agatctcaag gatcacatgc 240

gcgaggccgg cgacgtctgc ttcgcggaca cttacaagga tggttccggc gtcgttgagt 300

tcctgcgcca cgaggacatg aagtacgcaa tcaaaaaatt ggacgactct cgcttccga 359

<210> 58
<211> 381
<212> ADN
<213> secuencia artificial

<220>
<223> ID de FlyBase = FBst0450381

<400> 58

ggtccgataa agaaggcggc agtgcactac gatcgctccg gtcgctcgtt gggcaccgct 60

gacgtgattt tcgaacgtcg cgccgacgcc ttgaaggcca ttaaacagta ccatggcgta 120

cctttggacg gacgccctat gaccattcag ctggccgtct cagacgtggc cgtgttgacc 180

cgtcccgtag ccgccaccga tgtcaagcgt cgcgtgggtg gtactgcacc aacttcattc 240

aagcgtggtg gtggccaagc tggtggcacg gcgcgtcgcg gcttcaaacg tccggtcggt 300

ggcaagccgg cggcaggcgg ccagcgacgg gagcgcaagg ccccgcccac tgctgaggag 360

ctggacgccg aactggactc a 381

<210> 59
<211> 324
<212> ADN
<213> secuencia artificial

<220>
<223> ID de FlyBase = FBst0476329

<400> 59

gtcgaacttg ataaagcgca tttctaaata caataaatac agcatcaaat gtatttcagt 60

tatcttaaca tccgccgcat tggcaaaact aacaattaat ggataaatgc gcaagtggtt 120

gattgatttg atgtccgatg ctttcaaaga tctgctcctg ggcgcggcgt tgtcgatgcg 180

tttgcattta tgtaccatgc gggggggtgtc catatggtag gcttaaaact atagattggg 240

ctgctcttct attcttgtta gactaattca gactattcac tatttagatc ttcatgtcgt 300

tgatgtatga gtccagttcg gcgt 324

<210> 60
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 60
gcatctgaaa gataagcttc t 21

<210> 61
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 de humano/ratón - miR1 - Cadena superior

<400> 61

tgctgttctg ggcccatcaa ctttaagttt tggccactga ctgacttaaa gtttgggccc 60

agaa 64

<210> 62
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 de humano/ratón/rata - miR1 - Cadena inferior:

<400> 62

cctgttctgg gcccaaactt taagtcagtc agtggccaaa acttaaagtt gatgggccca 60

gaac 64

<210> 63
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 de humano - miR2 - Cadena superior

<400> 63

tgctgagcac ttggagtcat accattgttt tggccactga ctgacaatgg tatctccaag 60

tgct 64

<210> 64
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 de humano - miR2 - Cadena inferior

<400> 64

cctgagcact tggagatacc attgtcagtc agtggccaaa acaatggtat gactccaagt 60

gctc 64

<210> 65
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> Secuencia de miR2 dirigida a SRSF1 de ratón

<400> 65
aatgtctatt ctgctctggt t 21

<210> 66
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> Secuencia de miR2 dirigida a SRSF1 de ratón

<400> 66
aatgtctatt ctgctctggt t 21

<210> 67
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 - miR2 de ratón - Cadena superior

<400> 67

tgctgaacca gagcagaata gacattgttt tggccactga ctgacaatgt ctactgctct 60

ggtt 64

<210> 68
<211> 64
<212> ADN
<213> secuencia artificial

<220>
<223> SRSF1 - miR2 de ratón - Cadena inferior

<400> 68

cctgaaccag agcagtagac attgtcagtc agtggccaaa acaatgtcta ttctgctctg 60

gttc 64

<210> 69
<211> 6
<212> ADN
<213> secuencia artificial

<220>
<223> hexanucleótidos GGGGCCx38 sentido

<400> 69
ggggcc 6

<210> 70
<211> 6
<212> ADN
<213> secuencia artificial

<220>
<223> hexanucleótido CCCCGGx39 antisentido

<400> 70
ccccgg 6

<210> 71
<211> 12
<212> ADN
<213> secuencia artificial

<220>
<223> oligonucleótidos sintéticos

<400> 71
ccccggcccc gg 12

<210> 72
<211> 45
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 72
actctagagg taccacgtga tcattctcga gggtgctatc caggc 45

<210> 73
<211> 45
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 73
gcctggatag caccctcgag aatgatcacg tggtacctct agagt 45

<210> 74
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 74
taccgcgtca tggttactgg 20

<210> 75
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 75
gtacgcgaat gtaggcaacc 20

<210> 76
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 76
cgatatgtac gacggaccga a 21

<210> 77
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 77
cggaccaaag tcgttgaaga g 21

<210> 78
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 78
tgggcccgtc tggaccacaa 20

<210> 79
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 79
tcgccgtcac cggagtccat 20

<210> 80
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 80
ttccaaccta tggaactgat ga 22

<210> 81
<211> 23
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 81
ggttttcctc attaaaggca ttc 23

<210> 82
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 82
ccgcatctac gtgggtaact 20

<210> 83
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 83
tcgaactcaa cgaaggcgaa 20

<210> 84
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 84
tctggtcgca gcttaggaac 20

<210> 85
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 85
ccacctctgt ttacgctctg t 21

<210> 86
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 86
ccatgatcac gaaggtggtt 20

<210> 87
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 87
atgcagtcga gtttcccaca 20

<210> 88
<211> 31
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 88
cttgtgaaac aaaatgcttt ttaacatcca t 31

<210> 89
<211> 25
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 89
gaatgtgagc accttccttc ttttt 25

<210> 90
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 90
gaactgcaca bccagaacac 20

<210> 91
<211> 19
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 91
tgggttgaag ttgctgagg 19

<210> 92
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 92
gggcccttcg aacccccgtc 20

<210> 93
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 93
gggaggggca aacaacagat 20

<210> 94
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 94
tcaaacagcg acaagttccg 20

<210> 95
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 95
gtcgacatga ctgcattcca 20

<210> 96
<211> 20
<212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 96
ggagagaggg tgggaaaaac 20

<210> 97
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 97
gcatctacgt gggtaactta c 21

<210> 98
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 98
ggaacaacga ttgccgcatc t 21

<210> 99
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 99
ggagtttgta cggaaagaag a 21

<210> 100
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 100
ggaagttggc aggatttaaa g 21

<210> 101
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 101
ggtaggttat acacgtattc t 21

<210> 102
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 102
gcaggtgatg tatgttatgc t 21

<210> 103
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 103
gctgatgttt accgagatgg c 21

<210> 104
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 104
gaagttggca ggatttaaag g 21

<210> 105
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 105
gaaagaagat atgacctatg c 21

<210> 106
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 106
gcgtgaagca ggtgatgtat g 21

<210> 107
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 107
gtaagttacc cacgtagatg c 21

<210> 108
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 108
taagtta 7

<210> 109
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 109
agatgcggca atcgttgttc c 21

<210> 110
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 110
gatgcgg 7

<210> 111
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 111
tcttctttcc gtacaaactc c 21

<210> 112
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 112
cttcttt 7

<210> 113
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 113
ctttaaatcc tgccaacttc c 21

<210> 114
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 114
tttaaat 7

<210> 115
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 115
agaatacgtg tataacctac c 21

<210> 116
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 116
gaatacg 7

<210> 117
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 117
agcataacat acatcacctg c 21

<210> 118
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 118
gcataac 7

<210> 119
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 119
gccatctcgg taaacatcag c 21

<210> 120
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 120
ccatctc 7

<210> 121
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 121
cctttaaatc ctgccaactt c 21

<210> 122
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 122
ctttaaa 7

<210> 123
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 123
gcataggtca tatcttcttt c 21

<210> 124
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 124
cataggt 7

<210> 125
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 125
catacatcac ctgcttcacg c 21

<210> 126
<211> 7
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 126
atacatc 7

<210> 127
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 127
gcgctatccg cgacatcgac ct 22

<210> 128
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 128
ggucgau 7

<210> 129
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 129
tggccccgca gggaacaacg at 22

<210> 130
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 130
ucguugu 7

<210> 131
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 131
tcgcgacggc tatgattacg at 22

<210> 132
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 132
ucguaau 7

<210> 133
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 133
cggcgctatc cgcgacatcg ac 22

<210> 134
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 134
ucgaugu 7

<210> 135
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 135
agggaacaac gattgccgca tc 22

<210> 136
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 136
augcggc 7

<210> 137
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 137
gcggtgtatg gtcgcgacgg ct 22

<210> 138
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 138
gccgucg 7

<210> 139
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 139
ctgcgggtgg agtttcctcg aa 22

<210> 140
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 140
ucgagga 7

<210> 141
<211> 22
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 141
gtggagtttc ctcgaagcgg cc 22

<210> 142
<211> 7
<212> ARN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 142
gccgcuu 7

<210> 143
<211> 214
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 143

gaacgctgac gtcatcaacc cgctccaagg aatcgcgggc ccagtgtcac taggcgggaa 60

cacccagcgc ggtgcgccct ggcaggaaga tggctgtgag ggacagggga gtggcgccct 120

gcaatatttg catgtcgcta tgtgttctgg gaaatcacca taaacgtgaa atgtctttgg 180

atttgggaat cttataagtt ctgtatgaga ccac 214

<210> 144
<211> 91
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 144

atatttgcat gtcgctatgt gttctgggaa atcaccataa acgtgaaatg tctttggatt 60

tgggaatctt ataagttctg tatgagacca c 91

<210> 145
<211> 241
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 145

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60

ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180

atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

c 241

<210> 146
<211> 278
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 146

tcgaggtgag ccccacgttc tgcttcactc tccccatctc ccccccctcc ccacccccaa 60

ttttgtattt atttattttt taattatttt gtgcagcgat gggggcgggg gggggggggg 120

ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg 180

cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc 240

ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcg 278

<210> 147
<211> 659
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 147

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60

catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 120

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 180

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 300

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 360

tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct 420

ccccccccctc cccaccccca attttgtatt tatttatttt ttaattattt tgtgcagcga 480

tgggggcggg gggggggggg gggccccccc caggcggggc ggggcggggc gaggggcggg 540

gcggggcgag gcggaaaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc 600

cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcg 659

<210> 148
<211> 500
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 148

gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg 60

gcacttggcg ctacacaagt ggcctctggc ctcgcacaca ttccacatcc accggtaggc 120

gccaaccggc tccgttcttt ggtggcccct tcgcgccacc ttctactcct cccctagtca 180

ggaagttccc ccccgccccg cagctcgcgt cgtgcaggac gtgacaaatg gaagtagcac 240

gtctcactag tctcgtgcag atggacagca ccgctgagca atggaagcgg gtaggccttt 300

ggggcagcgg ccaatagcag ctttgctcct tcgctttctg ggctcagagg ctgggaaggg 360

gtgggtccgg gggcgggctc aggggcgggc tcaggggcgg ggcgggcgcc cgaaggtcct 420

ccggaggccc ggcattctgc acgcttcaaa agcgcacgtc tgccgcgctg ttctcctctt 480

cctcatctcc gggcctttcg 500

<210> 149

<211> 212
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 149

gggcagagcg cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgat 60

ccggtgccta gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc 120

gccttttttcc cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc 180

tttttcgcaa cgggtttgcc gccagaacac ag 212

<210> 150
<211> 380
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 150

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60

catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 120

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 180

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240

aagtgtatca tatgccaagt acgccccta ttgacgtcaa tgacggtaaa tggcccgcct 300

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 360

tagtcatcgc tattaccatg 380

<210> 151
<211> 204
<212> ADN
<213> secuencia artificial

<220>
<223> promotor

<400> 151

gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 60

ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac 120

tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg 180

tgggaggtct atataagcag agct 204

<210> 152
<211> 248
<212> PRT
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 152

```
Met Ser Gly Gly Gly Val Ile Arg Gly Pro Ala Gly Asn Asn Asp Cys
1               5                  10                  15

Arg Ile Tyr Val Gly Asn Leu Pro Pro Asp Ile Arg Thr Lys Asp Ile
            20                  25                  30

Glu Asp Val Phe Tyr Lys Tyr Gly Ala Ile Arg Asp Ile Asp Leu Lys
        35                  40                  45

Asn Arg Arg Gly Gly Pro Pro Phe Ala Phe Val Glu Phe Glu Asp Pro
    50                  55                  60

Arg Asp Ala Glu Asp Ala Val Tyr Gly Arg Asp Gly Tyr Asp Tyr Asp
65                  70                  75                  80

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

Arg Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr
            100                 105                 110

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

Pro Pro Ser Gly Ser Trp Gln Asp Leu Lys Asp His Met Arg Glu Ala
    130                 135                 140

Gly Asp Val Cys Tyr Ala Asp Val Tyr Arg Asp Gly Thr Gly Val Val
145                 150                 155                 160

Glu Phe Val Arg Lys Glu Asp Met Thr Tyr Ala Val Arg Lys Leu Asp
                165                 170                 175

Asn Thr Lys Phe Arg Ser His Glu Gly Glu Thr Ala Tyr Ile Arg Val
            180                 185                 190

Lys Val Asp Gly Pro Arg Ser Pro Ser Tyr Gly Arg Ser Arg Ser Arg
        195                 200                 205

Ser Arg Ser Arg Ser Arg Ser Arg Ser Arg Ser Asn Ser Arg Ser Arg
    210                 215                 220

Ser Tyr Ser Pro Arg Arg Ser Arg Gly Ser Pro Arg Tyr Ser Pro Arg
225                 230                 235                 240

His Ser Arg Ser Arg Ser Arg Thr
                245
```

<210> 153
<211> 61
<212> PRT
<213> secuencia artificial

<220>
<223> péptido antagonista

<400> 153

```
Pro Arg Ser Gly Arg Gly Thr Gly Arg Gly Gly Gly Gly Gly Gly Gly
1               5                  10                  15

Gly Gly Ala Pro Arg Gly Arg Tyr Gly Pro Pro Ser Arg Arg Ser Glu
            20                  25                  30

Gly Gly Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr
        35                  40                  45

Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
    50                  55                  60
```

<210> 154
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 154
ggaaagaaga tatgacctat g 21

<210> 155
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 155
agggagaaac tgcctacatc c 21

<210> 156
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 156
gggactaatg tgggaagaac t 21

<210> 157
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 157
gcaaccacga aacctgtaat a 21

<210> 158
<211> 21
<212> ADN
<213> secuencia artificial

<220>
<223> ARN inhibidor

<400> 158
gggatcagat taccaggaac a 21

REIVINDICACIONES

1. Un vector de expresión que comprende un casete de transcripción para su uso en el tratamiento de una enfermedad neurodegenerativa causada por la expansión polimórfica de G4C2 en el primer intrón del gen de C9ORF72 en donde dicho vector de expresión comprende:

    i) un ácido nucleico antisentido, un ARN o ARNhc inhibidor que es complementario a e inhibe la expresión de un ácido nucleico que codifica un factor de corte y empalme rico en serina/arginina [SRSF1];
    ii) una molécula de ácido nucleico que codifica un péptido que comprende o consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 15; o
    iii) una molécula de ácido nucleico que codifica una proteína negativa dominante que comprende o consiste en una secuencia de aminoácidos como se establece en la SEQ ID NO: 16 o 17;

    en donde dicha molécula de ácido nucleico está unida operativamente a un promotor adaptado para expresar el agente de acuerdo con i), ii) o iii).

2. El vector de acuerdo con el uso de la reivindicación 1, en donde dicho ARN inhibidor, ácido nucleico antisentido o ARNhc, comprende o consiste en una secuencia de nucleótidos como se establece en la SEQ ID NO: 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14 o 47, 48, 50, 99, 154, 155, 156, 157, 158, 60, 97, 98, 100, 101, 102, 103, 104, 105, 106, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139 o 141.

3. El vector de acuerdo con el uso de cualquiera de las reivindicaciones 1 a 2, en donde dicho vector de expresión es un vector de expresión de base viral.

4. El vector de expresión de acuerdo con el uso de la reivindicación 3, en donde dicho vector de base viral es un virus adenoasociado [AAV].

5. El vector de expresión de acuerdo con el uso de la reivindicación 4, en donde dicho vector de base viral es AAV9.

6. El vector de expresión de acuerdo con el uso de la reivindicación 3, en donde dicho vector de base viral es un lentivirus.

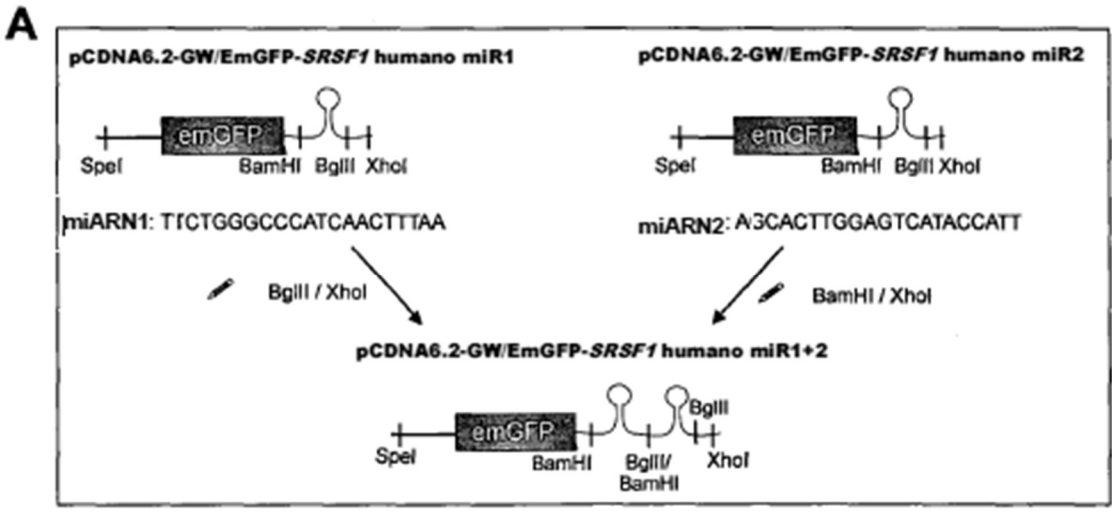
7. El vector de acuerdo con el uso de cualquiera de las reivindicaciones 1 a 6, en donde dicha enfermedad neurodegenerativa se selecciona del grupo que consiste en: esclerosis lateral amiotrófica, demencia frontotemporal, enfermedad de la neurona motora, demencia lobular frontotemporal, trastorno similar a Huntington, esclerosis lateral primaria, atrofia muscular progresiva, síndrome corticobasal, enfermedad de Alzheimer y demencia con cuerpos de Lewy.

8. El vector de acuerdo con el uso de la reivindicación 7, en donde dicha enfermedad neurodegenerativa es la enfermedad de la neurona motora.

9. El vector de acuerdo con el uso de la reivindicación 7, en donde dicha enfermedad neurodegenerativa es la esclerosis lateral amiotrófica.

10. El vector de acuerdo con el uso de la reivindicación 7, en donde dicha enfermedad neurodegenerativa es la demencia frontotemporal.

Figura 1


A
pCDNA6.2-GW/EmGFP-SRSF1 humano miR1
pCDNA6.2-GW/EmGFP-SRSF1 humano miR2
emGFP
SpeI
BamHI
BglII
XhoI
miARN1: TTCTGGGCCCATCAACTTTAA
miARN2: AGCACTTGGAGTCATACCATT
BglII / XhoI
BamHI / XhoI
pCDNA6.2-GW/EmGFP-SRSF1 humano miR1+2
BglII/ BamHI

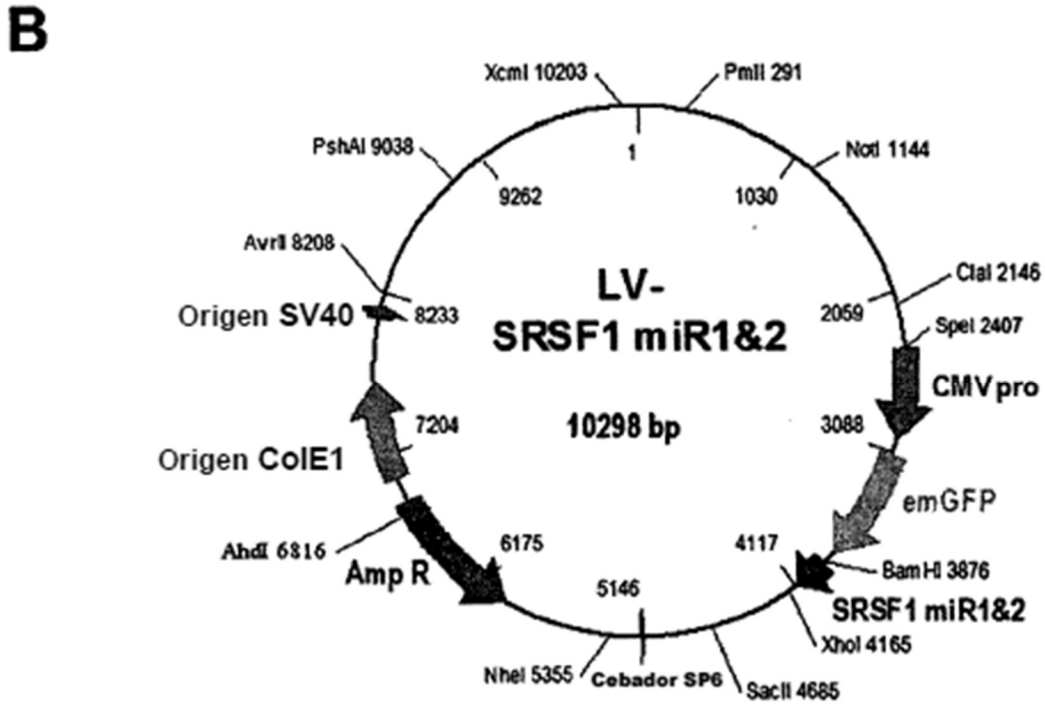
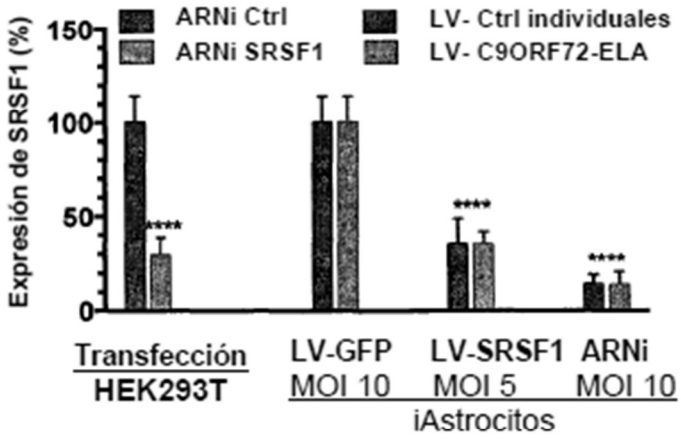

B
XcmI 10203
PmlI 291
PshAI 9038
NotI 1144
1
9262
1030
AvrII 8208
ClaI 2146
LV-
SRSF1 miR1&2
Origen SV40
8233
2059
SpeI 2407
CMV pro
10298 bp
7204
3088
Origen ColE1
emGFP
AhdI 6816
6175
4117
BamHI 3876
Amp R
5146
SRSF1 miR1&2
XhoI 4165
NheI 5355
Cebador SP6
SacII 4685

Figura 2

A


ARNi Ctrl
ARNi SRSF1
LV- Ctrl individuales
LV- C9ORF72-ELA
Expresión de SRSF1 (%)
150
100
50
0
****
****
****
Transfección
HEK293T
LV-GFP
MOI 10
LV-SRSF1
MOI 5
ARNi
MOI 10
iAstrocitos

B

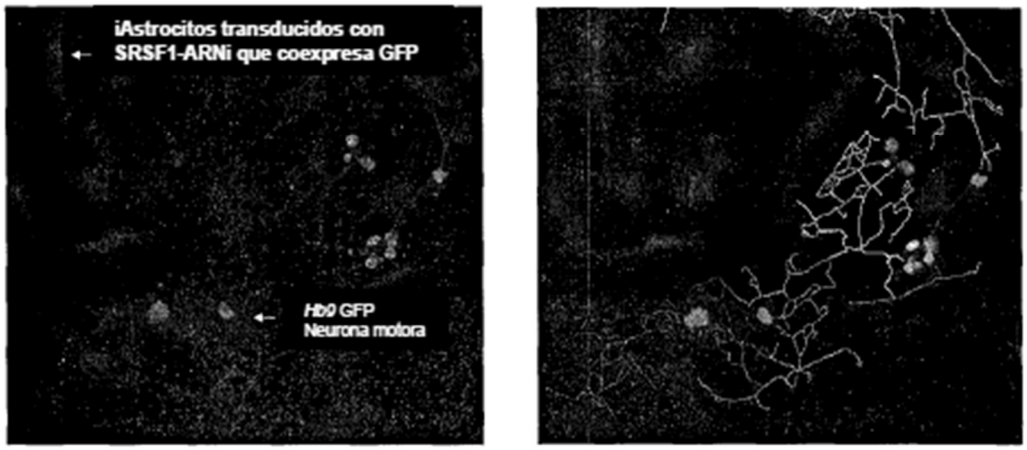

iAstrocitos transducidos con
SRSF1-ARNi que coexpresa GFP
Hb9 GFP
Neurona motora

Figura 3

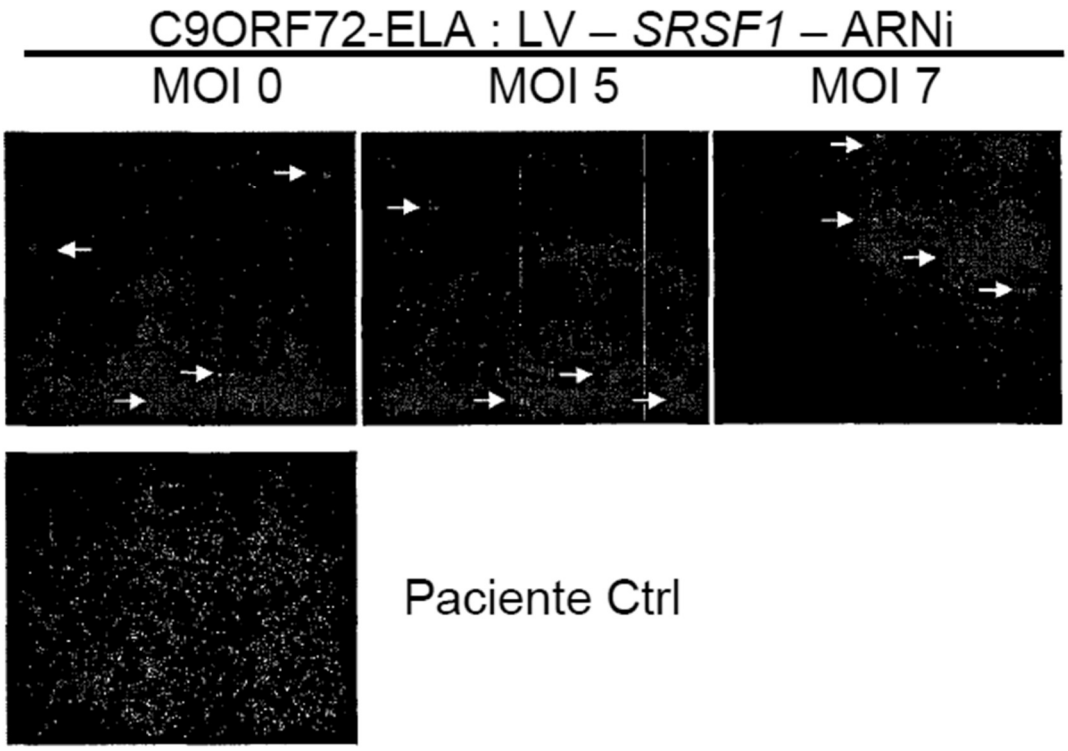
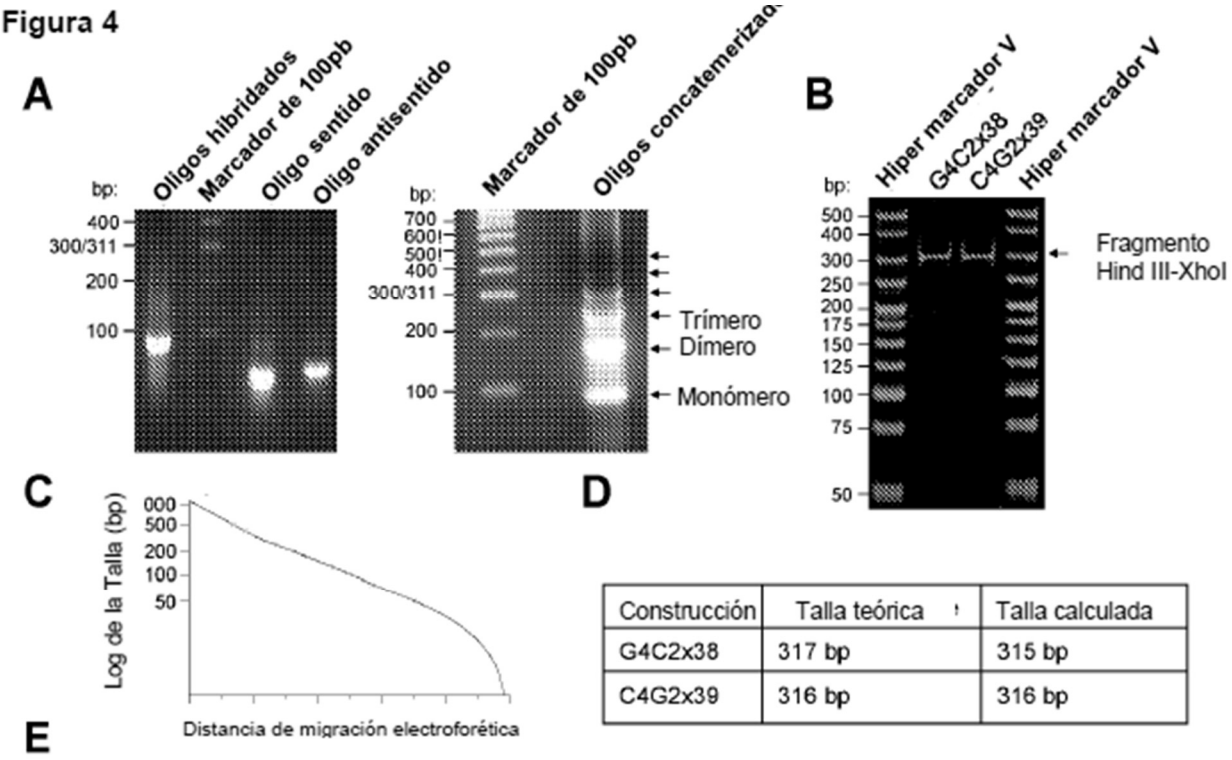


Figura 4



Construcción	Talla teórica	Talla calculada
G4C2x38	317 bp	315 bp
C4G2x39	316 bp	316 bp

E

RAN-G4C2x38 repeticiones-sentido:

...TGGCTT ATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCACTA
GTCCAGTGTGGTGGAATTGGGGCCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGG
GGAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATCTG
AATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCT CGACTG...

RAN-C4G2x39 repeticiones-antisentido:

...TGGCTT ATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCACTA
GTCCAGTGTGGTGGAATTCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGC
CCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCC
CCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCC
CAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATCTGA
ATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCT CGACTG...

Figura 5

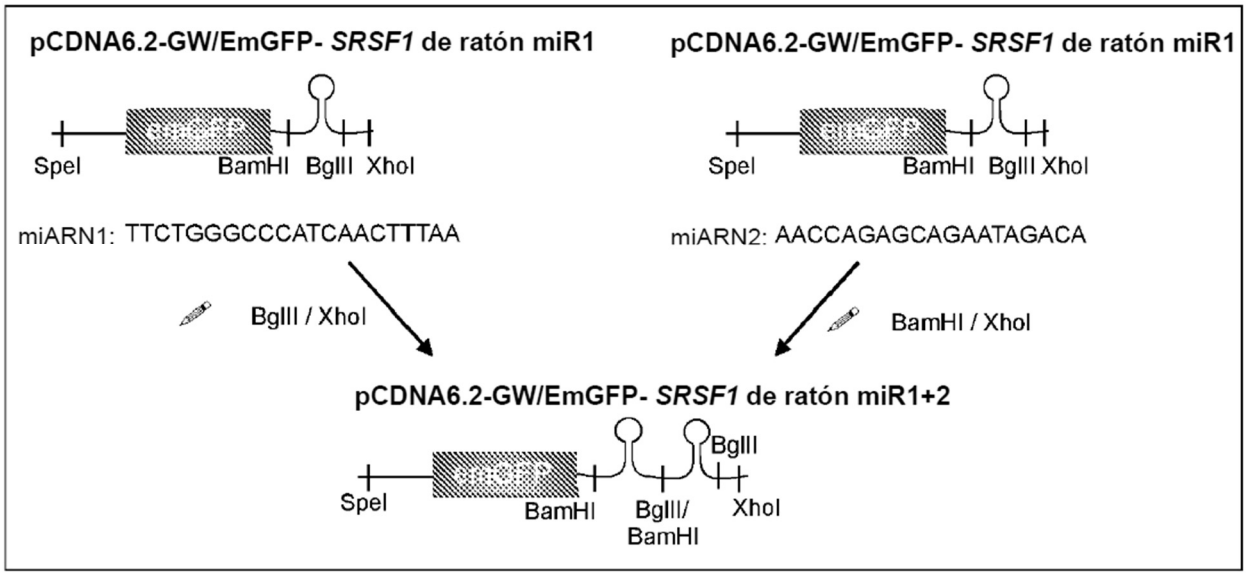

pCDNA6.2-GW/EmGFP- *SRSF1* de ratón miR1
SpeI
BamHI
BglII
XhoI
miARN1: TTCTGGGCCCATCAACTTTAA
BglII / XhoI
pCDNA6.2-GW/EmGFP- *SRSF1* de ratón miR1
SpeI
BamHI
BglII
XhoI
miARN2: AACCAGAGCAGAATAGACA
BamHI / XhoI
pCDNA6.2-GW/EmGFP- *SRSF1* de ratón miR1+2
SpeI
BamHI
BglII/
BamHI
BglII
XhoI

Figura 6

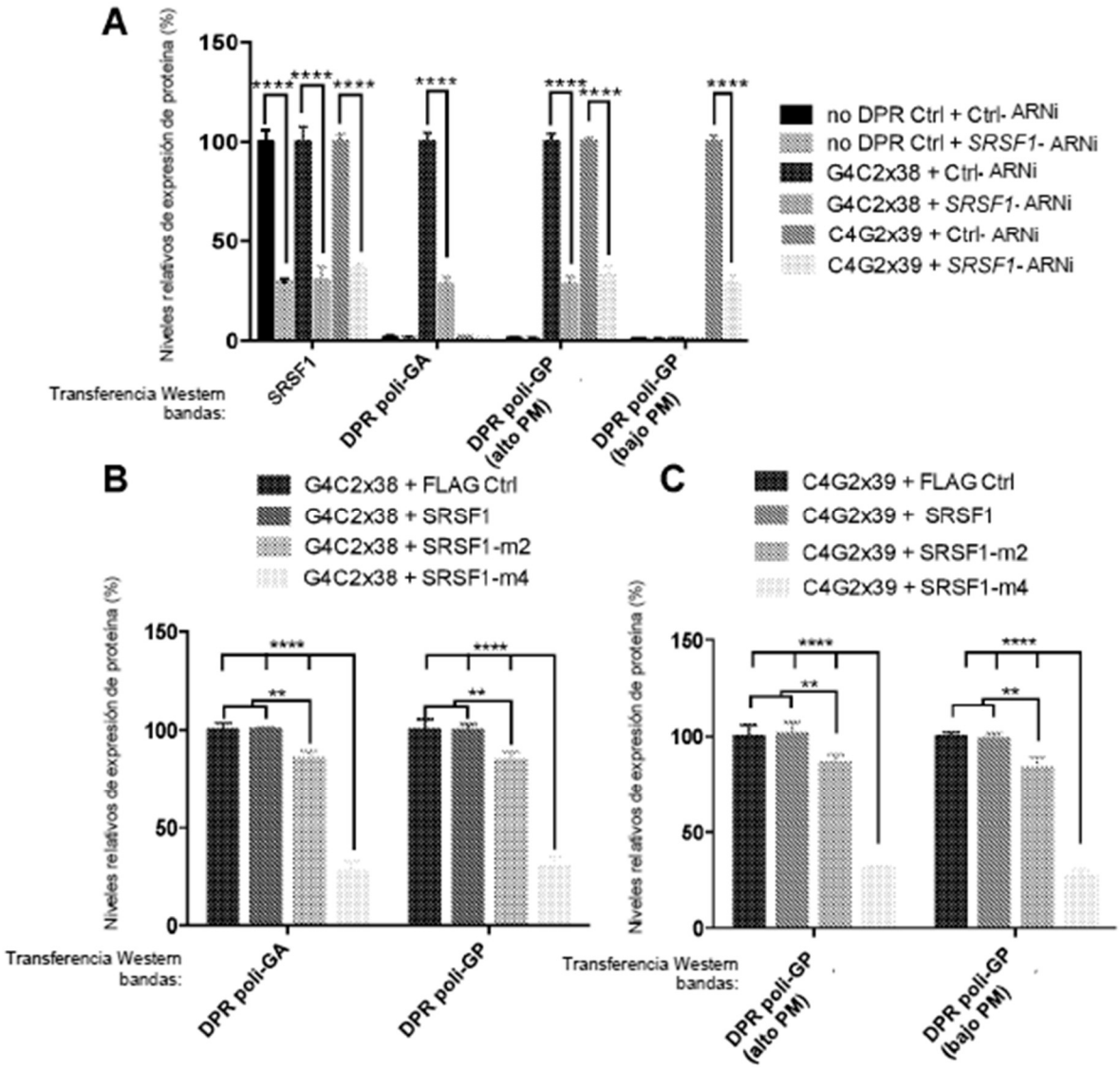

A
Niveles relativos de expresión de proteína (%)
150
100
50
0
****
no DPR Ctrl + Ctrl- ARNi
no DPR Ctrl + SRSF1- ARNi
G4C2x38 + Ctrl- ARNi
G4C2x38 + SRSF1- ARNi
C4G2x39 + Ctrl- ARNi
C4G2x39 + SRSF1- ARNi
Transferencia Western bandas:
SRSF1
DPR poli-GA
DPR poli-GP (alto PM)
DPR poli-GP (bajo PM)
B
G4C2x38 + FLAG Ctrl
G4C2x38 + SRSF1
G4C2x38 + SRSF1-m2
G4C2x38 + SRSF1-m4
Niveles relativos de expresión de proteína (%)
**
Transferencia Western bandas:
DPR poli-GA
DPR poli-GP
C
C4G2x39 + FLAG Ctrl
C4G2x39 + SRSF1
C4G2x39 + SRSF1-m2
C4G2x39 + SRSF1-m4
Niveles relativos de expresión de proteína (%)
Transferencia Western bandas:
DPR poli-GP (alto PM)
DPR poli-GP (bajo PM)

Figura 6 continuación

d)

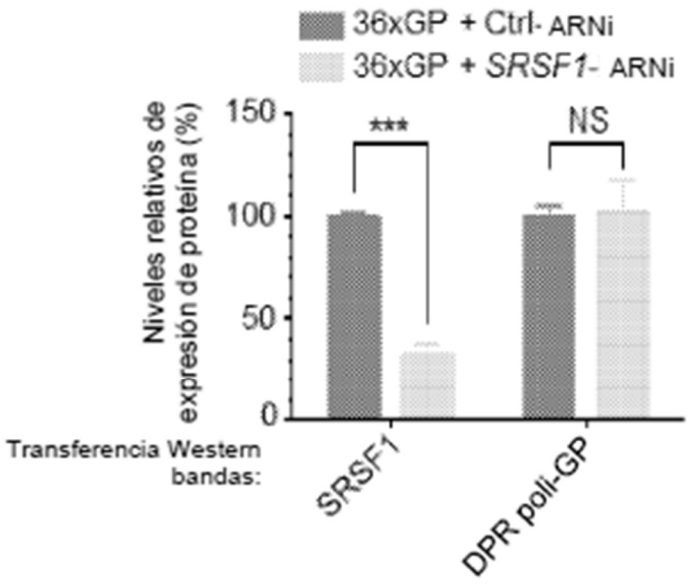


Figura 6 continuación

e)

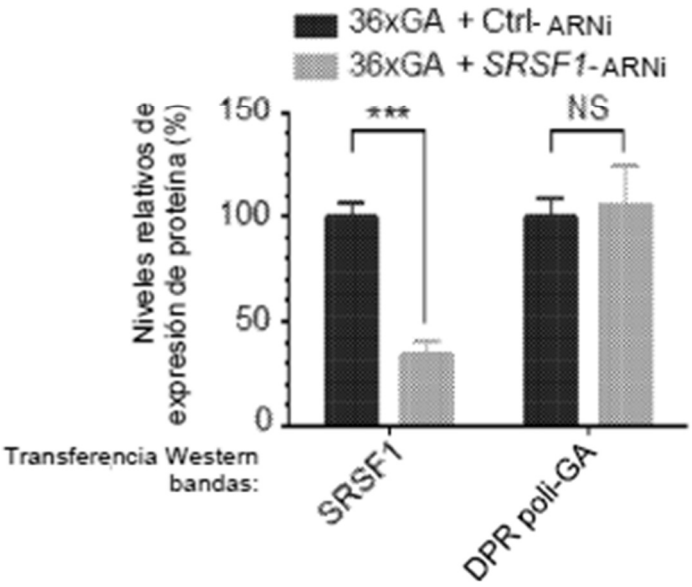


Figura 6 continuación

t)

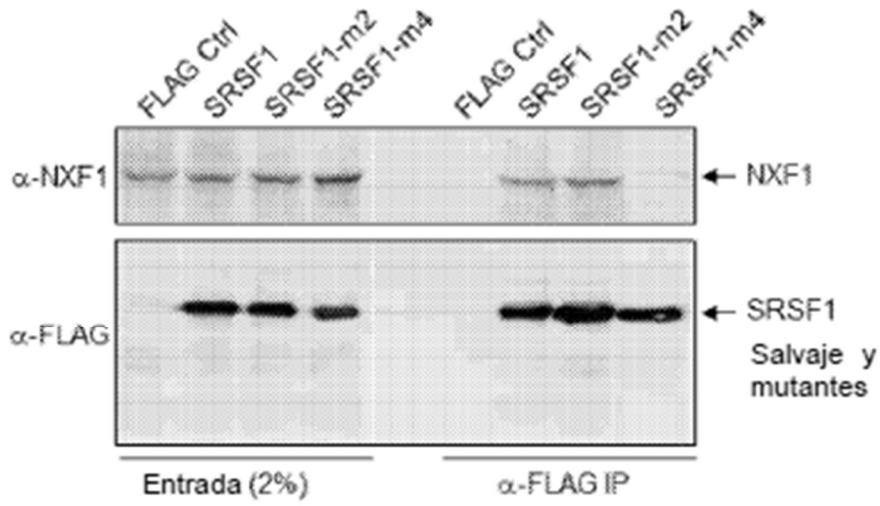


Figura 7

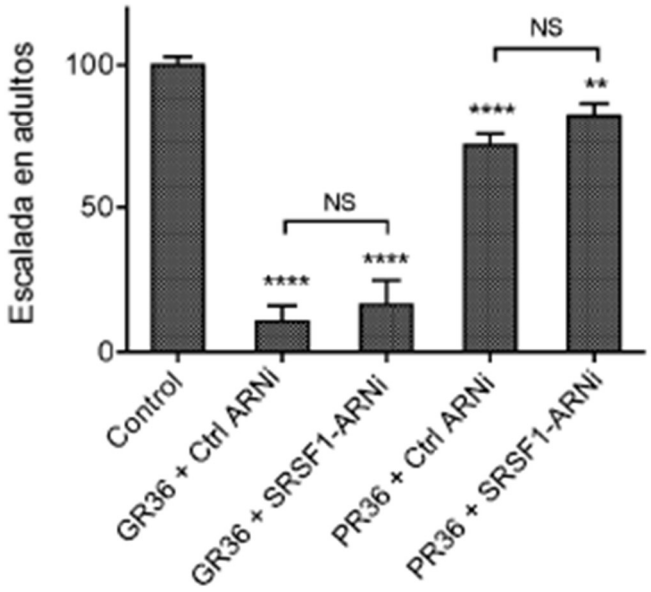
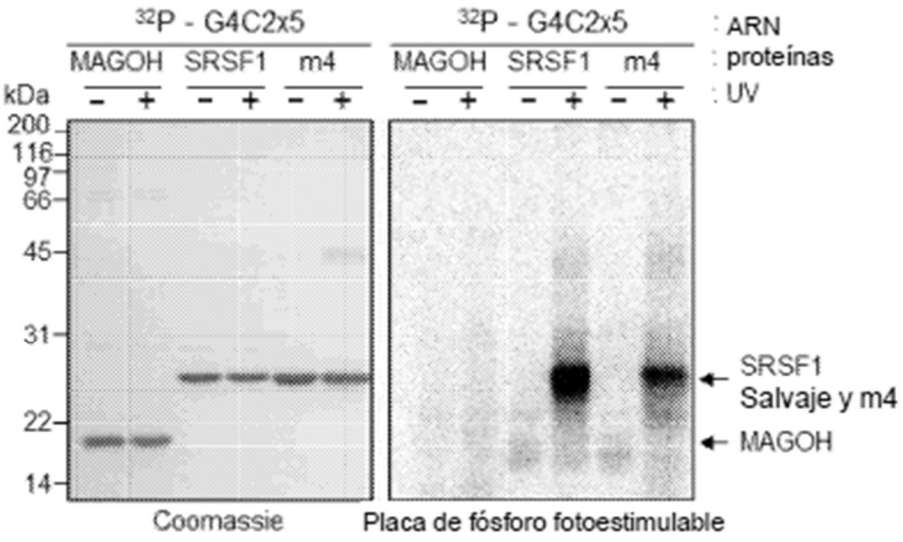


Figura 8

a)


32P - G4C2x5
32P - G4C2x5
: ARN
MAGOH SRSF1 m4
MAGOH SRSF1 m4
: proteínas
kDa
– + – + – +
– + – + – +
: UV
200
116
97
66
45
31
22
14
SRSF1
Salvaje y m4
MAGOH
Coomassie
Placa de fósforo fotoestimulable

b)

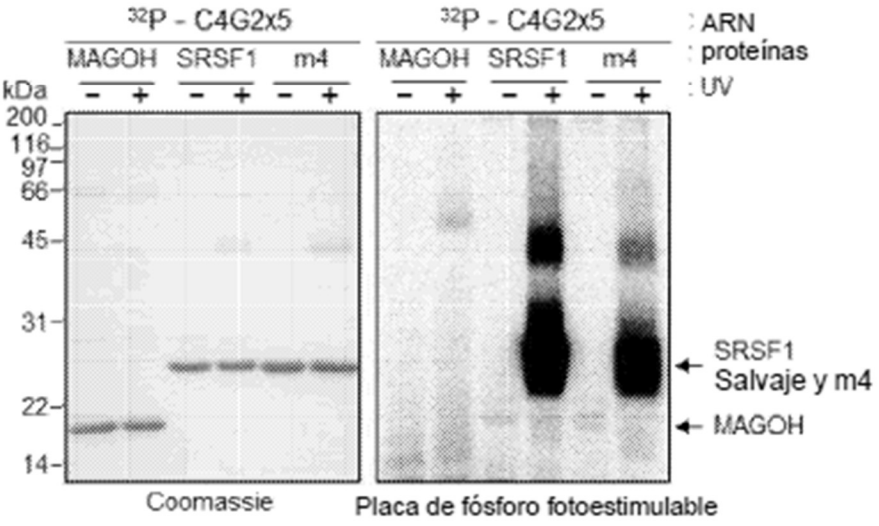
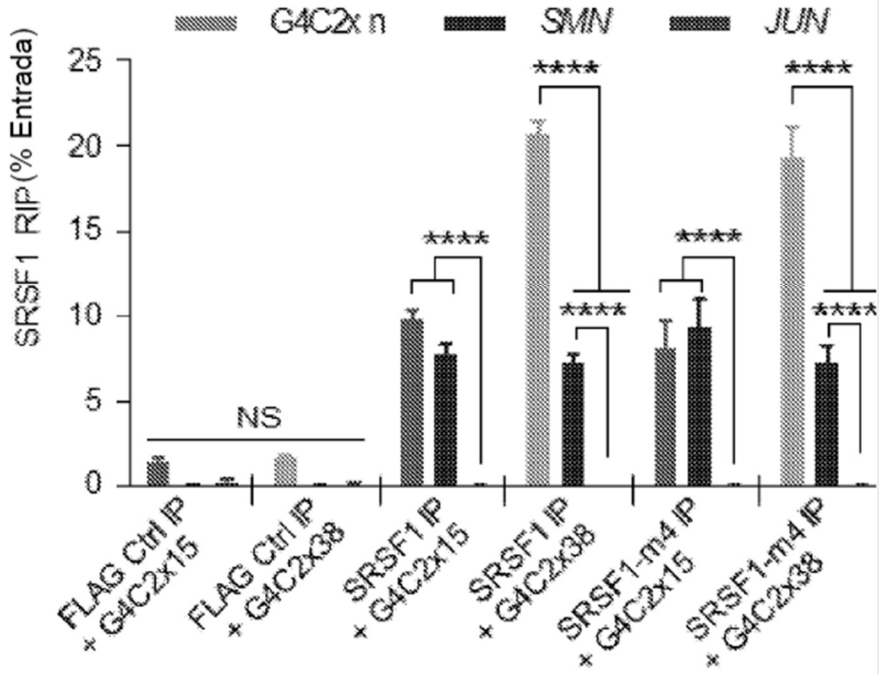

32P - C4G2x5
32P - C4G2x5
: ARN
MAGOH SRSF1 m4
MAGOH SRSF1 m4
: proteínas
kDa
– + – + – +
– + – + – +
: UV
200
116
97
66
45
31
22
14
SRSF1
Salvaje y m4
MAGOH
Coomassie
Placa de fósforo fotoestimulable

Figura 8 continuación

c)



d)

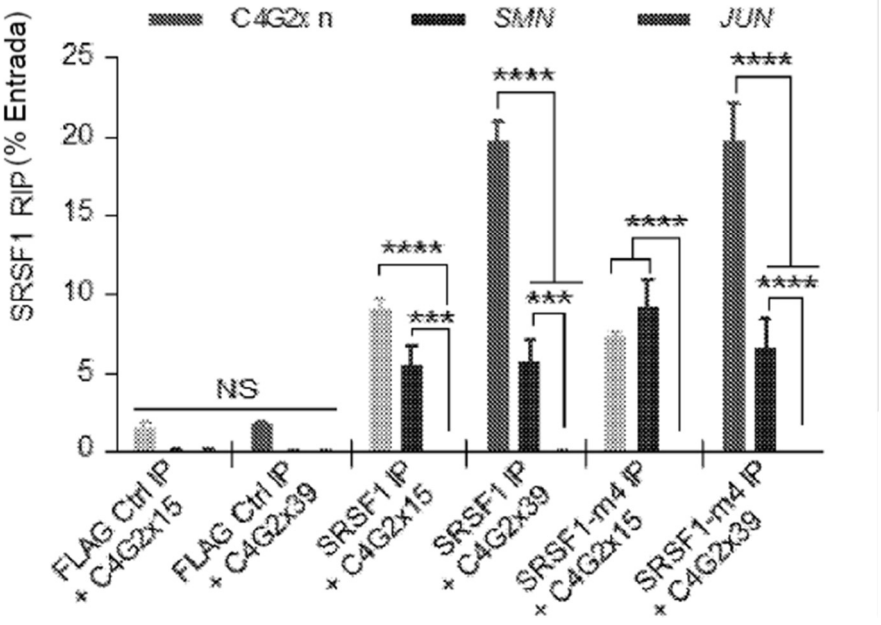
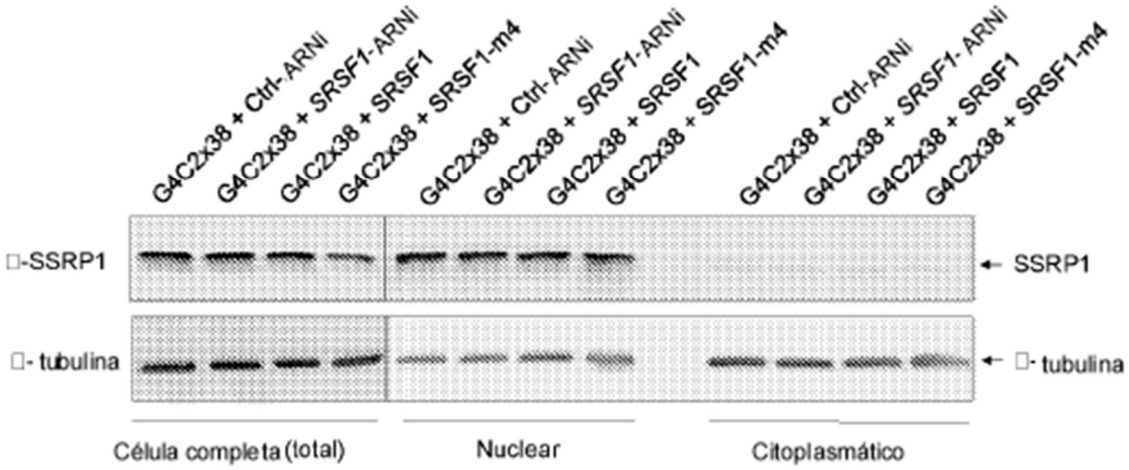


Figura 8 continuación

e)



g)

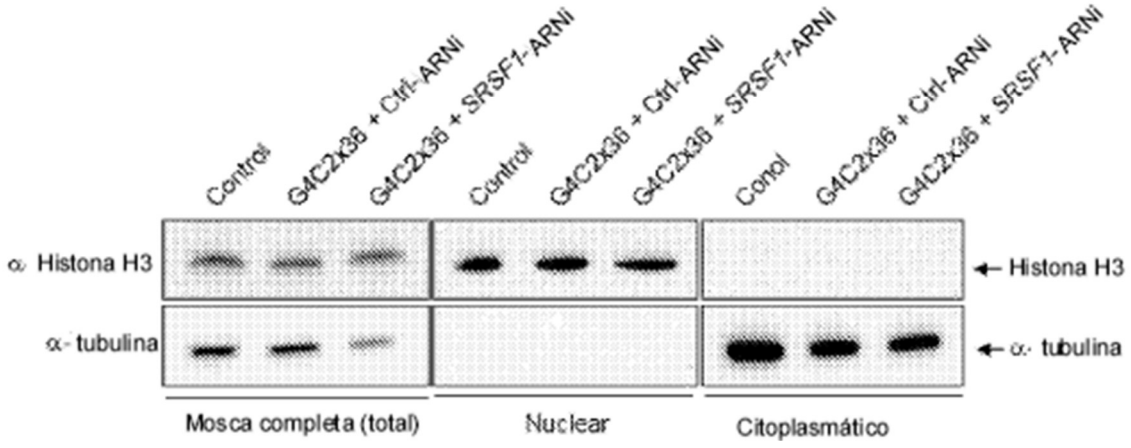


Figura 8 continuación

h)

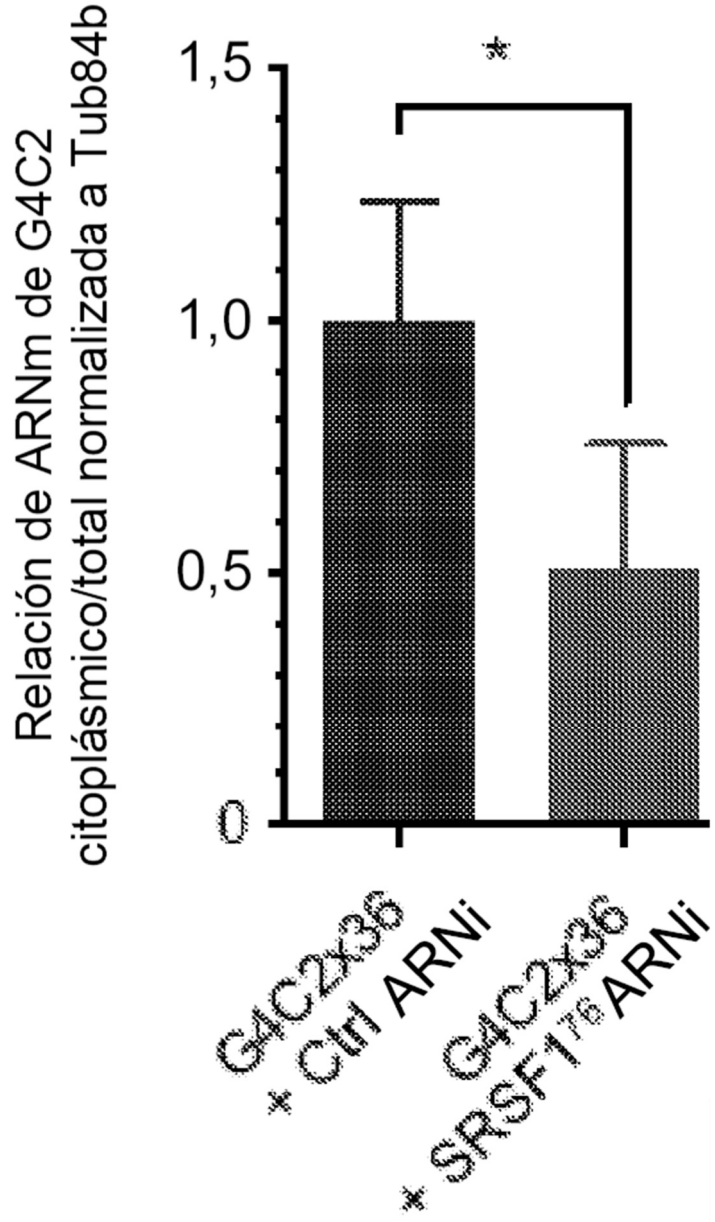


Figura 9

A

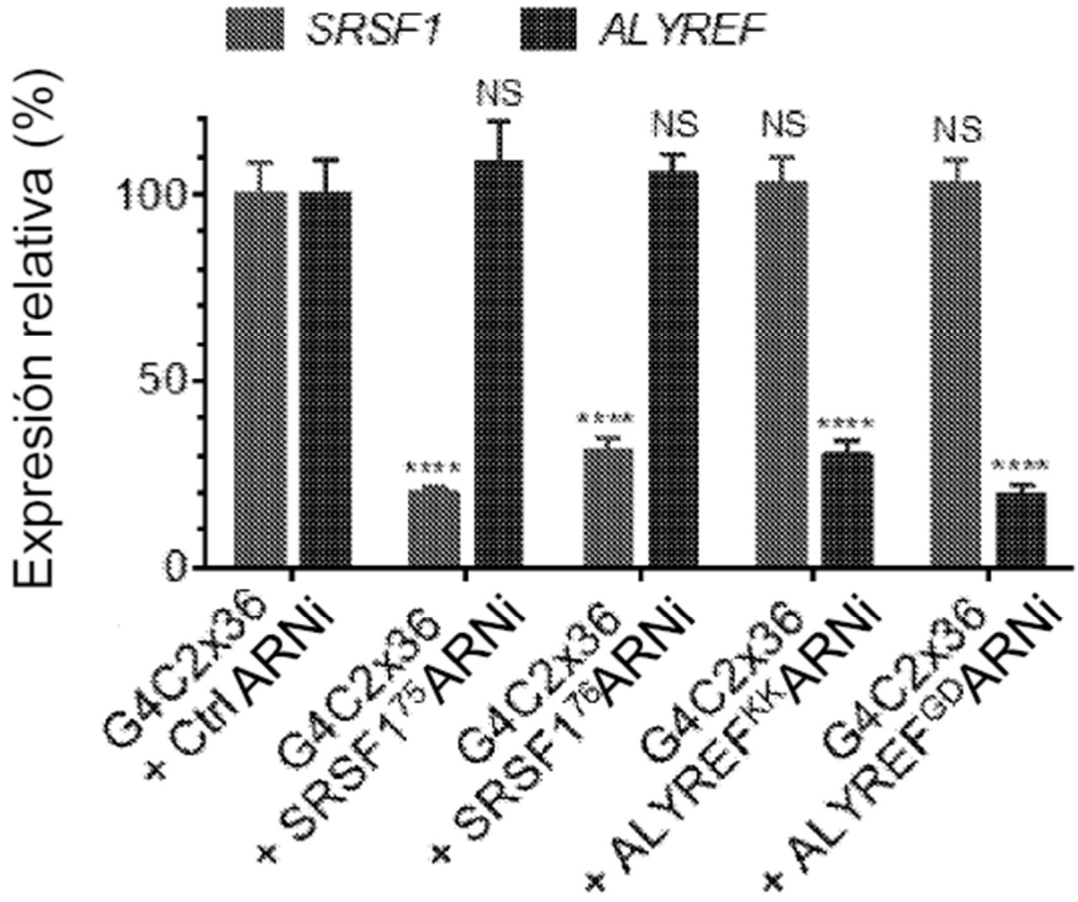

SRSF1
ALYREF
Expresión relativa (%)
100
50
0
NS
NS
NS
NS
****
****
****
****
G4C2x36 + CtrlARNi
G4C2x36 + SRSF1^175ARNi
G4C2x36 + SRSF1^176ARNi
G4C2x36 + ALYREF^KKARNi
G4C2x36 + ALYREF^GDARNi

Figura 9 continuación

B

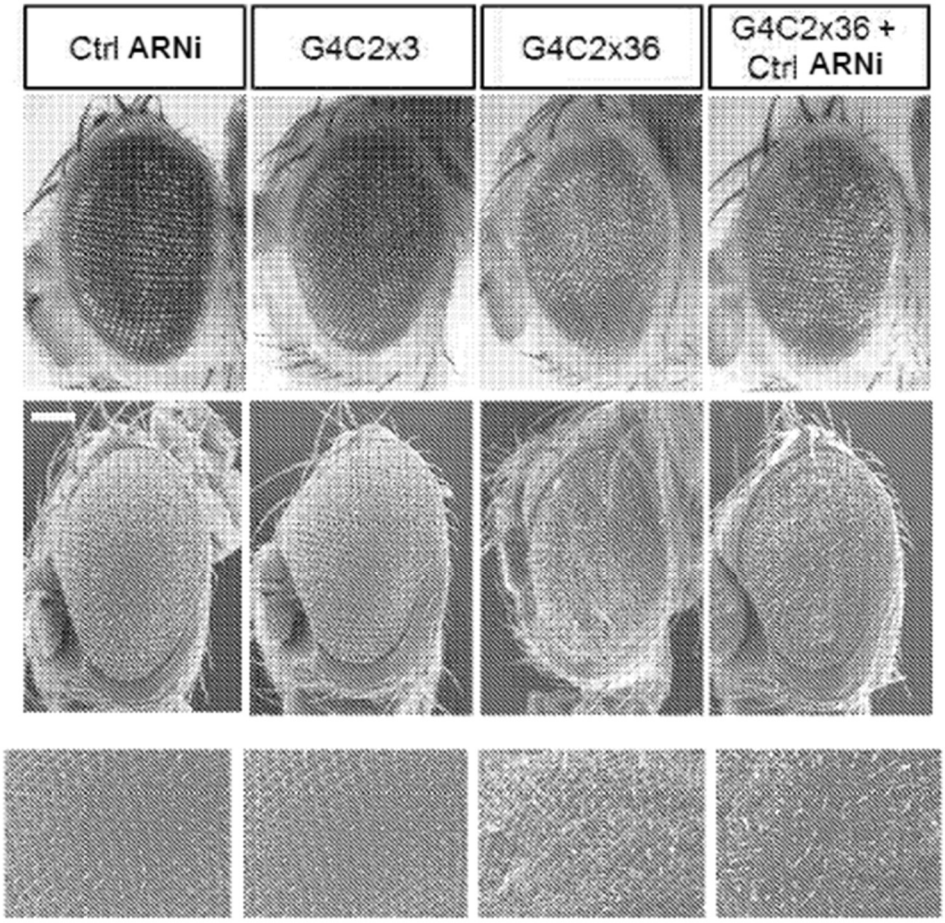


Figura 9B continuación

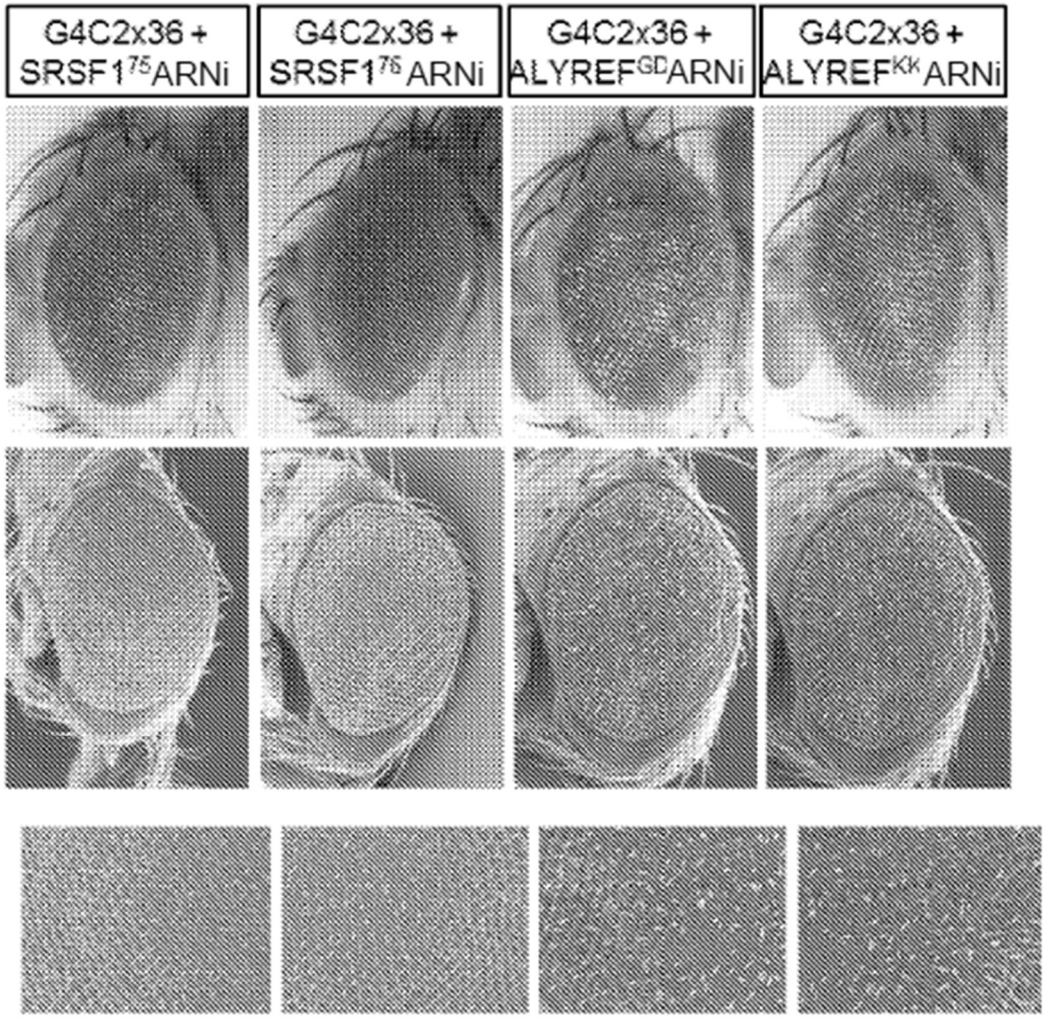


Figura 9 continuación

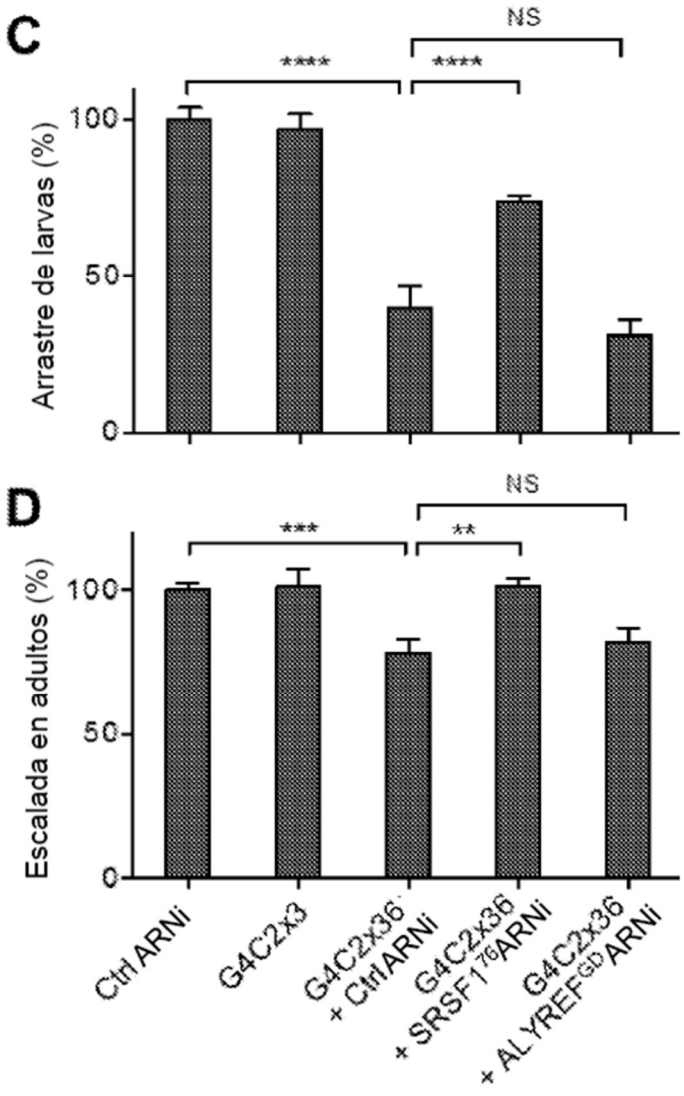
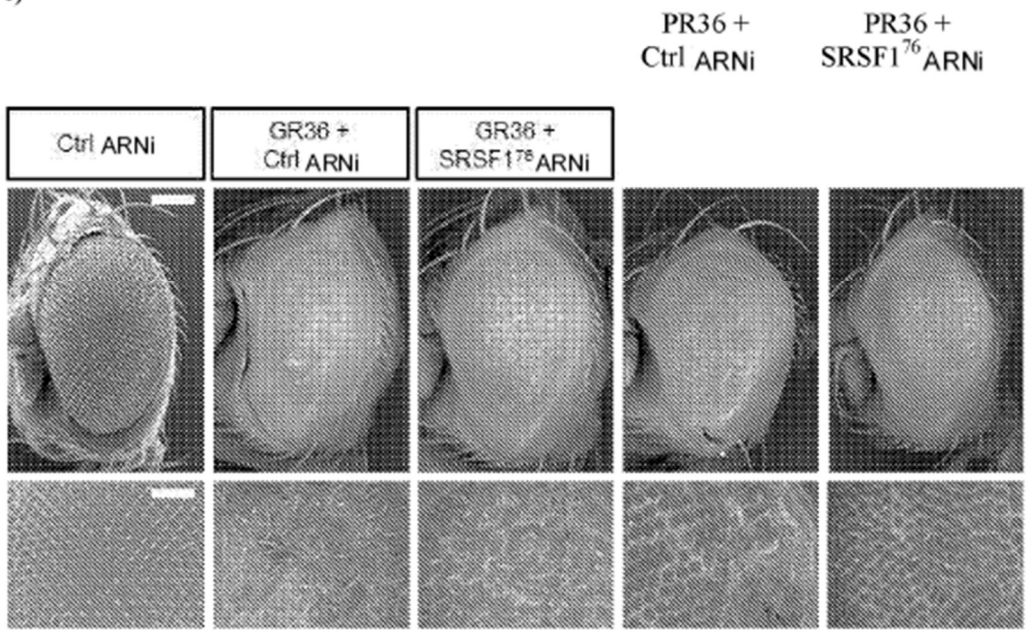


Figura 9 continuación

e)



f)

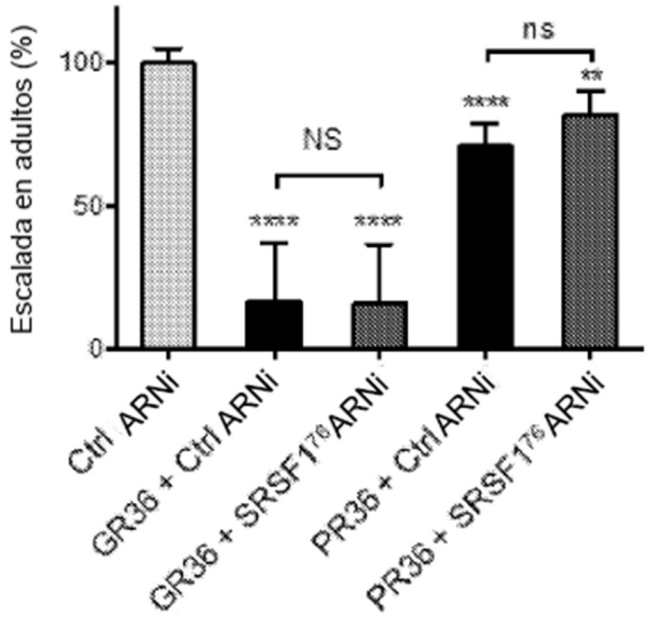
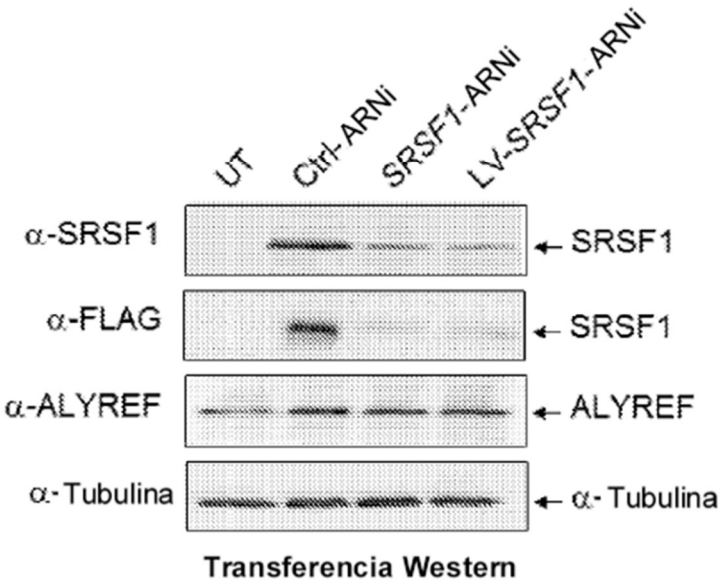


Figura 10

A


UT
Ctrl-ARNi
SRSF1-ARNi
LV-SRSF1-ARNi
α-SRSF1
SRSF1
α-FLAG
SRSF1
α-ALYREF
ALYREF
α-Tubulina
α-Tubulina
Transferencia Western

B

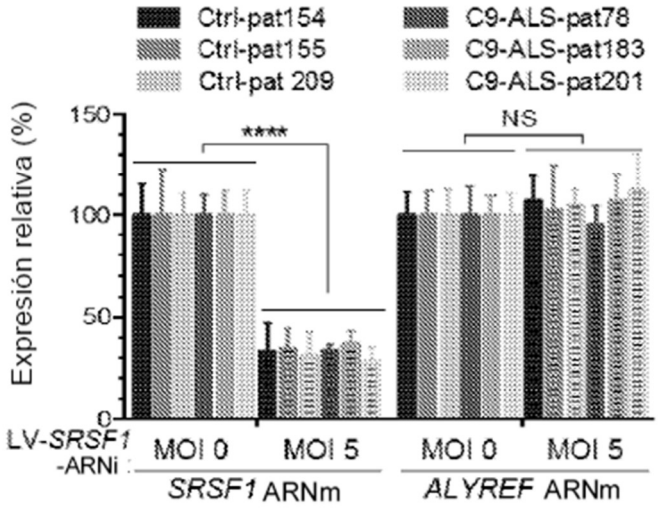
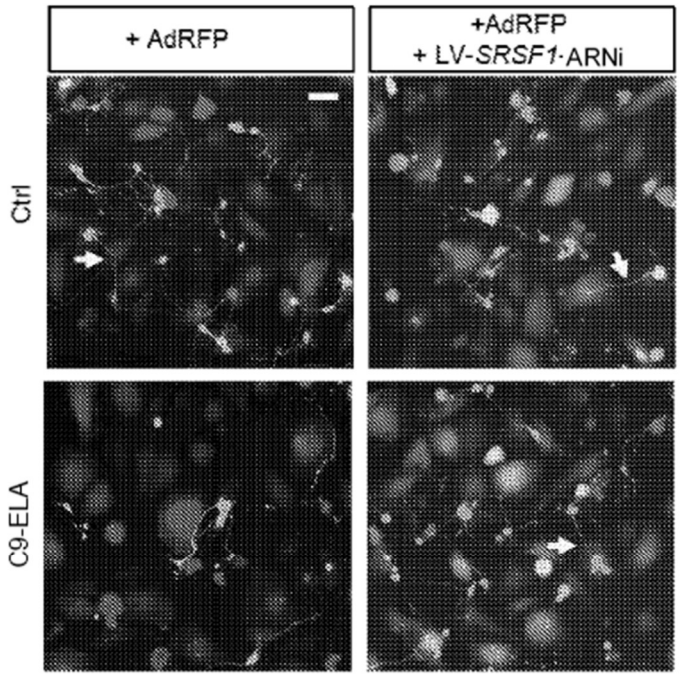

Ctrl-pat154
Ctrl-pat155
Ctrl-pat 209
C9-ALS-pat78
C9-ALS-pat183
C9-ALS-pat201
Expresión relativa (%)
150
100
50
0
****
NS
LV-SRSF1 -ARNi :
MOI 0
MOI 5
MOI 0
MOI 5
SRSF1 ARNm
ALYREF ARNm

Figura 10 continuación

C



D

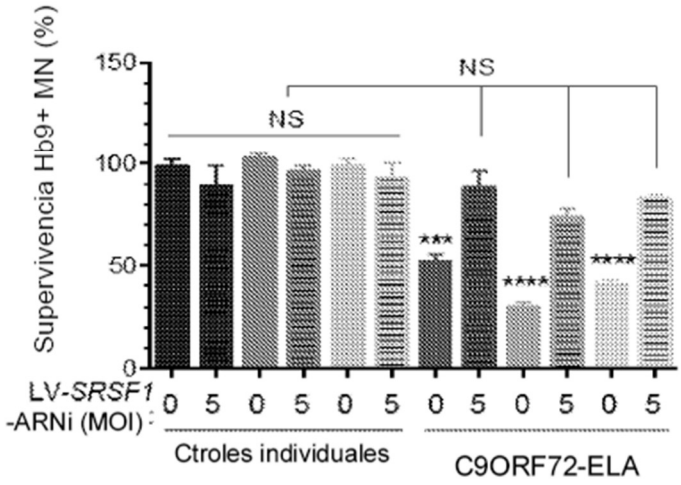


Figura 10 continuación

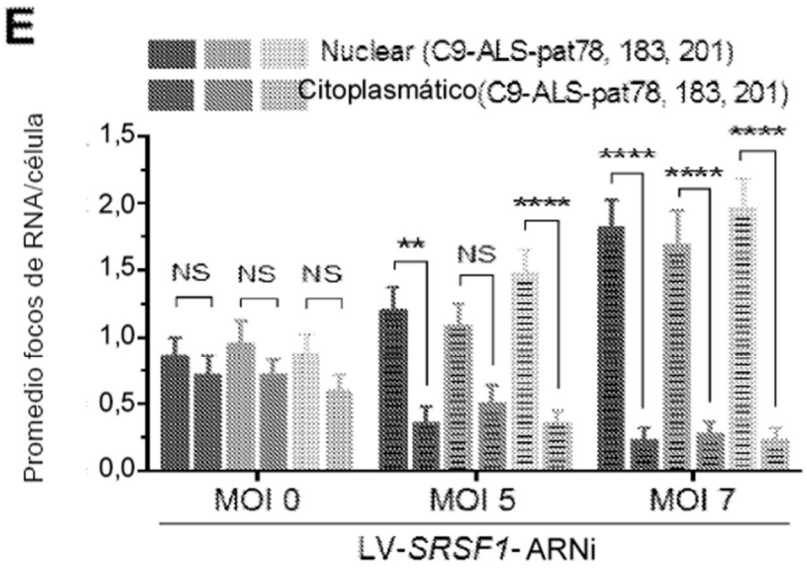
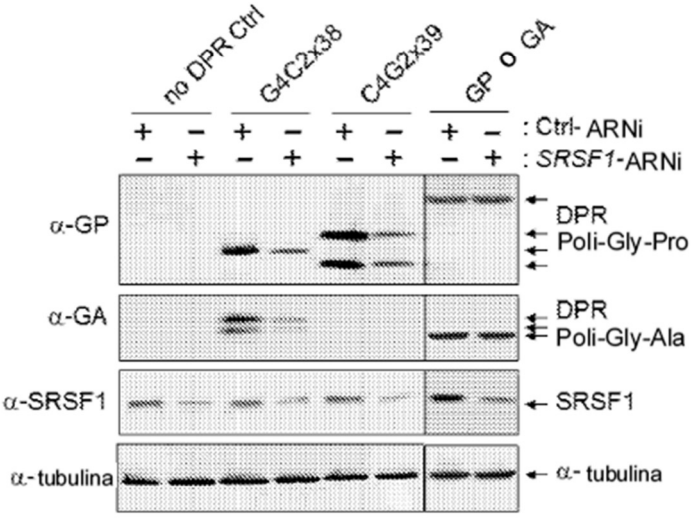
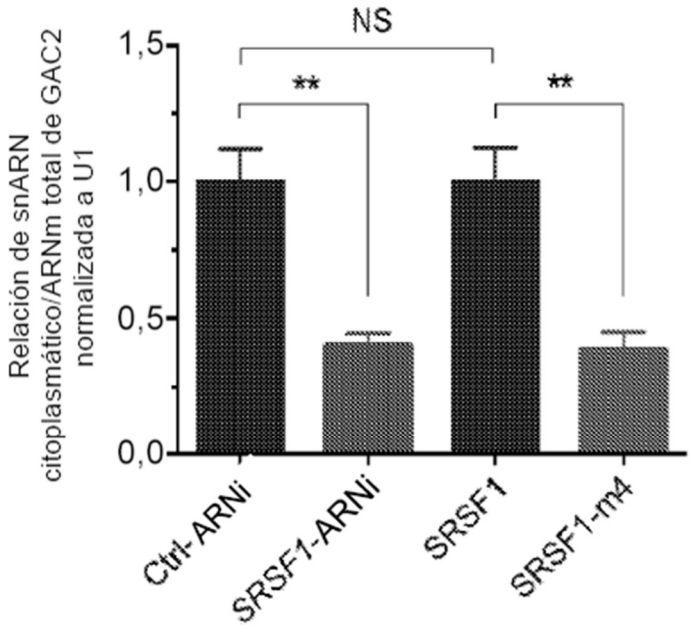


Figura 11

A



B


NS
**
**
Relación de snARN citoplasmático/ARNm total de GAC2 normalizada a U1
1,5
1,0
0,5
0,0
Ctrl-ARNi
SRSF1-ARNi
SRSF1
SRSF1-m4

C

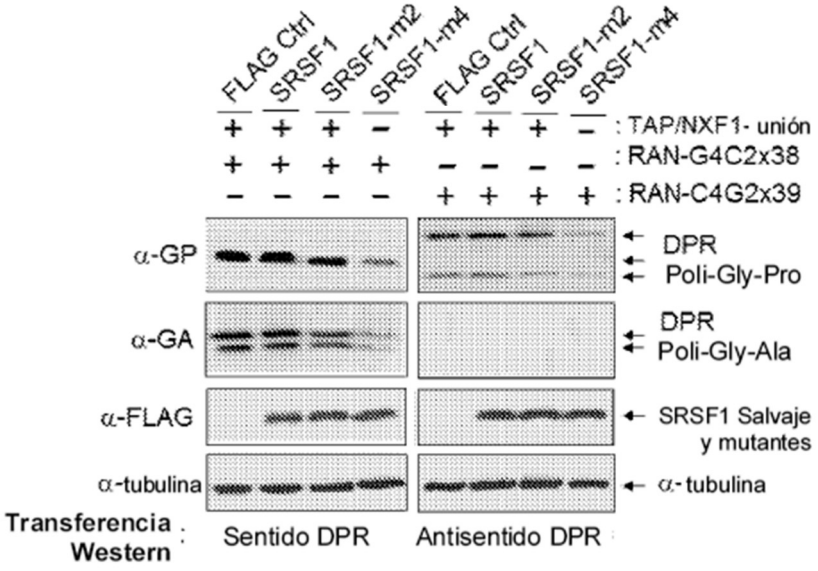

FLAG Ctrl
SRSF1
SRSF1-m2
SRSF1-m4
FLAG Ctrl
SRSF1
SRSF1-m2
SRSF1-m4
+ + + − + + + − : TAP/NXF1- unión
+ + + + − − − − : RAN-G4C2x38
− − − − + + + + : RAN-C4G2x39
α-GP
DPR
Poli-Gly-Pro
α-GA
DPR
Poli-Gly-Ala
α-FLAG
SRSF1 Salvaje y mutantes
α-tubulina
α-tubulina
Transferencia Western
Sentido DPR
Antisentido DPR

Figura 11

D

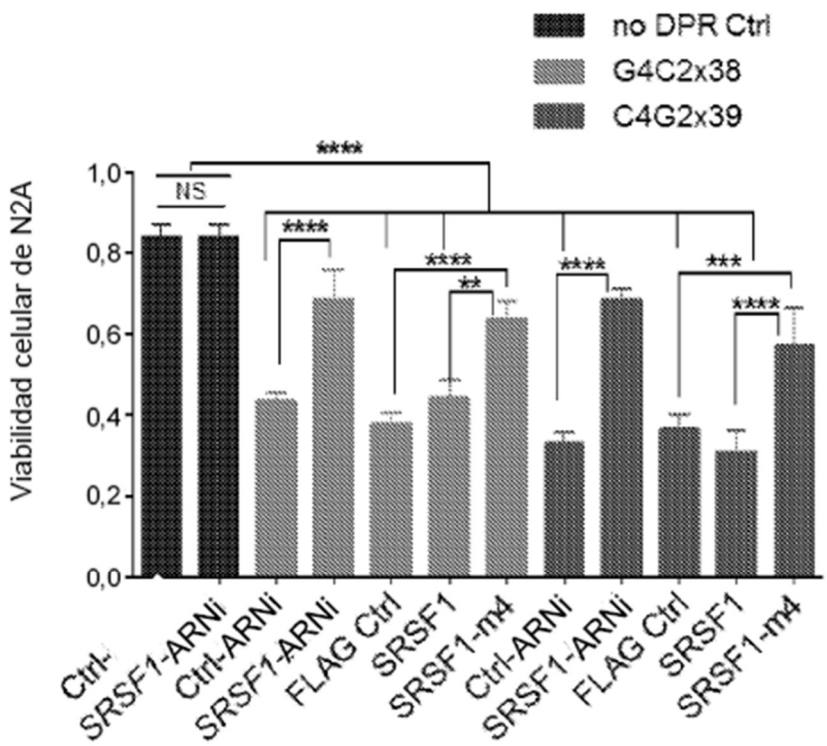


Figura 12

A1) (SEQ ID NO 1)

```
1     atgtcgggagg tggtgtgatt cgtggccccg
241   cagggaacaa cgattgccgc atctacgtgg gtaacttacc tccagacatc cgaaccaagg
301   acattgagga cgtgttctac aaatacggcg ctatccgcga catcgacctc aagaatcgcc
361   gcgggggacc gcccttcgcc ttcgttgagt tcgaggaccc gcgagacgcg gaagacgcgg
421   tgtatggtcg cgacggctat gattacgatg ggtaccgtct gcgggtggag tttcctcgaa
481   gcggccgtgg aacaggccga ggcggcggcg ggggtggagg tggcggagct ccccgaggtc
541   gctatggccc cccatccagg cggtctgaaa acagagtggt tgtctctgga ctgcctccaa
601   gtggaagttg gcaggattta aaggatcaca tgcgtgaagc aggtgatgta tgttatgctg
661   atgtttaccg agatggcact ggtgtcgtgg agtttgtacg gaaagaagat atgacctatg
721   cagttcgaaa actggataac actaagttta gatctcatga gggagaaact gcctacatcc
781   gggttaaagt tgatgggccc agaagtccaa gttatggaag atctcgatct cgaagccgta
841   gtcgtagcag aagccgtagc agaagcaaca gcaggagtcg cagttactcc ccaaggagaa
901  gcagaggatc accacgctat tctccccgtc atagcagatc tcgctctcgt acataa
```

Figura 12 continuación

A2) (SEQ ID NO 2):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGGAPRGRYGP
PSRRSENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGRSRSRSRSRSRSRSRSNSRSRSYS
PRRSRGSPRYSPRHSRSRSRT

A3) (SEQ ID NO: 3)

```
      atgtcgggagg tggtgtgatt cgtggccccg
241   cagggaacaa cgattgccgc atctacgtgg gtaacttacc tccagacatc cgaaccaagg
301   acattgagga cgtgttctac aaatacggcg ctatccgcga catcgacctc aagaatcgcc
361   gcgggggacc gcccttcgcc ttcgttgagt tcgaggaccc gcgagacgcg gaagacgcgg
421   tgtatggtcg cgacggctat gattacgatg ggtaccgtct gcgggtggag tttcctcgaa
481   gcggccgtgg aacaggccga ggcggcggcg ggggtggagg tggcggagct ccccgaggtc
541   gctatggccc cccatccagg cggtctgaaa acagagtggt tgtctctgga ctgcctccaa
601   gtggaagttg gcaggattta aaggatcaca tgcgtgaagc aggtgatgta tgttatgctg
661   atgtttaccg agatggcact ggtgtcgtgg agtttgtacg gaaagaagat atgacctatg
721   cagttcgaaa actggataac actaagttta gatctcatga ggtaggttat acacgtattc
781  ttttctttga ccagaattgg atacagtggt cttaa
```

A4)
(SEQ ID NO: 4)

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGGAPRGRYGP
PSRRSENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEVGYTRILFFDQNWIQWS

A5) (SEQ ID NO 23)

```
1     gttgcgctcc cgcgcgtgcg tgttgggatc gaatcgctgt ttcgttccgc ttctcttcct
61    ctgtctcccc cccatatccg tgcgccgagc tgataaaggc gccattttgg aggggccgcg
121   ggagacgtgg tgccgctgcg ggctcgctct gccgtgcgct aggcttggtg ggaaggcctg
181   ttctcgagtc cgcgcttttc gtcaccgcca tgtcgggagg tggtgtgatt cgtggccccg
241   cagggaacaa cgattgccgc atctacgtgg gtaacttacc tccagacatc cgaaccaagg
301   acattgagga cgtgttctac aaatacggcg ctatccgcga catcgacctc aagaatcgcc
361   gcgggggacc gcccttcgcc ttcgttgagt tcgaggaccc gcgagacgcg gaagacgcgg
421   tgtatggtcg cgacggctat gattacgatg ggtaccgtct gcgggtggag tttcctcgaa
481   gcggccgtgg aacaggccga ggcggcggcg ggggtggagg tggcggagct ccccgaggtc
541   gctatggccc cccatccagg cggtctgaaa acagagtggt tgtctctgga ctgcctccaa
601   gtggaagttg gcaggattta aaggatcaca tgcgtgaagc aggtgatgta tgttatgctg
661   atgtttaccg agatggcact ggtgtcgtgg agtttgtacg gaaagaagat atgacctatg
721   cagttcgaaa actggataac actaagttta gatctcatga ggtaggttat acacgtattc
781   ttttctttga ccagaattgg atacagtggt cttaacagtg gaatttcaag gtaaggattc
841   aggcaaggtt gtccaagtaa attgccagat ttctggtttt agttacattg tattcattca
901   gcatgtctga agatagatga aagcttagat ctttcaatgg aaagttctgt ctatccaata
```

```
961  gggagaaact gcctacatcc gggttaaagt tgatgggccc agaagtccaa gttatggaag
1021 atctcgatct cgaagccgta gtcgtagcag aagccgtagc agaagcaaca gcaggagtcg
1081 cagttactcc ccaaggagaa gcagaggatc accacgctat tctccccgtc atagcagatc
1141 tcgctctcgt acataagatg attggtgaca ctttttgtag aacccatgtt gtatacagtt
1201 ttcctttatt cagtacaatc ttttcatttt ttaattcaaa ctgttttgtt cagaatgggc
1261 taaagtgttg aattgcattc ttgtaatatc cccttgctcc taacatctac attcccttcg
1321 tgtctttgat aaattgtatt ttaagtgatg tcatagacag gattgtttaa atttagttaa
1381 ctccatactc ttcagactgt gatattgtgt aaatgtctat ctgccctggt ttgtgtgaac
1441 tgggatgttg ggggtgtttg tggttatctt acctggggaa gttcttatgt ttatcttgct
1501 tttcatgtgt ctttctgtag acatatctga agagatggat taagaatgct ttggattaag
1561 gattgtggag cacatttcaa tcattttagg attgtcaaaa ggaggattga ggaggatcag
1621 atcaataatg gaggcaatgg tttggattgg agagggctca ctggatccca atccttggag
1681 ctggatcatt ggattcaaat cataatgtgg ataggatagg gaggatgaat taccaggatt
1741 catggagcgg gatcagatta ccaggaacat aggagtggat tcctgcccca accaaaccgc
1801 attcgtgtgg attttttttat tcaacttaat tggctattcc aaagattttt tttttcctat
1861 ttttgacgat tggagccctt aagatgcacg atggaattgt gttttgcgtt ttttggtaaa
1921 aggagcaaag cgaggacctg gagataaacg ctggagcaat ctccttggaa ggattcagca
1981 cgagtagatg gtaaacattt aaaggggaaa gggggggttt gtttaaaata gtaaatcagt
2041 aagtcacttc taaatttaaa gaaaacaaaa ttggagttga agaataagta ggtttccaat
2101 tggctattgc cgttttcttt gaaaaaataa acatttttta aaaaactatg catggttgtc
2161 ctttttcctc ttcatgtaag attctaactg ggtctatcag ttaatcttta aattgttaag
2221 taagataaga ttttgactct tgtgttaatg tgttagcaaa ttaaaagttc ttaaaaggca
2281 atctaatggt attagccatc ttttattgtt aattgtaaaa gtcttcaggg gaaagcaaaa
2341 ggggagaata aggcatttgt gtatgtaact tggtaaatga cggtgggggа tggatctagc
2401 atctgaaaga taagcttctc tactttgtta taaagtggtt aaaaaactat agatgctgct
2461 tattttctgg tggtcataga caacataggc ttttgtgcaa aattggttga tggctactaa
2521 tgttcacttg gagatagctt ttgatattct caatgaaact catctcaaaa aaaggtaagt
2581 attaaatgtt aacatcagca cagatgtatt agaactgttt tttgtttttg agacagagtc
2641 tcgctctgtt ctccagactg gagggcagtg gtgtgatcta ggctcactgc aacctccacc
2701 cctggatttg agtgattctc gtgcctcagt ctcccaagta gctgagacta caagtgtgtg
2761 ccacccttgc ccggctaatt ttgtattttt agtagagatg tggtttctct gtgttaccca
2821 ggctggtccc aaaactcctg gcctcaagtg atctgcctgc cttggcctcc caaagtgtta
2881 ggattacagg tgtgagccac catgctcagc ctgtagaact tttaacccaa gtctcatttc
2941 ttttttgaaa gggaagagtg cacaagatta actgcttctt tggatgaatc attgttaata
3001 aaaagctggg catttagaat tttgccttat aagcccttct ccaaccataa gattattttg
3061 taccaaaaac tttggtgttc tctaccaaag cagttaaaaa cttttagcct gctacttctt
3121 gtatttgtct actgacagcc ccttggtact atttaggttg ggggagggga cctaaaataa
3181 atagacttta acatttccct tgggtgctaa tcatagttgg aagttgaatt taaggtgatt
3241 atttgggtga caattaaaaa cctaaggaaa accagaaatc ttggtagtgg aagaaatgtg
3301 taaggtcacc ccaatcggta gattttaatg aacgttgtgg aatgttggga agaggggatg
3361 ttaagttgaa tgcagaattt cactaagtac ttagtgtaag tttaaggatg tagctctttt
3421 tatctaagaa ttcaatgtaa tggccaaaag gcagatttac tgtttaaaaa tttgaataat
3481 tttacatgac attcttgaaa ttctaagaag ttttatgtgt agaacatttt aaaaattcat
3541 cagattatta aagggaaaat aaatgattaa tgataatttt ggaaggttaa tgtgagctag
3601 acttaagtaa actttggttc atttgtgttc attgaatgtt ttggaaatga ccaaaaaatg
3661 taaatggcct tcactcaagt ttgagtgttt aaagttgaaa gatgtgctct actaaaagtt
3721 atagtaattc taacctcaca ttgaaatgag acagtattcc ttgttataca ggctgaattt
3781 gaagattaga gaggatctaa tgtttactta ggtaaagggg cataggtttt gtagttaaga
3841 tgaccagaca gctaaaagct gtgatgggaa gtatggactg ctcctattta tagtctcaga
3901 aaatggacct ttaggtctct atccgtattg gcaattatta gaagaaagtt acacccttтt
3961 gaaactacaa aagctgtctt ggaatttccc cctcttctcc ctatttatgt ccccttagaa
4021 tattttaggg agcctataat tattttctaa ccaaggaaaa acttaagtct ctttaagaag
4081 caattacttt tcataacatc agattgaata acccaccttg ctgttcagcc cacatcctac
4141 tggaaacaaa aggtaagaaa cccatttтct ggttcttgat tgtttgggtc tgaattttgt
4201 ttttaaaact aagctaagtt taatgttttt taaaatgctg tttggaatat gaatagattc
4261 cccgtaaaat gatttttcct aagttttatg ctttagtaaa tattcagtgc tcacgtctgt
4321 gcatcatagt gcttgcgttt aatatgattt attgtagaat ctcaactttt cttggtgttt
```

```
4381 gttgtctttg aaacattgtc ttggtcatta gggctggtgt tttcacattt ctgtggtcaa
4441 ggtggatttc ttatgtgtgc ctttttgctt actttgtata tgaattttgt aatttaaatt
4501 gcaagtaagt tatatatatg tatttaccat aaatagtatt aaaagatgag aaactgttag
4561 actgaagttc tgttgtaaca taaccattat ttccatcaca gtatgaagac tgcaaacgca
4621 gaaaacagat tacagtctct tatccatttt ttgaaatcca aaaactacga aaacaaaaga
4681 ttttctgttg ttgagctaat taaatgtgaa ccctgaccag aaaaaaaaaa aaaaaaaaaa
4741 aaaaaaa
```

B1) (SEQ ID NO 24)

```
1     atgtcgggaggt ggtgtgatcc
121   gtggcccggc ggggaacaac gactgccgca tctacgtggg taacctacct ccggatatcc
181   gaaccaagga catcgaggac gtgttttaca aatacggcgc catccgcgac atcgacctga
241   agaaccgccg cgggggaccg cccttcgcct tcgttgagtt cgaggacccg cgagacgcgg
301   aagatgcggt gtacggtcgc gacggctacg actacgacgg ctaccggctg cgggtagagt
361   ttccccgaag cggccgcggg accggccgag gcggcggcgg gggtggaggc ggcggcgccc
421   cgagaggccg ctatggcccg ccgtccaggc ggtccgagaa cagagtggtt gtctctggac
481   tgcctccgag tggaagctgg caggacttaa aggatcacat gcgtgaggca ggtgatgtat
541   gttacgctga tgtttaccga gatggcactg gtgtcgtgga gtttgtacgg aaagaagata
601   tgacgtatgc agttcgaaaa ctggataaca ctaagtttag atctcacgag ggagaaactg
661   cctacatccg ggttaaagtt gatgggccca gaagtccaag ttatggaaga tctcgatctc
721   gaagccgtag tcgtagcaga agccgtagca gaagcaacag caggagtcgc agttactccc
781   caaggagaag cagaggatca ccacgctatt ctccccgtca tagcagatct cgctctcgta
841   cataa
```

B2) (SEQ ID NO 25)

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGGAPRGRYGP
PSRRSENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGRSRSRSRSRSRSRSRSNSRSRSYS
PRRSRGSPRYSPRHSRSRSRT

B3) (SEQ ID NO 26)

```
1    atgt cgggaggtgg
481   tgtgatccgt ggcccggcgg ggaacaacga ctgccgcatc tacgtgggta acctacctcc
541   ggatatccga accaaggaca tcgaggacgt gttttacaaa tacggcgcca tccgcgacat
601   cgacctgaag aaccgccgcg ggggaccgcc cttcgccttc gttgagttcg aggacccgcg
661   agacgcggaa gatgcggtgt acggtcgcga cggctacgac tacgacggct accggctgcg
721   ggtagagttt ccccgaagcg gccgcgggac cggccgaggc ggcggcgggg gtggaggcgg
781   cggcgccccg agaggccgct atggcccgcc gtccaggcgg tccgagaaca gagtggttgt
841   ctctggactg cctccgagtg gaagctggca ggacttaaag gatcacatgc gtgaggcagg
901   tgatgtatgt tacgctgatg tttaccgaga tggcactggt gtcgtggagt ttgtacggaa
961   agaagatatg acgtatgcag ttcgaaaact ggataacact aagtttagat ctcacgaggg
1021  agaaactgcc tacatccggg ttaaagttga tgggcccaga agtccaagtt atggaagatc
1081  tcgatctcga agccgtagtc gtagcagaag ccgtagcaga agcaacagca ggagtcgcag
1141  ttactcccca aggagaagca gaggatcacc acgctattct ccccgtcata gcagatctcg
1201 ctctcgtaca taa
```

Figura 12 continuación

B4) (SEQ ID NO 27)

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGGAPRGRYGP
PSRRSENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGRSRSRSRSRSRSRSRSNSRSRSYS
PRRSRGSPRYSPRHSRSRSRT

B5) (SEQ ID NO 28)

```
   1    atgt cgggaggtgg
481   tgtgatccgt ggcccggcgg ggaacaacga ctgccgcatc tacgtgggta acctacctcc
541   ggatatccga accaaggaca tcgaggacgt gttttacaaa tacggcgcca tccgcgacat
601   cgacctgaag aaccgccgcg ggggaccgcc cttcgccttc gttgagttcg aggacccgcg
661   agacgcggaa gatgcggtgt acggtcgcga cggctacgac tacgacggct accggctgcg
721   ggtagagttt ccccgaagcg gccgcgggac cggccgaggc ggcggcgggg gtggaggcgg
781   cggcgccccg agaggccgct atggcccgcc gtccaggcgg tccgagaaca gagtggttgt
841   ctctggactg cctccgagtg gaagctggca ggacttaaag gatcacatgc gtgaggcagg
901   tgatgtatgt tacgctgatg tttaccgaga tggcactggt gtcgtggagt ttgtacggaa
961   agaagatatg acgtatgcag ttcgaaaact ggataacact aagtttagat ctcacgaggt
1021 aggttataca cttattcttt tttttggcca gaattggata cagttttctt aa
```

B6) (SEQ ID NO 29)

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGGAPRGRYGP
PSRRSENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEVGYTLILFFGQNWIQFS

Figura 13

A) A (SEQ ID NO 16)

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFV
EFEDPRDAEDAVYGRDGYDYDGYRLRVEFP**A**SG**A**GTGRGGGGGGGGGAPRGRYGP
PS**AA**SENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTY
AVRKLDNTKFRSHEGETAYIRVKVDG

B) B (SEQ ID NO 17):

MAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRVEFP**A**SG**A**GTGRGGGGGGGGGAPRGRYGPPS**AA**SENRV
VVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTK
FRSHEGETAYIRVKVDG

c) C (SEQ ID NO 152):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPP
FAFVEFEDPRDAEDAVYGRDGYDYDGYRLRVEFP**A**SG**A**GTGRGGGGGGGGGA
PRGRYGPPS**AA**SENRVVVSGLPPSGSWQDLKDHMREAGDVCYADVYRDGTGV

VEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGRSRSRSRSRS
RSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

Figura 14: (SEQ Id NO 15):
PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE

Figura 15:(SEQ ID NO: 18)
A)

```
1    atggag cggaaagtgc ttgcgctcca ggcccgaaag aaaaggacca aggccaagaa
181  ggacaaagcc caaaggaaat ctgaaactca gcaccgaggc tctgctcccc actctgagag
241  tgatctacca gagcaggaag aggagattct gggatctgat gatgatgagc aagaagatcc
301  taatgattat tgtaaaggag gttatcatct tgtgaaaatt ggagatctat tcaatgggag
361  ataccatgtg atccgaaagt taggctgggg acacttttca acagtatggt tatcatggga
421  tattcagggg aagaaatttg tggcaatgaa agtagttaaa agtgctgaac attacactga
481  aacagcacta gatgaaatcc ggttgctgaa gtcagttcgc aattcagacc ctaatgatcc
541  aaatagagaa atggttgttc aactactaga tgactttaaa atatcaggag ttaatggaac
601  acatatctgc atggtatttg aagttttggg gcatcatctg ctcaagtgga tcatcaaatc
661  caattatcag gggcttccac tgccttgtgt caaaaaaatt attcagcaag tgttacaggg
721  tcttgattat ttacatacca agtgccgtat catccacact gacattaaac cagagaacat
781  cttattgtca gtgaatgagc agtacattcg gaggctggct gcagaagcaa cagaatggca
841  gcgatctgga gctcctccgc cttccggatc tgcagtcagt actgctcccc agcctaaacc
901  agctgacaaa atgtcaaaga ataagaagaa gaaattgaag aagaagcaga agcgccaggc
961  agaattacta gagaagcgaa tgcaggaaat tgaggaaatg gagaaagagt cgggccctgg
1021 gcaaaaaaga ccaaacaagc aagaagaatc agagagtcct gttgaaagac ccttgaaaga
1081 gaacccacct aataaaatga cccaagaaaa acttgaagag tcaagtacca ttggccagga
1141 tcaaacgctt atggaacgtg atacagaggg tggtgcagca gaaattaatt gcaatggagt
1201 gattgaagtc attaattata ctcagaacag taataatgaa acattgagac ataaagagga
1261 tctacataat gctaatgact gtgatgtcca aaatttgaat caggaatcta gtttcctaag
1321 ctcccaaaat ggagacagca gcacatctca agaaacagac tcttgtacac ctataacatc
1381 tgaggtgtca gacaccatgg tgtgccagtc ttcctcaact gtaggtcagt cattcagtga
1441 acaacacatt agccaacttc aagaaagcat tcgggcagag ataccctgtg aagatgaaca
1501 agagcaagaa cataacggac cactggacaa caaaggaaaa tccacggctg gaaattttct
1561 tgttaatccc cttgagccaa aaaatgcaga aaagctcaag gtgaagattg ctgaccttgg
1621 aaatgcttgt tgggtgcaca aacatttcac tgaagatatt caaacaaggc aatatcgttc
1681 cttggaagtt ctaatcggat ctggctataa tacccctgct gacatttgga gcacggcatg
1741 catggccttt gaactggcca caggtgacta tttgtttgaa cctcattcag gggaagagta
1801 cactcgagat gaagatcaca ttgcattgat catagaactt ctggggaagg tgcctcgcaa
1861 gctcattgtg gcaggaaaat attccaagga atttttcacc aaaaaaggtg acctgaaaca
1921 tatcacgaag ctgaaacctt ggggcctttt tgaggttcta gtggagaagt atgagtggtc
1981 gcaggaagag gcagctggct tcacagattt cttactgccc atgttggagc tgatccctga
2041 gaagagagcc actgccgccg agtgtctccg gcacccttgg cttaactcct aa
```

SEQ ID NO: 19
B)

MERKVLALQARKKRTKAKKDKAQRKSETQHRGSAPHSESDLPEQEEEILGSDDDEQE
DPNDYCKGGYHLVKIGDLFNGRYHVIRKLGWGHFSTVWLSWDIQGKKFVAMKVVKSA
EHYTETALDEIRLLKSVRNSDPNDPNREMVVQLLDDFKISGVNGTHICMVFEVLGHHLL
KWIIKSNYQGLPLPCVKKIIQQVLQGLDYLHTKCRIIHTDIKPENILLSVNEQYIRRLAAEAT
EWQRSGAPPPSGSAVSTAPQPKPADKMSKNKKKKLKKKQKRQAELLEKRMQEIEEME
KESGPGQKRPNKQEESESPVERPLKENPPNKMTQEKLEESSTIGQDQTLMERDTEGG
AAEINCNGVIEVINYTQNSNNETLRHKEDLHNANDCDVQNLNQESSFLSSQNGDSSTS

QETDSCTPITSEVSDTMVCQSSSTVGQSFSEQHISQLQESIRAEIPCEDEQEQEHNGPL
DNKGKSTAGNFLVNPLEPKNAEKLKVKIADLGNACWVHKHFTEDIQTRQYRSLEVLIGS
GYNTPADIWSTACMAFELATGDYLFEPHSGEEYTRDEDHIALIIELLGKVPRKLIVAGKY
SKEFFTKKGDLKHITKLKPWGLFEVLVEKYEWSQEEAAGFTDFLLPMLELIPEKRATAAE
CLRHPWLNS

Figura 16a [SEQ ID NO: 20]

secuencia de ARNm:

```
1    gtgcatgacc cgccccgcgg cggagacgcg ctcgctgcgt catcagtgtt ttcgagacga
61   gtctcgacgc agcagctgtc agctccattt tgttgttggt gcgcgacgca gtcagctgcg
121  tgattcccgt gattgcgtta caagctttgt ctccttcgac ttggagtctt tgtccaggac
181  gatgagacac tcaaagagaa cttactgtcc tgattgggat gacaaggatt gggattatgg
241  aaaatggagg agcagcagca gtcataaaag aaggaagaga tcacatagca gtgcccagga
301  gaacaagcgc tgcaaataca atcactctaa aatgtgtgat agccattatt tggaaagcag
361  gtctataaat gagaaagatt atcatagtcg acgctacatt gatgagtaca gaaatgacta
421  cactcaagga tgtgaacctg gacatcgcca aagagaccat gaaagccggt atcagaacca
481  tagtagcaag tcttctggta gaagtggaag aagtagttat aaaagcaaac acaggattca
541  ccacagtact tcacatcgtc gttcacatgg gaagagtcac cgaaggaaaa gaaccaggag
601  tgtagaggat gatgaggagg gtcacctgat ctgtcagagt ggagacgtac taagtgcaag
661  atatgaaatt gttgatactt taggtgaagg agcttttgga aaagttgtgg agtgcatcga
721  tcataaagcg ggaggtagac atgtagcagt aaaaatagtt aaaaatgtgg atagatactg
781  tgaagctgct cgctcagaaa tacaagttct ggaacatctg aatacaacag accccaacag
841  tactttccgc tgtgtccaga tgttggaatg gtttgagcat catggtcaca tttgcattgt
901  ttttgaacta ttgggactta gtacttacga cttcattaaa gaaaatggtt ttctaccatt
961  tcgactggat catatcagaa agatggcata tcagatatgc aagtctgtga atttttgca
1021 cagtaataag ttgactcaca cagacttaaa gcctgaaaac atcttatttg tgcagtctga
1081 ctacacagag gcgtataatc ccaaaataaa acgtgatgaa cgcaccttaa taaatccaga
1141 tattaaagtt gtagactttg gtagtgcaac atatgatgac gaacatcaca gtacattggt
1201 atctacaaga cattatagag cacctgaagt tattttagcc ctagggtggt cccaaccatg
1261 tgatgtctgg agcataggat gcattcttat tgaatactat cttgggttta ccgtatttcc
1321 aacacacgat agtaaggagc atttagcaat gatggaaagg attcttggac ctctaccaaa
1381 acatatgata cagaaaacca ggaaacgtaa atattttcac cacgatcgat tagactggga
1441 tgaacacagt tctgccggca gatatgtttc aagacgctgt aaacctctga aggaatttat
1501 gctttctcaa gatgttgaac atgagcgtct ctttgacctc attcagaaaa tgttggagta
1561 tgatccagcc aaaagaatta ctctcagaga agccttaaag catcctttct ttgaccttct
1621 gaagaaaagt atatagatct gtaattggac agctctctcg aagagatctt acagactgta
1681 tcagtctaat ttttaaattt taagttattt tgtacagctt tgtaaattct taacattttt
1741 atattgccat gtttattttg tttgggtaat ttggttcatt aagtacatag ctaaggtaat
1801 gaacatcttt ttcagtaatt gtaaagtgat ttattcagaa taaatttttt gtgcttatga
1861 agttgatatg tatctgaaca gtttgttcta agtaccattt ttcttcctac ttctattaaa
1921 gaatggacat aga
```

Figura 16b [SEQ ID NO: 21]

MRHSKRTYCPDWDDKDWDYGKWRSSSSHKRRKRSHSSAQENKRCKYNHSKMCDS
HYLESRSINEKDYHSRRYIDEYRNDYTQGCEPGHRQRDHESRYQNHSSKSSGRSGRS
SYKSKHRIHHSTSHRRSHGKSHRRKRTRSVEDDEEGHLICQSGDVLSARYEIVDTLGE
GAFGKVVECIDHKAGGRHVAVKIVKNVDRYCEAARSEIQVLEHLNTTDPNSTFRCVQM
LEWFEHHGHICIVFELLGLSTYDFIKENGFLPFRLDHIRKMAYQICKSVNFLHSNKLTHTD
LKPENILFVQSDYTEAYNPKIKRDERTLINPDIKVVDFGSATYDDEHHSTLVSTRHYRAP
EVILALGWSQPCDVWSIGCILIEYYLGFTVFPTHDSKEHLAMMERILGPLPKHMIQKTRK
RKYFHHDRLDWDEHSSAGRYVSRRCKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRI
TLREALKHPFFDLLKKSI

Figura 17a [SEQ ID NO: 22]

Secuencia de ARNm (secuencia codificante 30-1022 nt)

```
1    gggcaaggag ctgctggctg gacggcggca ggtccgacag cgagaagctc aacctggact
61   cgatcatcgg gcgcctgctg gaagtgcagg gctcgcggcc tggcaagaat gtacagctga
121  cagagaacga gatccgcggt ctgtgcctga aatcccggga gatttttctg agccagccca
181  ttcttctgga gctggaggca cccctcaaga tctgcggtga catacacggc cagtactacg
241  accttctgcg actatttgag tatggcggtt tccctcccga gagcaactac ctctttctgg
301  gggactatgt ggacaggggc aagcagtcct tggagaccat ctgcctgctg ctggcctata
361  agatcaagta ccccgagaac ttcttcctgc tccgtgggaa ccacgagtgt gccagcatca
421  accgcatcta tggtttctac gatgagtgca agagacgcta caacatcaaa ctgtggaaaa
481  ccttcactga ctgcttcaac tgcctgccca tcgcggccat agtggacgaa aagatcttct
541  gctgccacgg aggcctgtcc ccggacctgc agtctatgga gcagattcgg cggatcatgc
601  ggcccacaga tgtgcctgac cagggcctgc tgtgtgacct gctgtggtct gaccctgaca
661  aggacgtgca gggctggggc gagaacgacc gtggcgtctc ttttaccttt ggagccgagg
721  tggtggccaa gttcctccac aagcacgact tggacctcat ctgccgagca caccaggtgg
781  tagaagacgg ctatgagttc tttgccaagc ggcagctggt gacacttttc tcagctccca
841  actactgtgg cgagtttgac aatgctggcg ccatgatgag tgtggacgag accctcatgt
901  gctctttcca gatcctcaag cccgccgaca agaacaaggg gaagtacggg cagttcagtg
961  gcctgaaccc tggaggccga cccatcaccc caccccgcaa ttccgccaaa gccaagaaag
1021 agcccccgca caccaccctg tgccccagat gatggattga ttgtacagaa atcatgctgc
1081 catgctgggg gggggtcacc ccgaccccta aggcccacct gtcacgggga acatggagcc
1141 ttggtgtatt tttcttttct ttttttaatg aatcaatagc agcgtccagt cccccagggc
1201 tgcttcctgc ctgcacctgc ggtactgtga gcaggatcct ggggccgagg ctgcagctca
1261 gggcaacggc aggccaggtc gtgggtctcc agccgtgctt ggcctcaggc tggcagcccg
1321 gatcctgggg caacccatct ggtctcttga ataaaggtca aagctgg
```

Figura 17b (SEQ ID NO 30)

MSDSEKLNLDSIIGRLLEVQGSRPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPLKIC
GDIHGQYYDLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFLLR
GNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
MEQIRRIMRPTDVPDQGLLCDLLWSDPDKDVQGWGENDRGVSFTFGAEVVAKFLHKH
DLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFDNAGAMMSVDETLMCSFQILK
PADKNKGKYGQFSGLNPGGRPITPPRNSAKAKK

Figura 18 (SEQ ID NO 31)

A

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAG
CCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCA
CTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGA
CGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTA
CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCC
ACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTATTTA
TTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGCCCC
CCCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAAAGG
TGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGG
CGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCG

Figura 18 continuación

B (SEQ ID NO 32)

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
TTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT
GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCAT
AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAA
CTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC
GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA
CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG
GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA
GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC
TTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT
ACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTT
ACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGT
TTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTG

C) (SEQ ID NO 33)

TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCC
CCAATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGG
GGGGGGGGGCCCCCCCCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGG
GCGAGGCGGAAAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTT
CCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGG
CGGGCG

D) (SEQ ID NO 34)

CTGCAGAGGGCCCTGCGTATGAGTGCAAGTGGGTTTTAGGACCAGGATGAGGCGG
GGTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCCAACC
CCCATTCCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGATGC
GGCGAGGCGCGTGCGCACTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCCCC
CGCCTGGCGGCGCGCGCCACCGCCGCCTCAGCACTGAAGGCGCGCTGACGTCAC
TCGCCGGTCCCCCGCAAACTCCCCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGC
CGCCGGCCCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGGCACGGGC
GCGACCATCTGCGCTGCGGCG

E) (SEQ ID NO 35)

GTCTGCAAGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCCAGGGCCTCCTCTTC
ATGCCCAGTGAATGACTCACCTTGGCACAGACACAATGTTCGGGGTGGGCACAGT
GCCTGCTTCCCGCCGCACCCCAGCCCCCCTCAAATGCCTTCCGAGAAGCCCATTG
AGTAGGGGGCTTGCATTGCACCCCAGCCTGACAGCCTGGCATCTTGGGATAAAAG
CAGCACAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGAGA
ACAAGGCTCTATTCAGCCTGTGCCCAGGAAAGGGGATCAGGGGATGCCCAGGCAT
GGACAGTGGGTGGCAGGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGAAGTCCAA
GGACACAAATGGGTGAGGGGACTGGGCAGGGTTCTGACCCTGTGGGACCAGAGT

GGAGGGCGTAGATGGACCTGAAGTCTCCAGGGACAACAGGGCCCAGGTCTCAGG
CTCCTAGTTGGGCCCAGTGGCTCCAGCGTTTCCAAACCCATCCATCCCCAGAGGTT
CTTCCCATCTCTCCAGGCTGATGTGTGGGAACTCGAGGAAATAAATCTCCAGTGGG
AGACGGAGGGGTGGCCAGGGAAACGGGGCGCTGCAGGAATAAAGACGAGCCAGC
ACAGCCAGCTCATGCGTAACGGCTTTGTGGAGCTGTCAAGGCCTGGTCTCTGGGA
GAGAGGCACAGGGAGGCCAGACAAGGAAGGGGTGACCTGGAGGGACAGATCCAG
GGGCTAAAGTCCTGATAAGGCAAGAGAGTGCCGGCCCCCTCTTGCCCTATCAGGA
CCTCCACTGCCACATAGAGGCCATGATTGACCCTTAGACAAAGGGCTGGTGTCCAA
TCCCAGCCCCCAGCCCCAGAACTCCAGGGAATGAATGGGCAGAGAGCAGGAATGT
GGGACATCTGTGTTCAAGGGAAGGACTCCAGGAGTCTGCTGGGAATGAGGCCTAG
TAGGAAATGAGGTGGCCCTTGAGGGTACAGAACAGGTTCATTCTTCGCCAAATTCC
CAGCACCTTGCAGGCACTTACAGCTGAGTGAGATAATGCCTGGGTTATGAAATCAA
AAAGTTGGAAAGCAGGTCAGAGGTCATCTGGTACAGCCCTTCCTTCCCTTTTTTTTT
TTTTTTTTTTTTTGTGAGACAAGGTCTCTCTCTGTTGCCCAGGCTGGAGTGGCGCAA
ACACAGCTCACTGCAGCCTCAACCTACTGGGCTCAAGCAATCCTCCAGCCTCAGCC
TCCCAAAGTGCTGGGATTACAAGCATGAGCCACCCCACTCAGCCCTTTCCTTCCTT
TTTAATTGATGCATAATAATTGTAAGTATTCATCATGGTCCAACCAACCCTTTCTTGA
CCCACCTTCCTAGAGAGAGGGTCCTCTTGATTCAGCGGTCAGGGCCCCAGACCCA
TGGTCTGGCTCCAGGTACCACCTGCCTCATGCAGGAGTTGGCGTGCCCAGGAAGC
TCTGCCTCTGGGCACAGTGACCTCAGTGGGGTGAGGGGAGCTCTCCCCATAGCTG
GGCTGCGGCCCAACCCCACCCCCTCAGGCTATGCCAGGGGGTGTTGCCAGGGGC
ACCCGGGCATCGCCAGTCTAGCCCACTCCTTCATAAAGCCCTCGCATCCCAGGAG
CGAGCAGAGCCAGAGCAT

6F) (SEQ ID NO 36)
GAATCCTATGCTTCGAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCC
CAGTGTCACTAGGCGGGAACACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGC
TGTGAGGGACAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTG
GGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAGTCTTATAAGTTCTGTAT
GAGACCACTCTTTCCCATGGTCTTCATCCTATCTAGACA

G) (SEQ ID NO 37)
GCCGCCCCCTTCACCGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGA
TACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTA
GTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT
TATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTT
GGCTTTATATATCTTGTGGAAAGGACGAAAC

h) promotor H1_pLVTHM (SEQ ID NO 143)
GAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCCCAGTGTCACTA
GGCGGGAACACCCAGCGCGGTGCGCCCTGGCAGGAAGATGGCTGTGAGGGA
CAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATC

ACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAG
ACCAC

j) promotor H1 en cadena principal de scAAV (SEQ ID NO 144)

ATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCT
TTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCAC

k) promotor U6 (SEQ ID NO 145)

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGT
TAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAA
AATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATT
ATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATT
TCTTGGCTTTATATATCTTGTGGAAAGGAC

l) promotor β-actina de pollo (CBA) (SEQ ID NO 146)

TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCA
CCCCCAATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCG
GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCG
GGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCT
CCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAG
CGAAGCGCGCGGCGGGCG

m) promotor CAG (promotor β-actina de pollo híbrido) (SEQ ID NO 147)

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTT
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTT
CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT
ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACG
CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT
ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTC
CCCCCCCTCCCCACCCCCAATTTTGTATTTATTTATTTTTTAATTATTTTGTGCA
GCGATGGGGGCGGGGGGGGGGGGGGGCCCCCCCCCAGGCGGGGCGGGGCG
GGGCGAGGGGCGGGGCGGGGCGAGGCGGAAAGGTGCGGCGGCAGCCAATC
AGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCG
GCCCTATAAAAAGCGAAGCGCGCGGCGGGCG

n) PKG de ratón_promotor de LV-LacZ (SEQ ID NO 148)

GGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTTTAGCAGC
CCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCC
ACATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCGCC
ACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCGCCCCGCAGCTCGCGT
CGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCACTAGTCTCGTGCAGAT
GGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGGCCA
ATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGT

CCGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGT
CCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTCTGCCGCGCT
GTTCTCCTCTTCCTCATCTCCGGGCCTTTCG

o) Factor de elongación corto 1α (EFS1α) (SEQ ID NO 149)

GGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCG
GCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTG
ATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATA
AGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAAC
ACAG

p) Potenciador CMV (SEQ ID NO 150)

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTT
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTT
CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT
ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACG
CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT
ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATG

q) Promotor CMV (SEQ ID NO 151)

GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAC
GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCAC
CAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC
AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT

Figura 19
A)

(SEQ ID NO 38)
AAAGCTTCCCGGGGGGATCTGGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACT
GAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCA
GTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCC
TGGAGGGGTGGAGTCGTGACCCCTAAAATGGGCAAACATTGCGCTAGCGCGGCCG
CGGTACCTACCGATGTACGGGCCAGATATACGCGTGGAGCTAGTTATTAATAGTAA
TCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTT
ACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAA
TAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGG
TGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAA
GTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG
TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT
ATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC
ACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAA
ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGA

ACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACAC
CGGGACCGATCCAGCCTGAATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCC
ACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC
CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC
GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCA
CCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCA
ACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG
GAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGG
CATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCC
GACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGC
GCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAAAGCGGCCGCGACTCTAGAGTCGACCTGCAGGGAAGA
CCAAGCTGACCTGACTCGATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT
GTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGT
TTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAACTAGTCCACTCCCTCTCTGC
GCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCT
TTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGACAGATCCGGG
CCCGCATGCGTCGACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAA
CCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAA
TGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCAC
ACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAG
CCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCG
GCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGT
TTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTT
TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATC
CGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTA
TGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCC
TGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGT
TTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACT
ATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATG
GCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCG
GCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCA
CAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAA
GCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGT
TGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAG
ACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGC

TGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA
TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACG
GGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCT
CACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGA
TTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCA
TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAA
AAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAA
ACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC
TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCT
AGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTT
ACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAA
CGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAG
ATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCG
GACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACT
TGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCC
AGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTC
TTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCT
GATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAA
GCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATT
AATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGC
AATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCC
GGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCT
ATGACCATGATTACGCCAAGCTCTCGAGATCTAG

B) (SEQ ID NO 39)
ATGCCGGGGTTTTACGAGATTGTGATTAAGGTCCCCAGCGACCTTGACGGGCATCT
GCCCGGCATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG
CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTGACCGTGG
CCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGTGTGAGTAAGGCCCC
GGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGAGAGAGCTACTTCCACATGCACG
TGCTCGTGGAAACCACCGGGGTGAAATCCATGGTTTTGGGACGTTTCCTGAGTCAG
ATTCGCGAAAAACTGATTCAGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAAA
CTGGTTCGCGGTCACAAAGACCAGAAATGGCGCCGGAGGCGGGAACAAGGTGGT
GGATGAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCCAGT
GGGCGTGGACTAATATGGAACAGTATTTAAGCGCCTGTTTGAATCTCACGGAGCGT
AAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCAGACGCAGGAGCAGAACA
AAGAGAATCAGAATCCCAATTCTGATGCGCCGGTGATCAGATCAAAAACTTCAGCC
AGGTACATGGAGCTGGTCGGGTGGCTCGTGGACAAGGGGATTACCTCGGAGAAGC
AGTGGATCCAGGAGGACCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAACTCG
CGGTCCCAAATCAAGGTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAAA
ACCGCCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTCCAGCAATC
GGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCAATATGCGGCTTCCGTCT

TTCTGGGATGGGCCACGAAAAAGTTCGGCAAGAGGAACACCATCTGGCTGTTTGG
GCCTGCAACTACCGGGAAGACCAACATCGCGGAGGCCATAGCCCACACTGTGCCC
TTCTACGGGTGCGTAAACTGGACCAATGAGAACTTTCCCTTCAACGACTGTGTCGA
CAAGATGGTGATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCGTGGAGTC
GGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCAAGTCC
TCGGCCCAGATAGACCCGACTCCCGTGATCGTCACCTCCAACACCAACATGTGCG
CCGTGATTGACGGGAACTCAACGACCTTCGAACACCAGCAGCCGTTGCAAGACCG
GATGTTCAAATTTGAACTCACCCGCCGTCTGGATCATGACTTTGGGAAGGTCACCA
AGCAGGAAGTCAAAGACTTTTTCCGGTGGGCAAAGGATCACGTGGTTGAGGTGGA
GCATGAATTCTACGTCAAAAAGGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGAC
GCAGATATAAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGT
CAGACGCGGAAGCTTCGATCAACTACGCGGACAGGTACCAAAACAAATGTTCTCGT
CACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGCGAGAGACTGAATCA
GAATTCAAATATCTGCTTCACTCACGGTGTCAAAGACTGTTTAGAGTGCTTTCCCGT
GTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGGCGTATCAGAAACTGTGCTACAT
TCATCACATCATGGGAAAGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATG
TGGACTTGGATGACTGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCC
GATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCGCGAGTG
GTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGACA
ACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTC
GACAAGGGGGAGCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAG
GCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACG
CCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCT
CGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTTG
AGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCA
GGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAAAAG
AGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCAACCAAT
CGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGT
GGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCT
CGGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAG
CACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCA
ACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCC
TGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCA
GCGACTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCT
TCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAAT
AACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGT
GCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATG
ATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTC
GTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACTT
CCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCC
AAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAA
AGACTATTAACGGTTCTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGA
CCCAGCAACATGGCTGTCCAGGGAAGAAACTACATACCTGGACCCAGCTACCGAC
AACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCT

GGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGC
TATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCCTTTGTCTGGATCTTTAAT
TTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAA
CCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAA
GTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAA
ACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGA
CCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCTTCTCCGCTGAT
GGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTG
TACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACC
CAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAA
ACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACTATTACAAGTCTAAT
AATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGC
ACCAGATACCTGACTCGTAATCTGTAATTGCTTGTTAATCAATAAACCGTTTAATTCG
TTTCAGTTGAACTTTGGTCTCTGCGAAGGGCGAATTCGTTTAAACCTGCAGGACTA
GAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCGACACCATGTG
GTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCATTTTGAAGCGG
GAGGTTTGAACGCGCAGCCGCCAAGCCGAATTCTGCAGATATCCATCACACTGGC
GGCCGCTCGACTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTT
TAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG
AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAA
AGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC
CCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACT
CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGT
AATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAG
GCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT
CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
GGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTAT
ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAAC
TACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCG
AGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC

GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATT
GCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGG
TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTT
CTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCG
AGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTT
AAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCA
TCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGC
AAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCA
ATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGT
ATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCT
AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCA
TTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACC
GAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGT
GGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGT
GAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCG
GAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTG
GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGT
GTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTAC
AGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGT
GCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGA
TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAG
TGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTC
GATCGAGGTCGACGGTATCGGGGGAGCTCGGATCCACTAGTAACGGCCGCCAGT
GTGCTGGATTCGGCTTTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCATTTTGAA
GCGGGAGGTTTGAACGCGCAGCCGCC

Figura 20

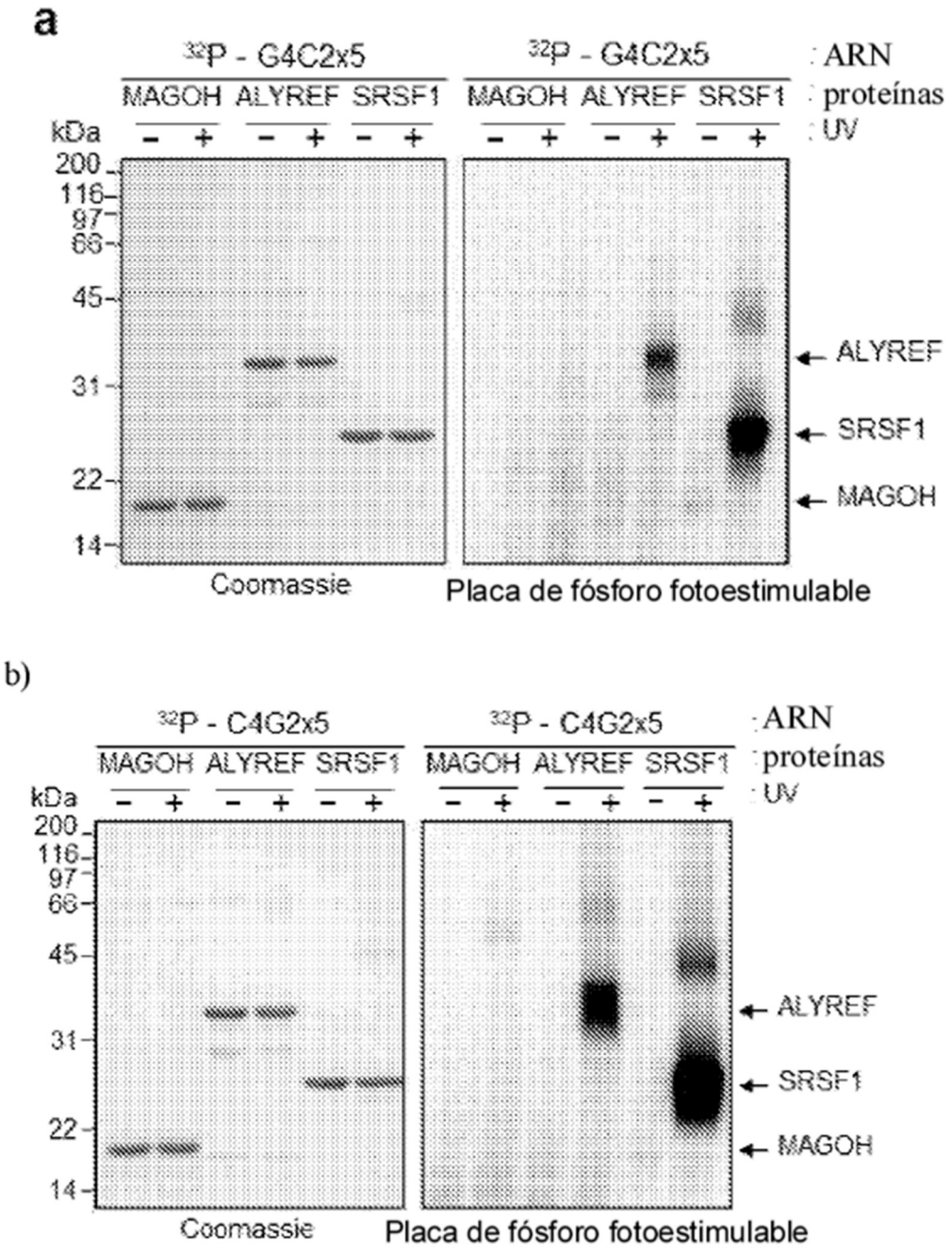
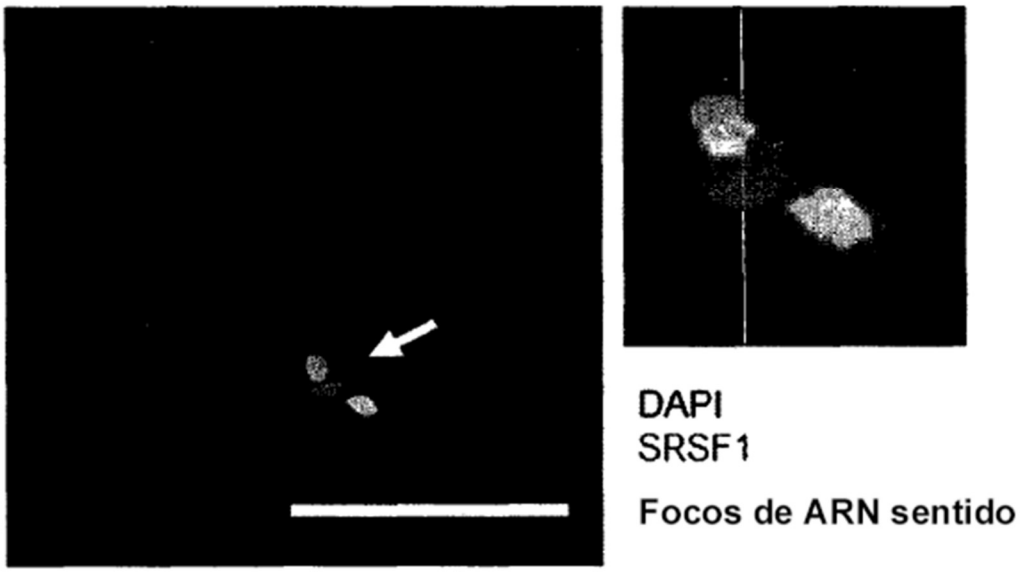

a
32P - G4C2x5
32P - G4C2x5
: ARN
MAGOH ALYREF SRSF1
MAGOH ALYREF SRSF1
: proteínas
kDa
- + - + - +
- + - + - +
: UV
200
116
97
66
45
31
22
14
ALYREF
SRSF1
MAGOH
Coomassie
Placa de fósforo fotoestimulable
b)
32P - C4G2x5
32P - C4G2x5
: ARN
MAGOH ALYREF SRSF1
MAGOH ALYREF SRSF1
: proteínas
kDa
- + - + - +
- + - + - +
: UV
200
116
97
66
45
31
22
14
ALYREF
SRSF1
MAGOH
Coomassie
Placa de fósforo fotoestimulable

Figura 20 continuación

c) Caso 1



Caso 2

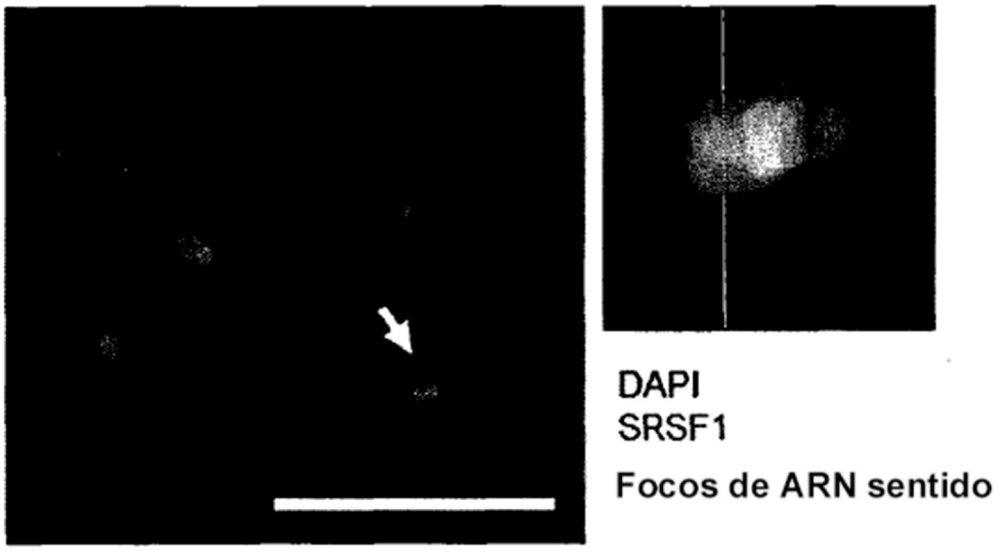
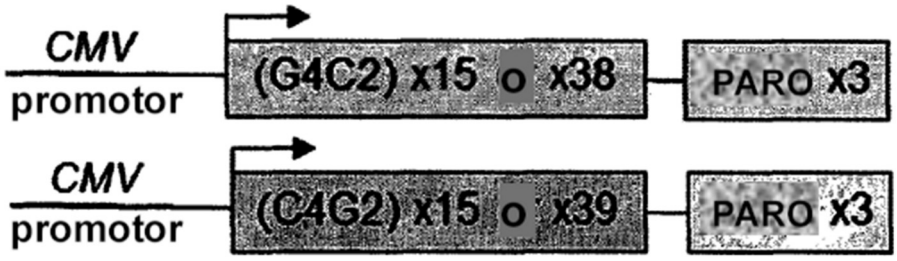
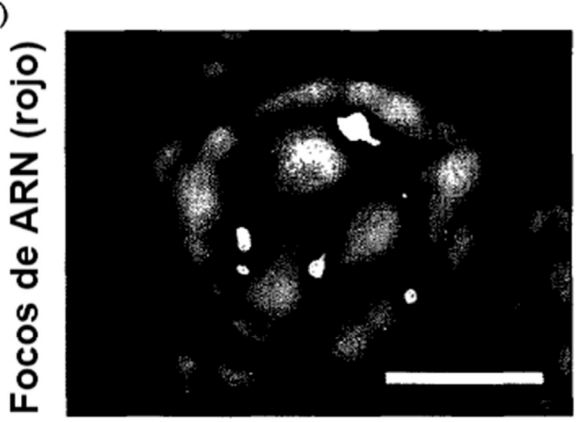


Figura 21

a)



b)



c)

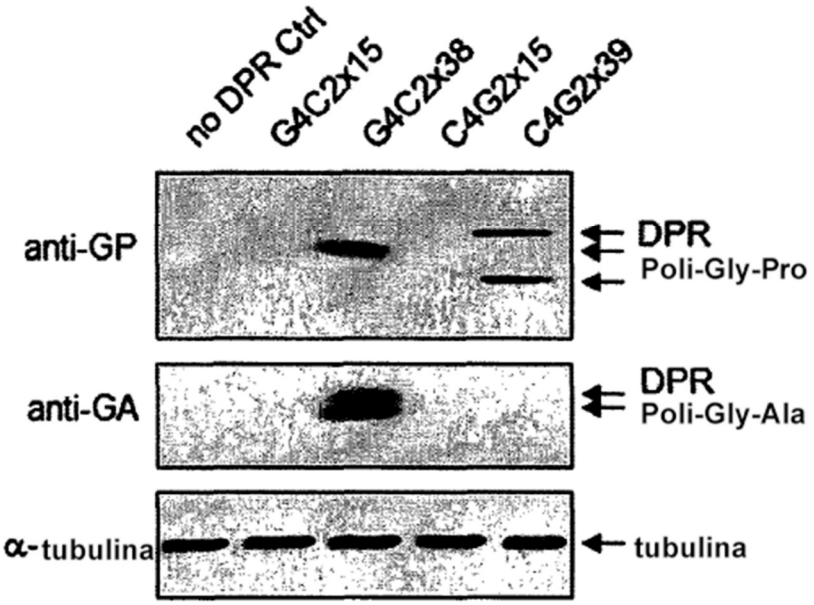


Figura 21 continuación

d)

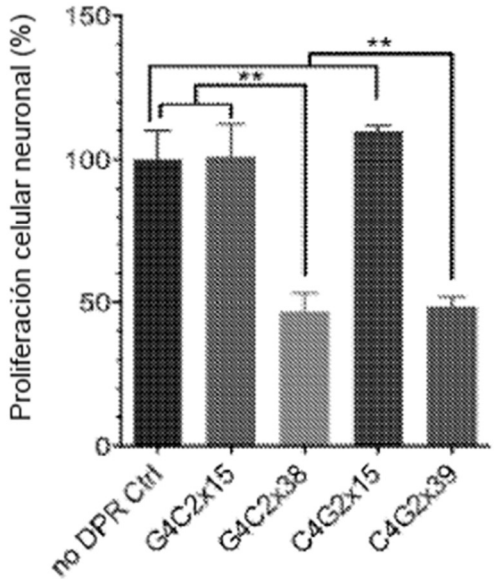


Figura 22

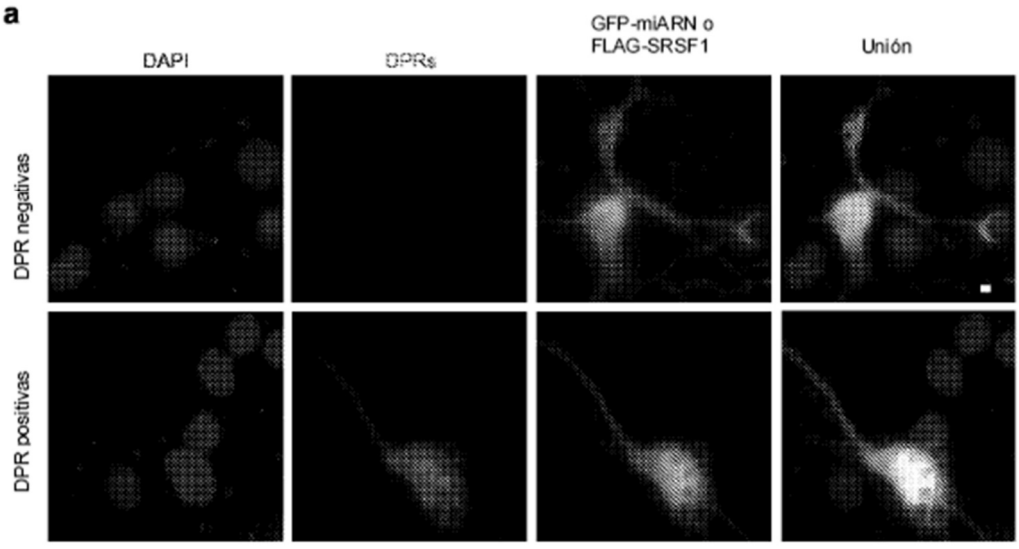
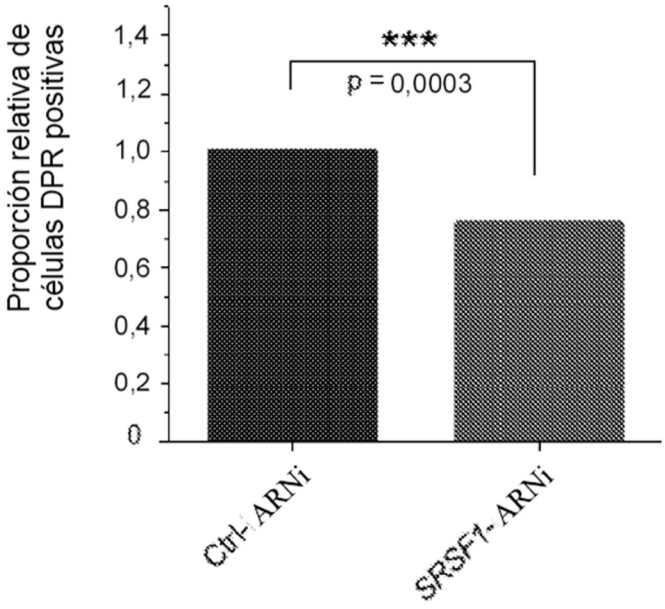


Figura 22 continuación

b)



c)

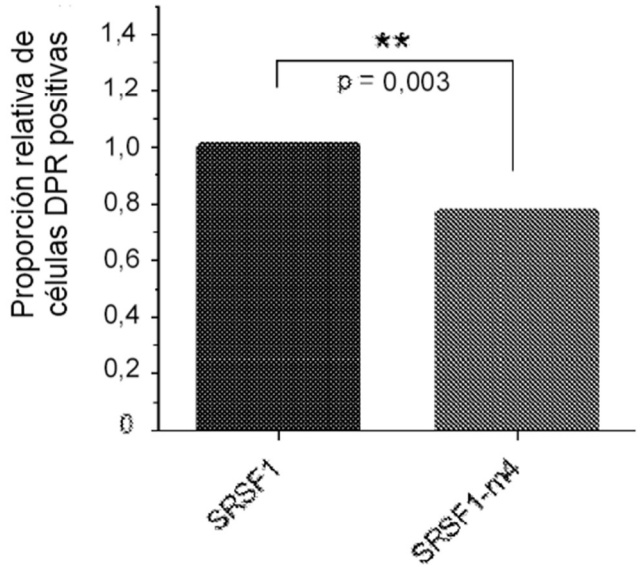
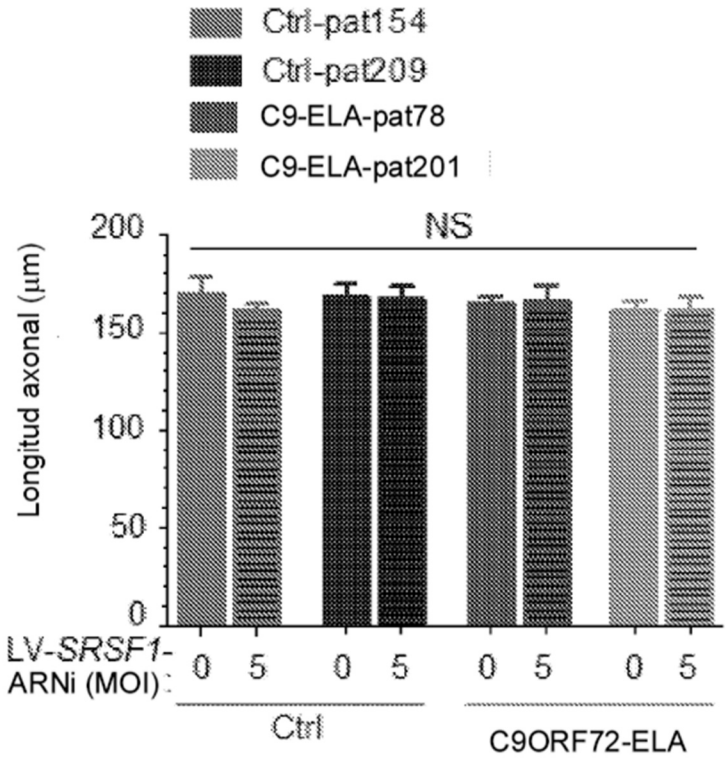


Figura 23

a)


Ctrl-pat154
Ctrl-pat209
C9-ELA-pat78
C9-ELA-pat201
NS
Longitud axonal (µm)
200
150
100
50
0
LV-SRSF1-ARNi (MOI):
0 5 0 5 0 5 0 5
Ctrl
C9ORF72-ELA

b)

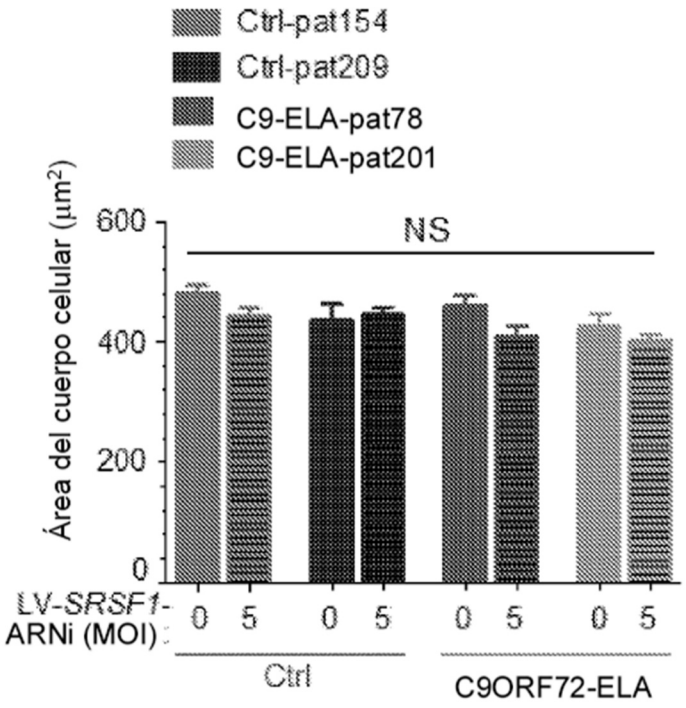

Ctrl-pat154
Ctrl-pat209
C9-ELA-pat78
C9-ELA-pat201
NS
Área del cuerpo celular (µm²)
600
400
200
0
LV-SRSF1-ARNi (MOI):
0 5 0 5 0 5 0 5
Ctrl
C9ORF72-ELA

Figura 23 continuación

c)

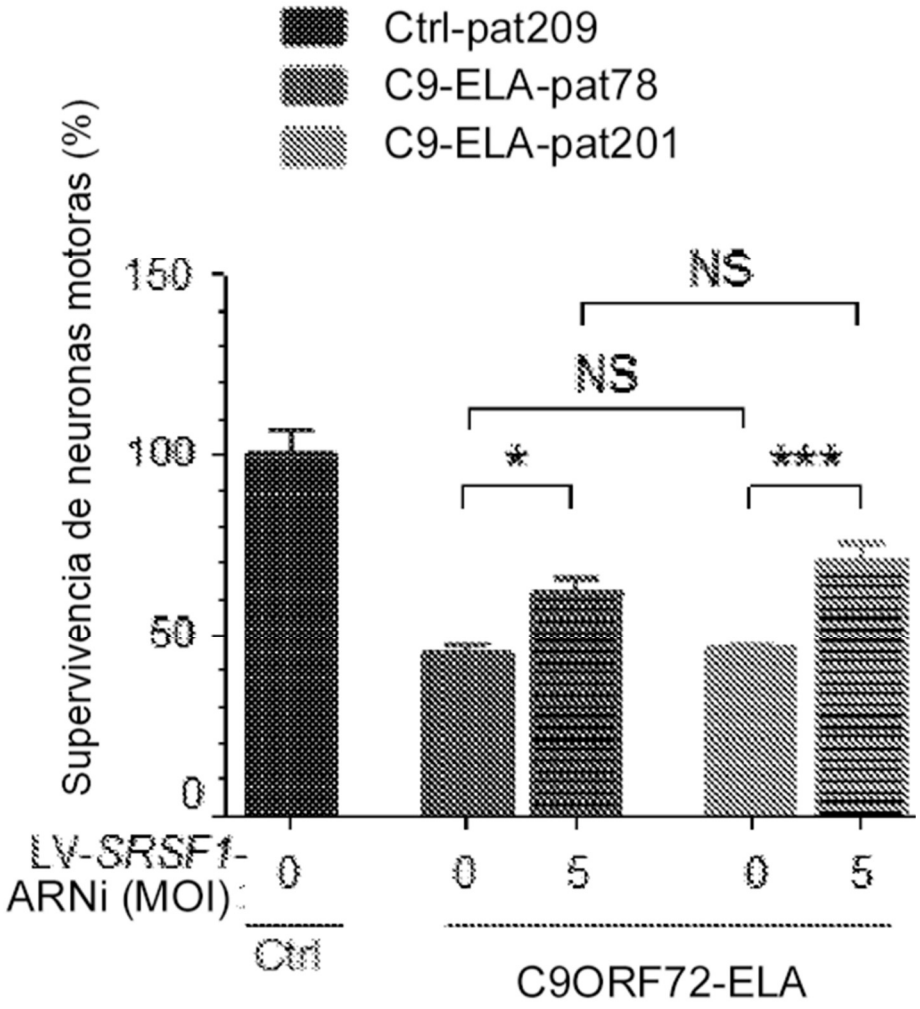


Figura 23 continuación

d)

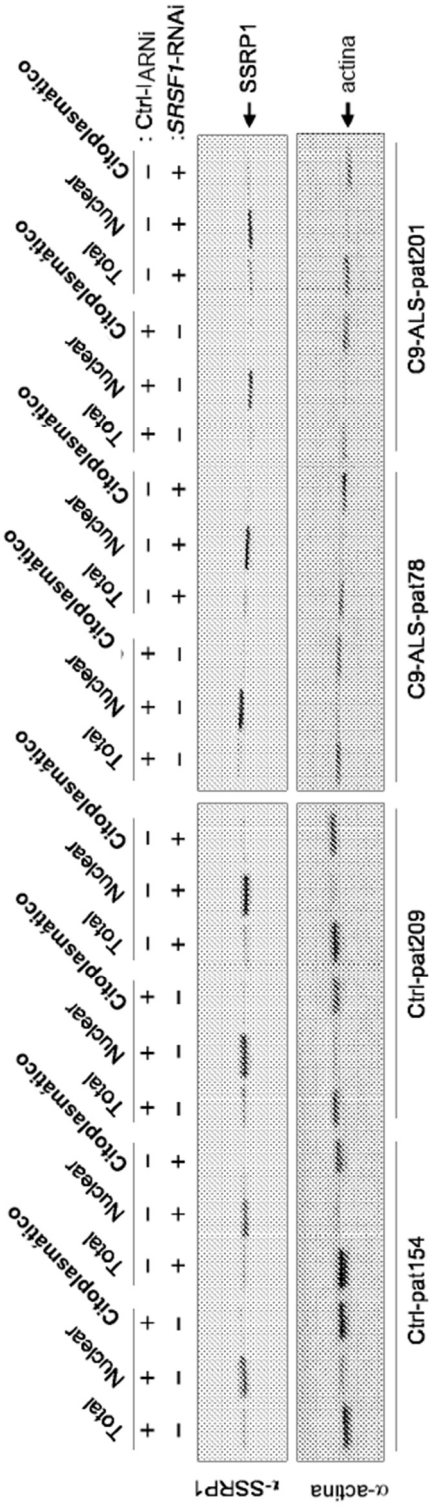
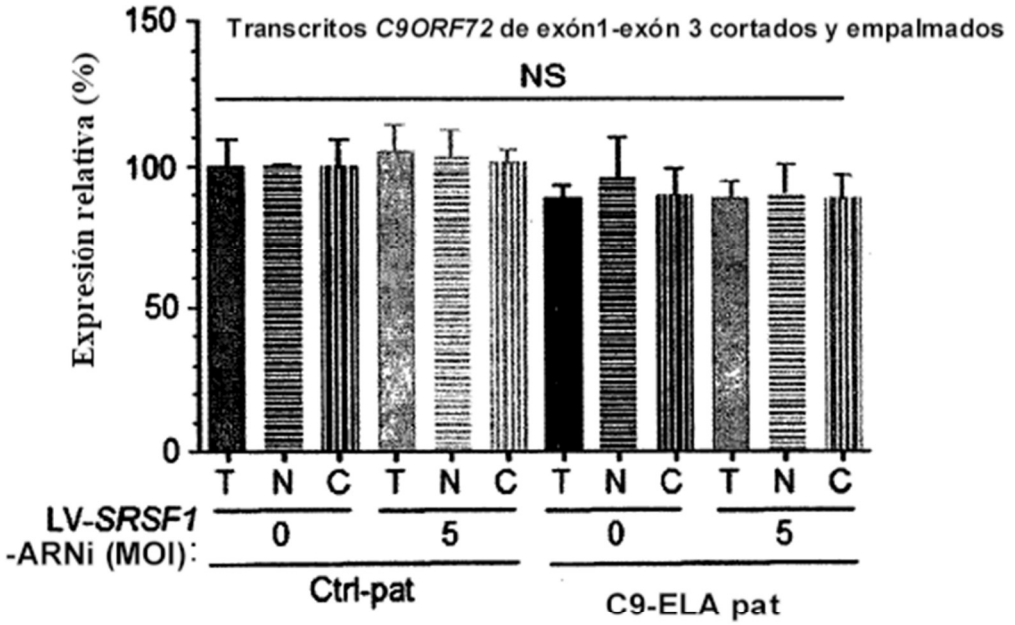


Figura 23 continuación

e)



f

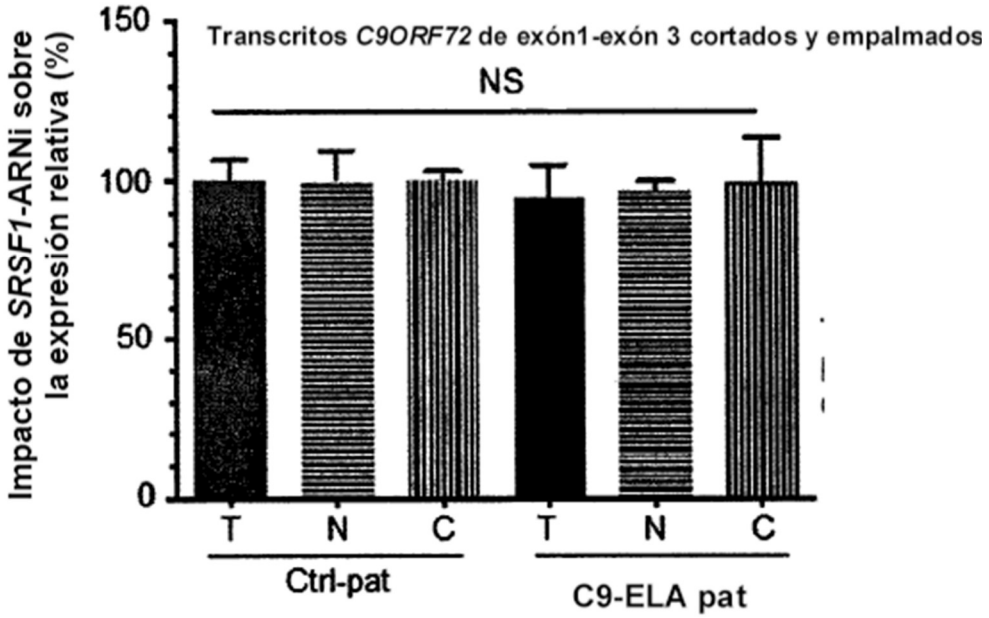


Figura 23 continuación

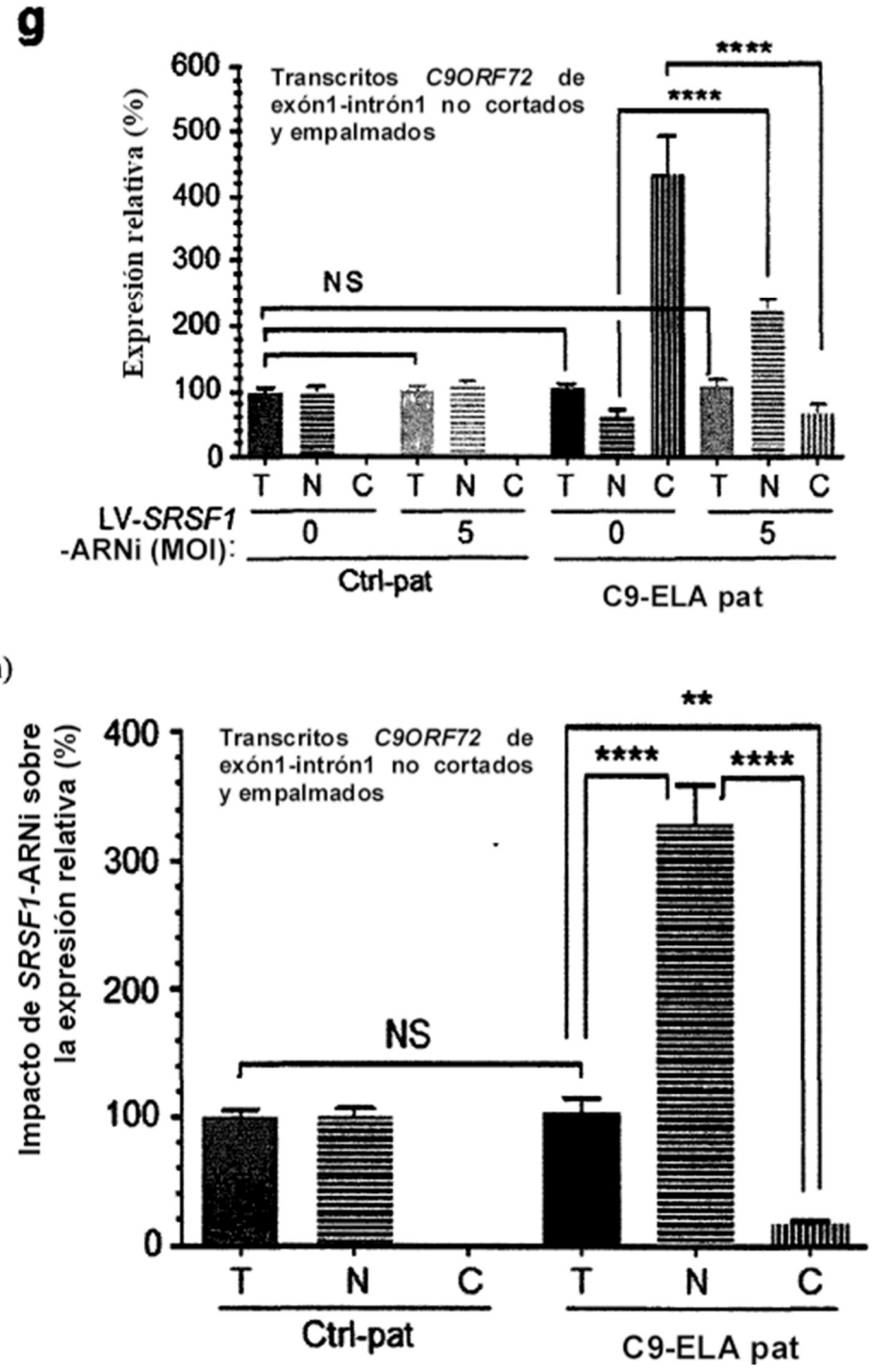


Figura 24

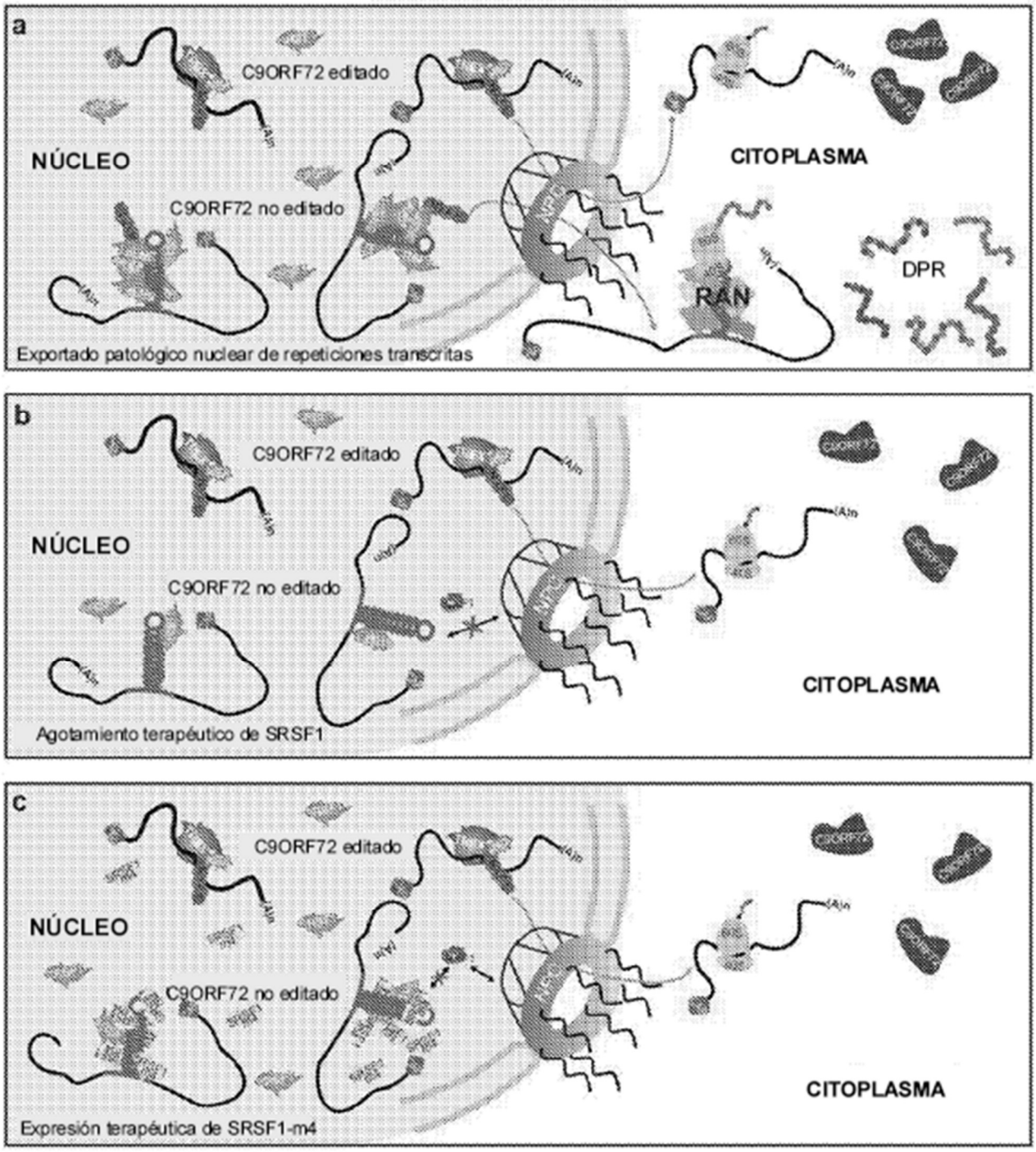
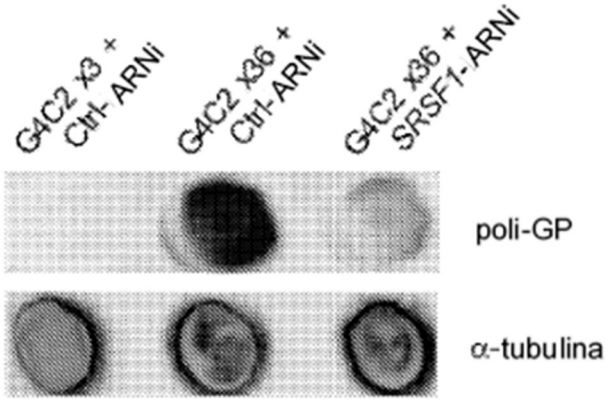

a
C9ORF72 editado
NÚCLEO
C9ORF72 no editado
CITOPLASMA
RAN
DPR
Exportado patológico nuclear de repeticiones transcritas
b
C9ORF72 editado
NÚCLEO
C9ORF72 no editado
CITOPLASMA
Agotamiento terapéutico de SRSF1
c
C9ORF72 editado
NÚCLEO
C9ORF72 no editado
CITOPLASMA
Expresión terapéutica de SRSF1-m4

Figura 25

a)


G4C2 x3 + Ctrl- ARNi
G4C2 x36 + Ctrl-ARNi
G4C2 x36 + SRSF1- ARNi
poli-GP
α-tubulina

b)

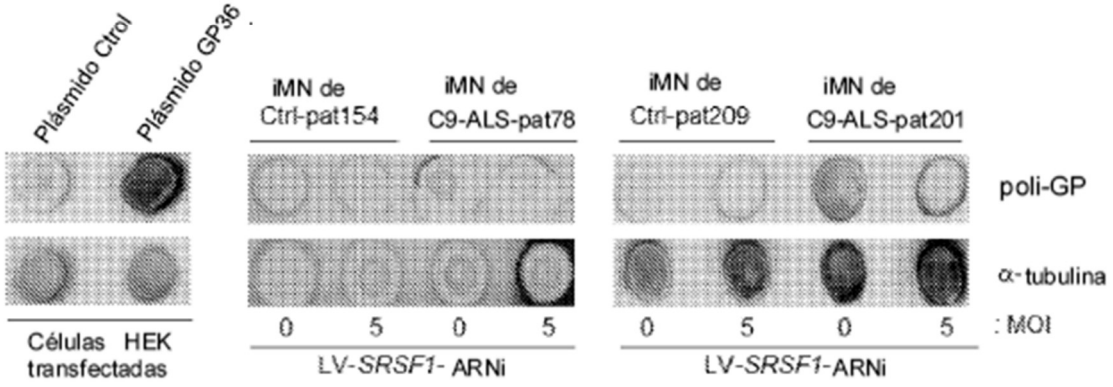

Plásmido Ctrol
Plásmido GP36
iMN de Ctrl-pat154
iMN de C9-ALS-pat78
iMN de Ctrl-pat209
iMN de C9-ALS-pat201
poli-GP
α-tubulina
0 5 0 5
0 5 0 5
: MOI
Células HEK transfectadas
LV-SRSF1- ARNi
LV-SRSF1-ARNi

Figura 26

a

Expresión de poli-Gly-Pro x36 DPR sintéticas independiente de repeticiones G4C2

ATG GGC AAA CCG ATT CCG AAC CCG CTG CTG GGC CTG GAT AGC ACC CTC GAG AAT GAT CCC ACC ATG
GGC CCT GGC CCT GGA CCA GGA CCT GGC CCC GGA CCC GGT CCA GGT CCC GGC CCA GGC CCC GGT CCC
GGC CCT GGA CCA GGC CCA GGA CCA GGA CCA GGC CCA GGT CCC GGA CCA GGA CCC GGA CCT GGC CCA
GGC CCT GGC CCT GGC CCT GGC CCC GGA CCA GGC CCT GGA CCC GGC CCT GGT CCC GGC CCA GGA CCC
GGA CCA GGA CCT GGC CCT **TAA**

b

Expresión de poli-Gly-Ala x36 DPRs sintéticas independiente de repeticiones G4C2

ATG GGC AAA CCG ATT CCG AAC CCG CTG CTG GGC CTG GAT AGC ACC CTC GAG AAT GAT CCC ACC ATG
GGA GCT GGT GCT GGT GCA GGC GCT GGC GCA GGG GCA GGC GCT GGT GCT GGG GCT GGT GCC GGG GCT
GGG GCA GGC GCA GGG GCT GGT GCC GGT GCA GGC GCA GGG GCT GGG GCT GGC GCT GGT GCC GGC GCA
GGC GCG GGT GCC GGC GCA GGG GCT GGT GCA GGG GCC GGT GCT GGC GCG GGT GCA GGG GCC GGT GCA
GGG GCA GGC GCA GGC GCT **TAA**

Figura 27

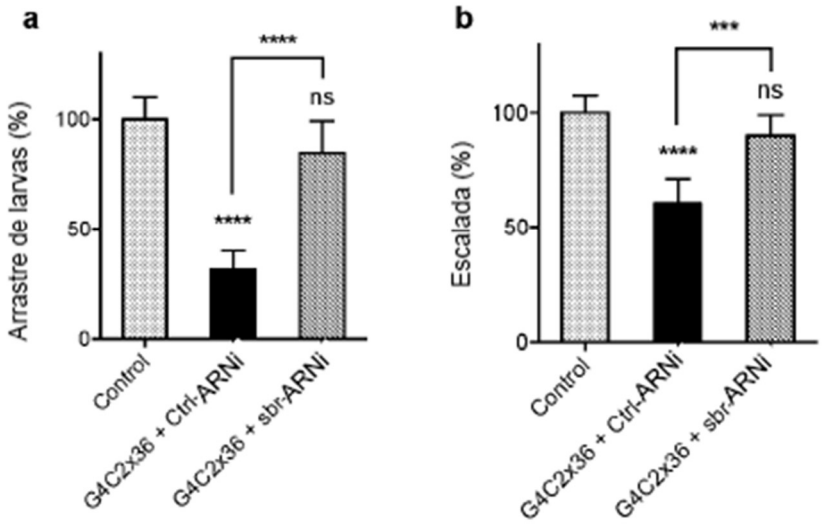
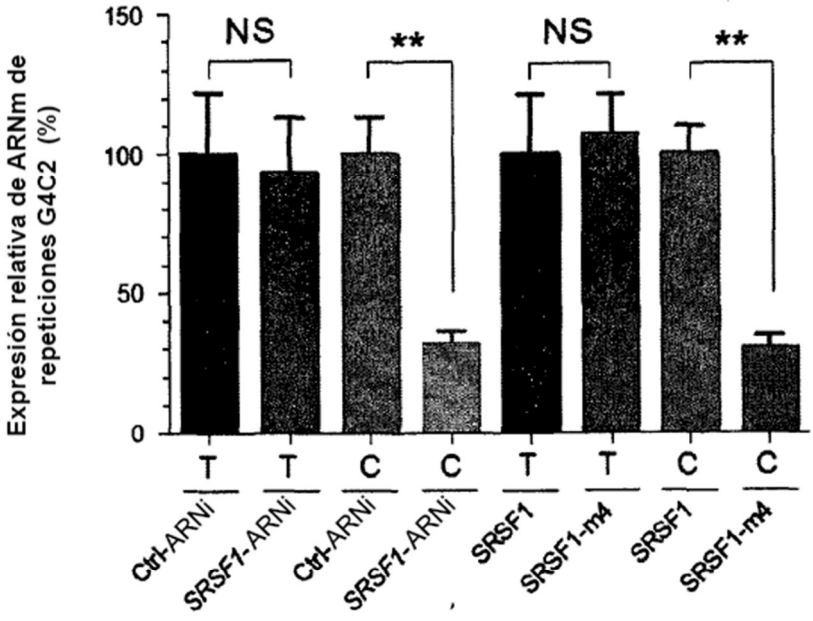


Figura 28

a)


Expresión relativa de ARNm de repeticiones G4C2 (%)
150
100
50
0
NS
**
NS
**
T T C C T T C C
Ctrl-ARNi
SRSF1-ARNi
Ctrl-ARNi
SRSF1-ARNi
SRSF1
SRSF1-m4
SRSF1
SRSF1-m4

b

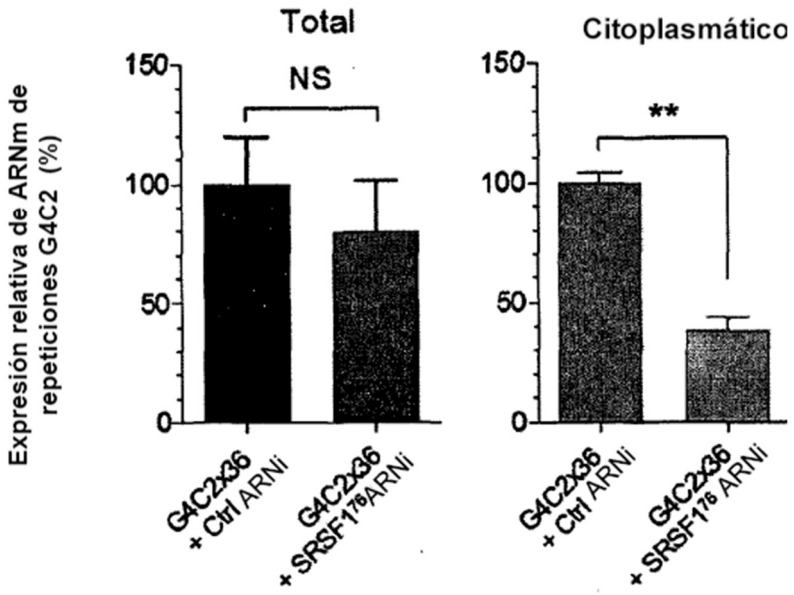

Total
Citoplasmático
Expresión relativa de ARNm de repeticiones G4C2 (%)
150
100
50
0
NS
**
G4C2x36 + Ctrl ARNi
G4C2x36 + SRSF1^76 ARNi
G4C2x36 + Ctrl ARNi
G4C2x36 + SRSF1^76 ARNi

Figura 29

Repeticiones-sentido RAN-G4C2x36 con 3x etiqueta V5

...CTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAA
TTGGGGCCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGC
CGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGAATTCTGCAGATATCCAGCACAGTG GCGGCCGC TG GGCAAGCCCATC
CCCAACCCCCTGCTCGGTCTGGACAGCACCGGCTAA C GGCAAGCCCATCCCCAACCCCCTGCTCGGTCTGGACAGCACCGGCTAA C GGCAAGC
CCATCCCCAACCCCCTGCTCGGTCTGGACAGCACCGGCTAA CTCGAGTCTAGATCTAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATC
TGAATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCT CGA...

Figura 30

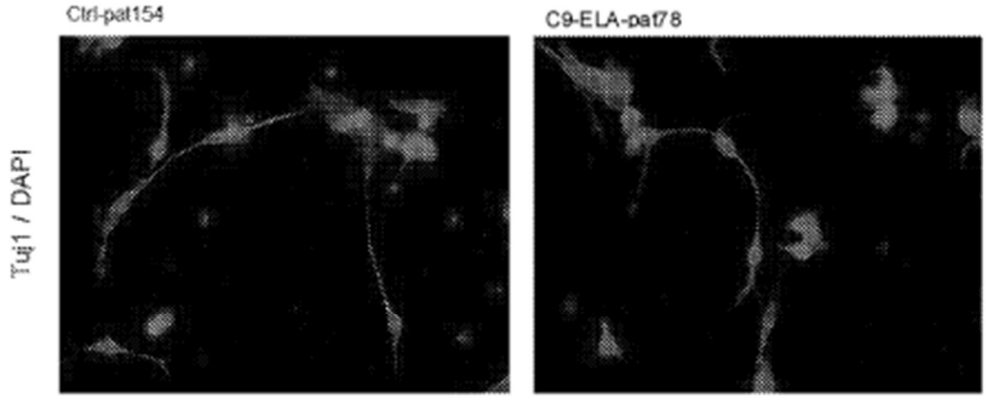


Figura 31

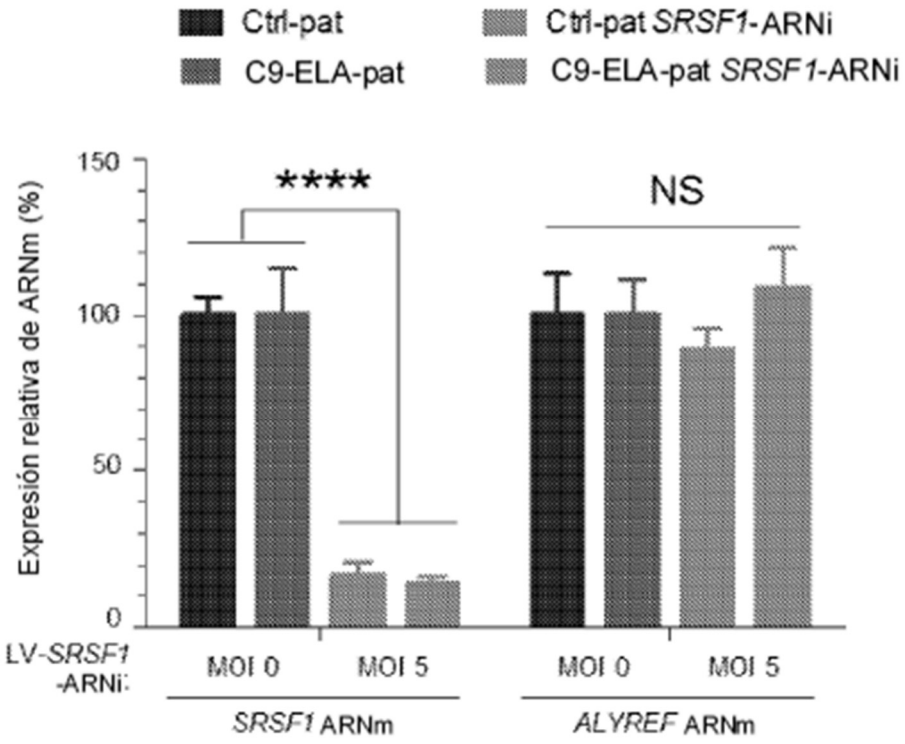
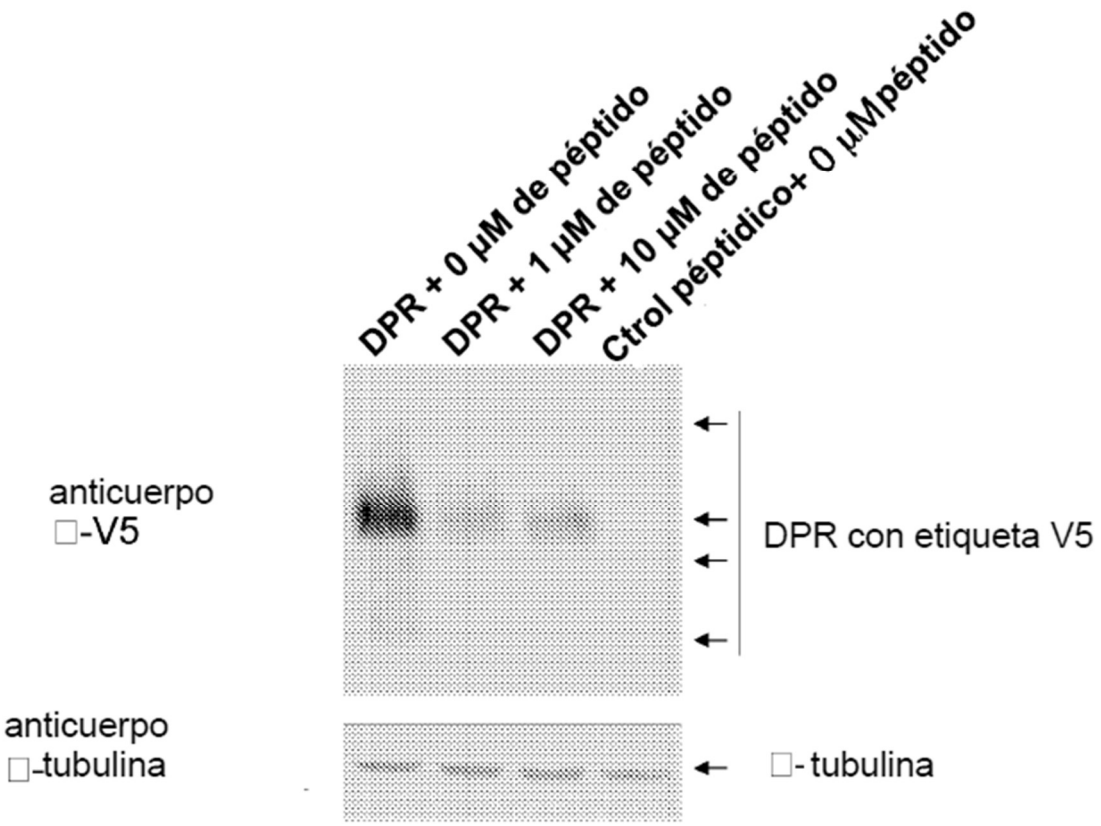


Figura 32


DPR + 0 µM de péptido
DPR + 1 µM de péptido
DPR + 10 µM de péptido
Ctrol péptidico+ 0 µMpéptido
anticuerpo
□-V5
DPR con etiqueta V5
anticuerpo
□-tubulina
□- tubulina