
(12) UK Patent Application (19) GB 2623164 (13) A
(19) (11) (43) Date of A Publication 10.04.2024

(21) Application No:	2312083.5			(51) INT CL:	
(22) Date of Filing:	07.08.2023			C12N 15/864 (2006.01)	A61K 31/713 (2006.01)
(30) Priority Data:	(31) 2211638 (32) 09.08.2022 (33) GB (31) 2211673 (32) 10.08.2022 (33) GB			A61K 38/16 (2006.01)	A61K 38/17 (2006.01)
				A61P 25/28 (2006.01)	C07K 14/47 (2006.01)
				C12N 15/113 (2010.01)	
(71) Applicant(s):	<p>University Of Sheffield (Incorporated in the United Kingdom) Western Bank, Firth Court, Sheffield, S10 2TN, United Kingdom</p>				
(72) Inventor(s):	<p>Guillaume Hautbergue Mimoun Azzouz Pamela Shaw</p>				
(74) Agent and/or Address for Service:	<p>Symbiosis IP Limited Innovation Centre, Innovation Way, Heslington, York, YO10 5DG, United Kingdom</p>				
(54) Title of the Invention: Viral vector					
	Abstract Title: Antagonists of Serine/Arginine Rich Splicing Factor 1 (SRSF1)				

(57) The present disclosure relates to antagonists that target or inhibit Serine/Arginine Rich Splicing Factor 1 (SRSF1). A viral vector, such as an AAV, is disclosed which comprises a non-expressed sequence (e.g. a stuffer sequence) and a nucleic acid encoding an antagonist of SRSF1 which is operably linked to a promotor. The antagonist may be an inhibitory RNA, such as shRNA, miRNA, antisense RNA or may be an inhibitory peptide which is, for example, fused to a cell penetrating peptide, CPP. The viral vector may be formulated as a pharmaceutical. Also disclosed are an inhibitory RNA to SRSF1, a cell penetrating peptide antagonist of SRSF1 and an agent comprising an inhibitory RNA to SRSF1, each of which may be formulated into a pharmaceutical composition. The SRSF1 antagonists may be used in the treatment of neurodegenerative diseases such as Amyotrophic Lateral Sclerosis (ALS, motor neurone disease, MND), sporadic Amyotrophic Lateral Sclerosis which is not caused by a pathological C9ORF72 hexanucleotide repeat expansion, sporadic frontotemporal dementia (FTD) or Fragile X-associated tremor/ataxia syndrome (FXTAS).

GB 2623164 A

Figure 1

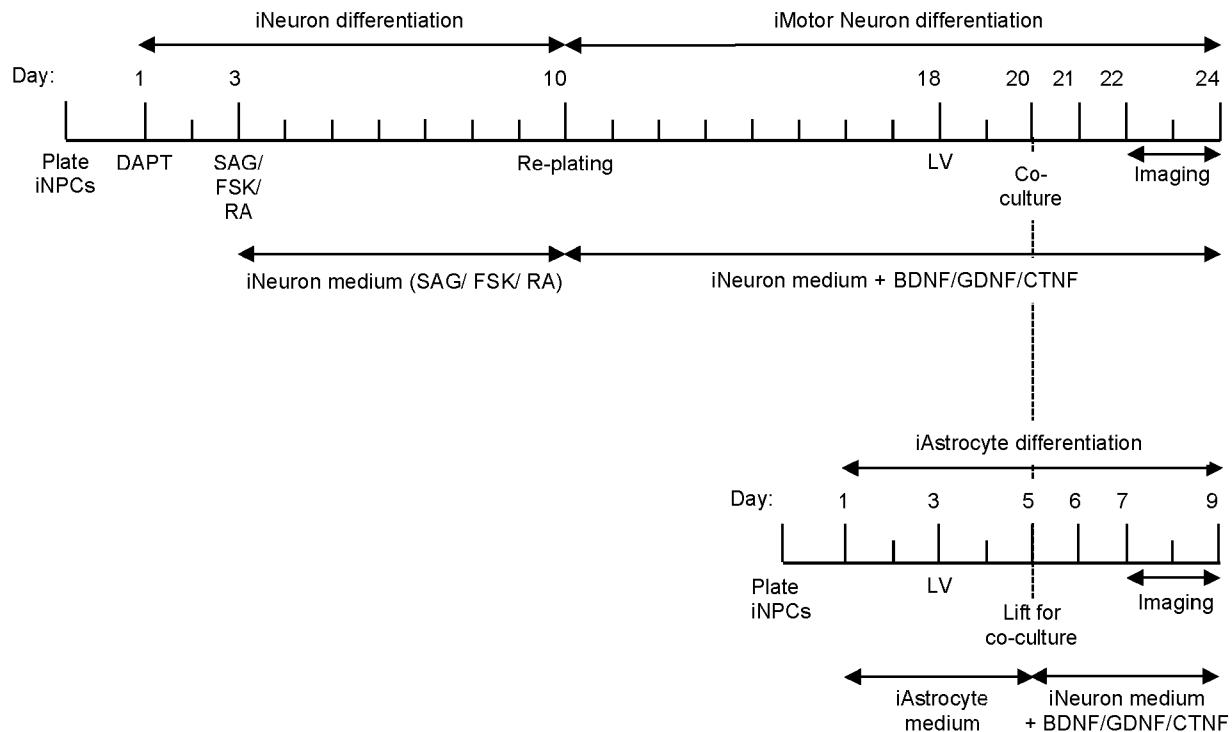


Figure 2

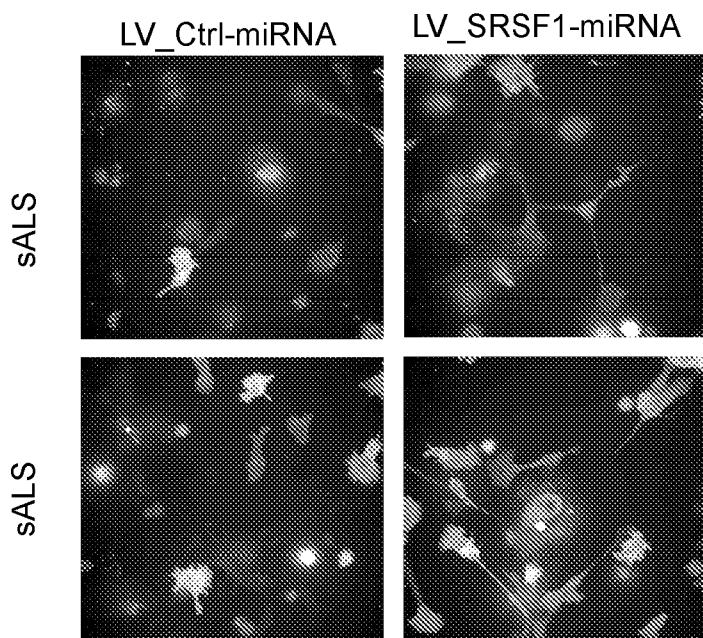


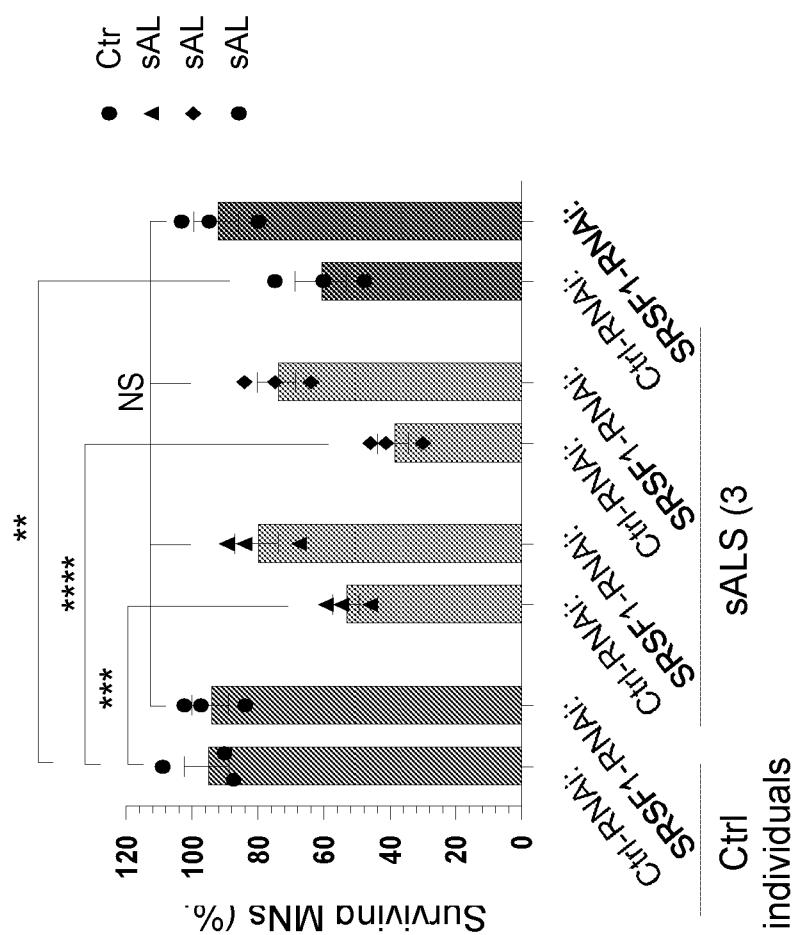
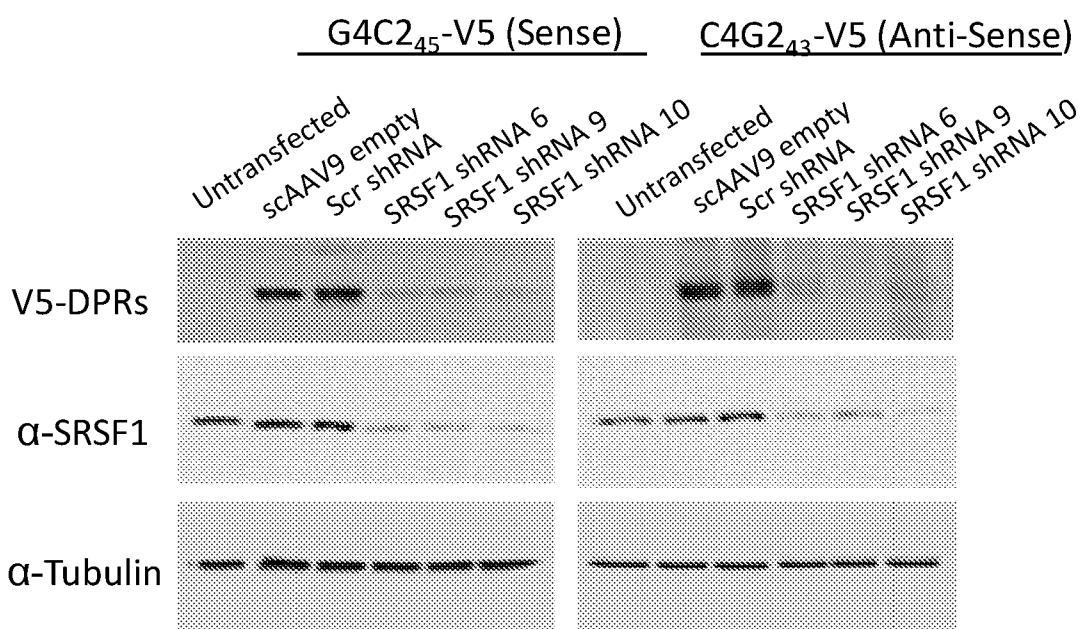
Figure 3**Figure 4**

Figure 5

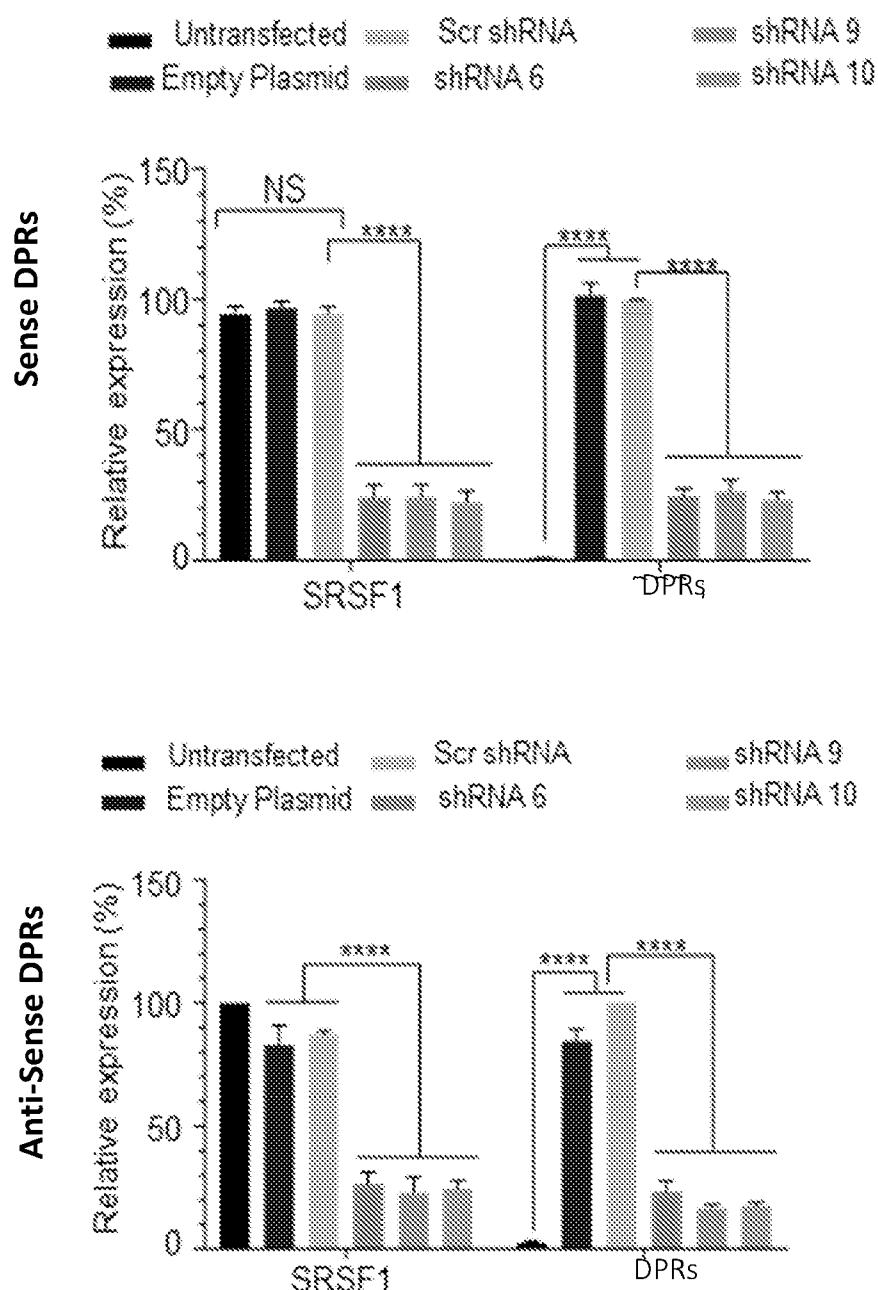


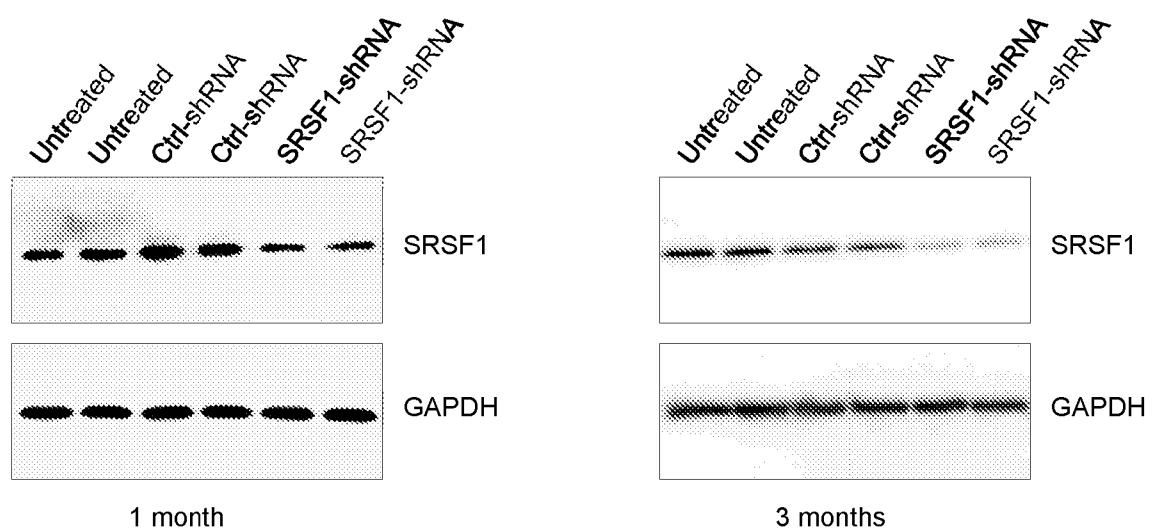
Figure 6

Figure 7

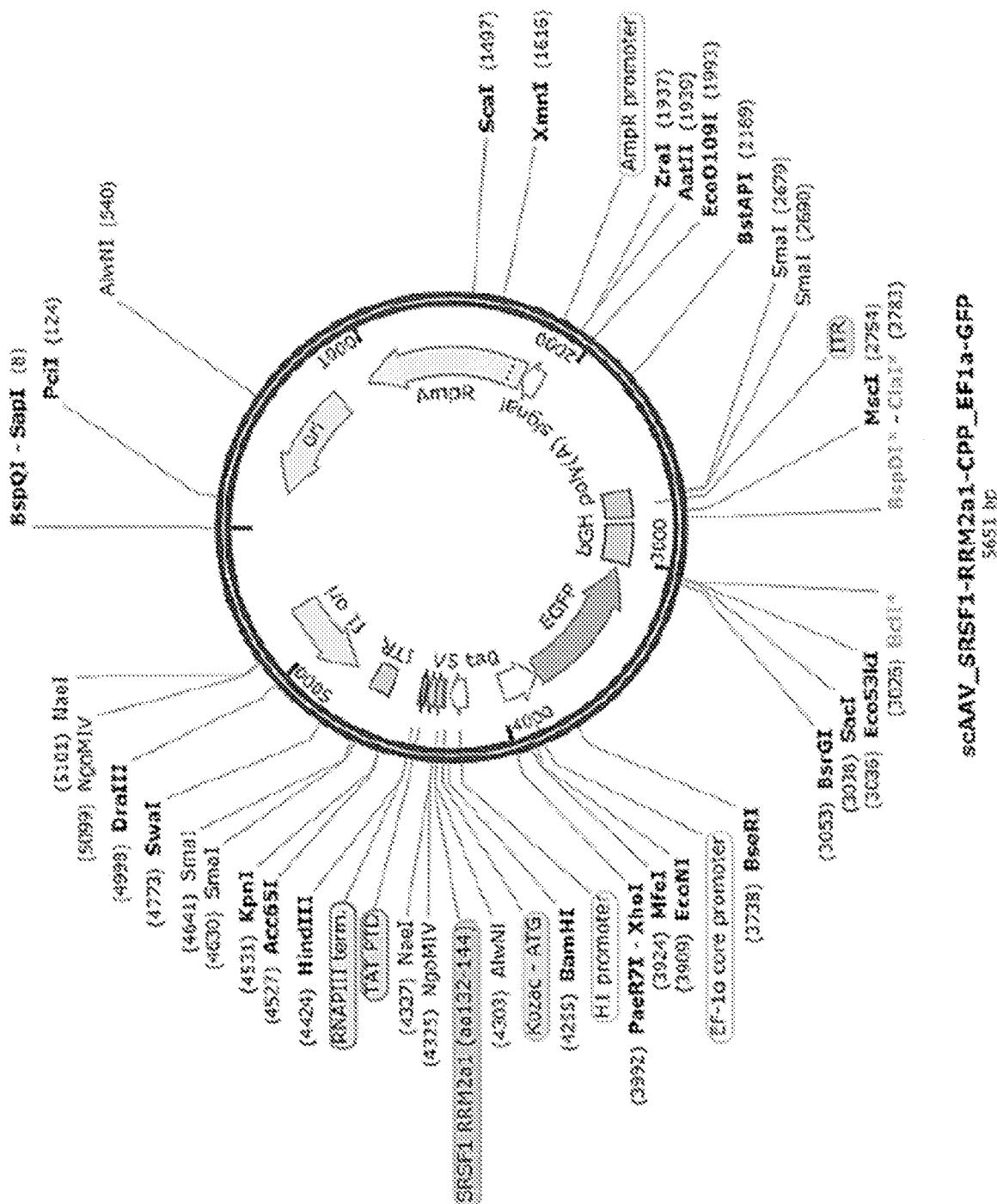


Figure 7 continued



Figure 7 continued

AluNI

68

89

109

209

309

409

509

609

709

809

909

925

936

285 300 325

W H K I L S A G T E D F R N

AAG

Figure 7 continued

ttcgttcatccatagttgcctgactccccgtcgtagataactacgataacggga
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 aagcaagttaggtatcaacggactgagggggcagcacatctattgtatgttatgcct
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 220 265 260 265 265
 R E D M T A Q S G T T Y I V V I R S
 AmpR

1045

gggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccg
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 cccgaatggtagaccggggtcacgacgttactatggcgctctgggtgcgagtggc
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 260 265 260 265 265
 P K G D F G I A A I I G R S G R S G
 AmpR

1100

gctccagatttatcagcaataaaccageccagccggaaaggcccggcggcagaagtg
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 cgagggtctaaatagtcgttatttggtcggtcggccttccggctcgctgtttcac
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 235 230 225 220 220
 A G S K D A I F W G A P L A S R T T P
 AmpR

1155

gtcctgcaactttatccgcctccatccagtcattatattgttgcggaaagctag
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 caggacgttgasataggcggaggtagtcagataattaacaacggcccttcgatc
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 215 210 205 200 200
 G A V K D A E M W D I L Q Q A S A I
 AmpR

1210

sgtaagttagttcgccagttaatagttgcgcacgttgcattgtacaggc
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 tcattcatcaaggcggtcaattatcaaaccgcgttgcacacaacggtaacgtccg
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 195 190 185 185 185
 T L L E G T I L K R L T T A M A V P
 AmpR

1265

atcgtgggtcacgctcgctgttttgtatggcttcattcagctccggttccaaac
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 tagcaccacagtgcgagcagcaaccataccgaagtaagtcgaggccaaagggttg
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 180 175 170 165 165
 M T T D R E D N P I A E N L E P E W R
 AmpR

1320

gatcaaggcgagttacatgatccccatgttgtgcaaaaagcggttagctcctt
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ctatgtttccgtcaatgtactaggggtacacacgttttcgcacatcgaggaa
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 160 155 150 145
 D L R T V H D S M N H L F A T L E K
 AmpR

1375

Figure 7 continued

Figure 7 shows five sequence alignments of DNA fragments with predicted protein domains. Each alignment includes the sequence, a schematic of the domain organization, and the position of the restriction enzyme cleavage site.

1. SmaI:

Sequence: cggtcctccgatcggtgtcagaagtaagttggccgcagtgttatcactcatggtt
 Domains: P G S I T T L L N A A T N D S M T
 Position: 1430

2. SmaI:

Sequence: atggcagcactgcataatttcttactgtcatgccatccgtaaagatgtttctg
 Domains: I A A S C L E R V T M G D T L H K F Y
 Position: 1485

3. ScaI:

Sequence: tgactggtagtactcaaccaagtcatctgagaatagtgtatgcggcgaccgag
 Domains: V P S Y E V L D N Q S Y H I R R G L
 Position: 1540

4. XbaI:

Sequence: ttgcctttccccgggtcaatacggataataccggccacatagcagaacttta
 Domains: Q E Q G A D I R S L V A G C T L V K
 Position: 1595

5. XbaI:

Sequence: aaagtgcctcatcattggaaaacgttctcgccggcgaaaactctcaaggatttac
 Domains: F T S M N P F R E E P R F S E L I K G
 Position: 1650

6. XbaI:

Sequence: cgctgttgagatccagttcgatgtacccactcggtgcacccaaactgatcttac
 Domains: S N L D L E I Y G V R A G L Q D E A
 Position: 1705

Figure 7 continued

atcttttactttcaccagcgtttctgggtgagcaaaaacaggasggcaaaatgcc
 tagaaaatgaaagtggtcgcaaagacccactcgtttgccttccgtttacgg
 1760

D K V K V L T E P H A F V R L C F A
 signal sequence

gcaaaaaaggaaataaggcgacacggaaatgtgaatactcatactttcc
 cgttttttcccttattccccgcgtgcctttacaacttatgagatgagaaggaaa
 1815

A F F P I K A V R F H Q I S M
 signal sequence AmpR promoter

ttaaatattattgaagcatttcatgggttattgtctcatgagcgatataatt
 aagttataataacttcgtaaatagtcctaataacagagtagactcgccatgtataa
 1870

AmpR promoter

tgaatgtatTTAGAAAAATAACAAATAGGGTTCCCGCGCACATTCCCCGAAAAA
 acttacataaaatTTTTATTGTTATCCCCAAGGCGCGTGTAAAGGGGTTTT
 1925

AmpR promoter

ZraI AatII
 gtgcacacctgacgtctaaagaaaccattattatcatgacattaaacctataaaaata
 cacgggtggactgcagatttttgtataatagtagactgttaattggatattttat
 1980

EcoO109I
 ggcgtatcacgaggcccTTTcgctcgcggtttcggtgtacggtaaaacct
 ccgcatagtgtccggaaagcagagcgccaaagccactactgccacttttgg
 2035

ctgacacatgcagctccggagacggcacagttgtctgttaagcgatggccgg
 gactgtgtacgtcgagggcctgtccagtgtaacagatcgccatcgccactacggccc
 2090

agcagacaagcccgtaaggcgccatcagcggtgttggcggtgtcgccggctgg
 tcgtctgttccggcagtcggcgccatcgccacacccggccacagcccccgg
 2145

BstAPI
 ttaactatgcggcatcagagcagattgtactgagagtgcacccatatgcgggtgtga
 aattgataacgcctgtactctcgtaacatgactctcacgtggatacggccacact
 2200

Figure 7 continued

aataccggcacagatgogtaaggagaaataccgcattcaggcgatccaaatcca
 2255
 ttatggcggtgtctacgcatttcctttatggcgtagccgcatacggttgcgggt

 atcaatccatcacagggaaaggaaaagaatrragcaaaatcaagcaatacgcctcaga
 2319
 tatttagtatgtccgttccgtttcttatacgttttatcgttataitcggagkt

 gcttaaaagctaaatccgttgttaccaaaaaacattatgacccctgttaacttttgcg
 2365
 cgtatttcgattttgcacaaatgglttttataactgggacattatgaaacgc

 ggagaagccctttatttcaacgcacggataaaatrragaaccccatataattt
 2420
 ectttcggaatataagttgcgttccattttttttttttttttttttttttttt

 taatgcataatgcctcaagtaatgttaggttttttttttttttttttttttttt
 2476
 atttacgttacggatcttttttttttttttttttttttttttttttttttttt

 ggagacagtccatcaccaatcaatatgtatattcaacccgtttagctgttttttt
 2530
 ctttttttttttttttttttttttttttttttttttttttttttttttttttttt

 atggccggagagatgtatatttttttttttttttttttttttttttttttttttt
 2585
 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt

 gacttgagagtcttggggcaaaacaaagagatcgcccccccccccccccccccc
 2640
 ccggacttccatcgataaaaacttttttttttttttttttttttttttttttt



 SmaI SmaI
 ctcccttcgtcgccgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc
 2695
 ggggggggggggggggggggggggggggggggggggggggggggggggggg



 tggggggacctttgttttttttttttttttttttttttttttttttttttttt
 2750
 agcccccgttggaaaccacagggggccggagccatccatccatccatccatcc



 SacII SacII
 stggcccaactccatcaactagggttttttttttttttttttttttttttttt
 2805
 caacgggttttttttttttttttttttttttttttttttttttttttttttt



Figure 7 continued

TTA~~C~~tccccagcatgcctgttatccattttccaaatccctcccccgttgtgttctg
 2860
 AATGGAGGGGTCGTACCGACCAATGAGAGAAGGGTLAGGAGGGGGGACCGACGGGGC
 SGN BOX(1) signal

ccccccaccccccaccccccagaaatagaatgacacctactcagacaatggatgcant
 2915
 GCGGTGGGGTGGTTATCTTACTGTGGATGGTCTGTACGCTACGTTAA
 SGN BOX(1) signal

TCCCATTTATTAGGAAAGGACAGTGGGAGTAGGCACCTTCAGGGTCAAGGAA
 2970
 AGGAGTAAATTAATCTTCTCTGTACCCCTACCGTGGAGGTCCCAGTTCTTC
 SGN BOX(1) signal

GCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 3035
 CGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 SGN BOX(1) signal

BamH ^I	Eco53kI	SacI	BglII										
↓	↓	↓	↓										
toatcagggaaatcttagaaatttacttgtacagtcgtccatggccggggggat													3090
acttagtcgtcgagatcccttaaatgaacatgtcgagtcgggttcacta													
235 239													
R K Y L E D M G I T V													
EGFP													
3135													
G A A T V F E L L N H H D R K E N P													
EGFP													
3180													
D K S L A E Q T S L Y H N D P L L V F													
EGFP													

Figure 7 continued

ggcggctcgcccgatgggggtgttcgtggtagtggttggcgagctgcacgcgtgcs
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 cccggcaggcggttacccccacaagacgaccatccaggccgtcgacgtgcgacgg
 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250
 G D G T P Y N Q Q Y H D A L Q V S G
 EGFP

 gtcctcgatgttgtggatcttgaagttaaccttgcgttttttgttgc
 +-----+-----+-----+-----+-----+-----+-----+-----+
 cggggactagaacaccccccgtggaaaccttcaagtggaaacttccgggg
 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240
 G E I N H R I K F N V K I G N R Q K
 EGFP

 tggccatgatataggacgttgtggctgttgtggactccatgtgtccccca
 +-----+-----+-----+-----+-----+-----+-----+-----+
 agccgggtactatactgcacccggacacatcaatggggtcggaaacacgggg
 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220
 S A M I Y V N R S S Y N E L K H G L
 EGFP

 ggatgttgcgttcttgcaggatgcgtggcccttcagtcgtatgcgggttaccag
 +-----+-----+-----+-----+-----+-----+-----+-----+
 ccataacggcaggaggaaacttcagtcggggatgcggatccggcaatgttc
 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205
 I W G D E K F D I L G K L S I R S V L
 EGFP

 gatgtcgccctcgaaatccacccgtggccggatgtgtgcgttgcgttgc
 +-----+-----+-----+-----+-----+-----+-----+-----+
 ccacayccggggatggatggatggatggatggatggatggatggatggatgg
 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185
 T D C E F K V E A R T K Y N G D B K
 EGFP

 aagaaatgtggcccttgcgtggatggatggatggatggatggatggatggatgg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 ttcttctaccacggggggatggatggatggatggatggatggatggatggatgg
 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170
 F F I T R E Q V Y G S S M A S X F F D
 EGFP

 aatgtgtgtcatgtgtgtgggttgcgtggatggatggatggatggatggatgg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 gatggatggatggatggatggatggatggatggatggatggatggatggatggatgg
 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150
 H Q K M H D P Y R S F C Q V G Y T L
 EGFP

Figure 7 continued

Figure 7 continued

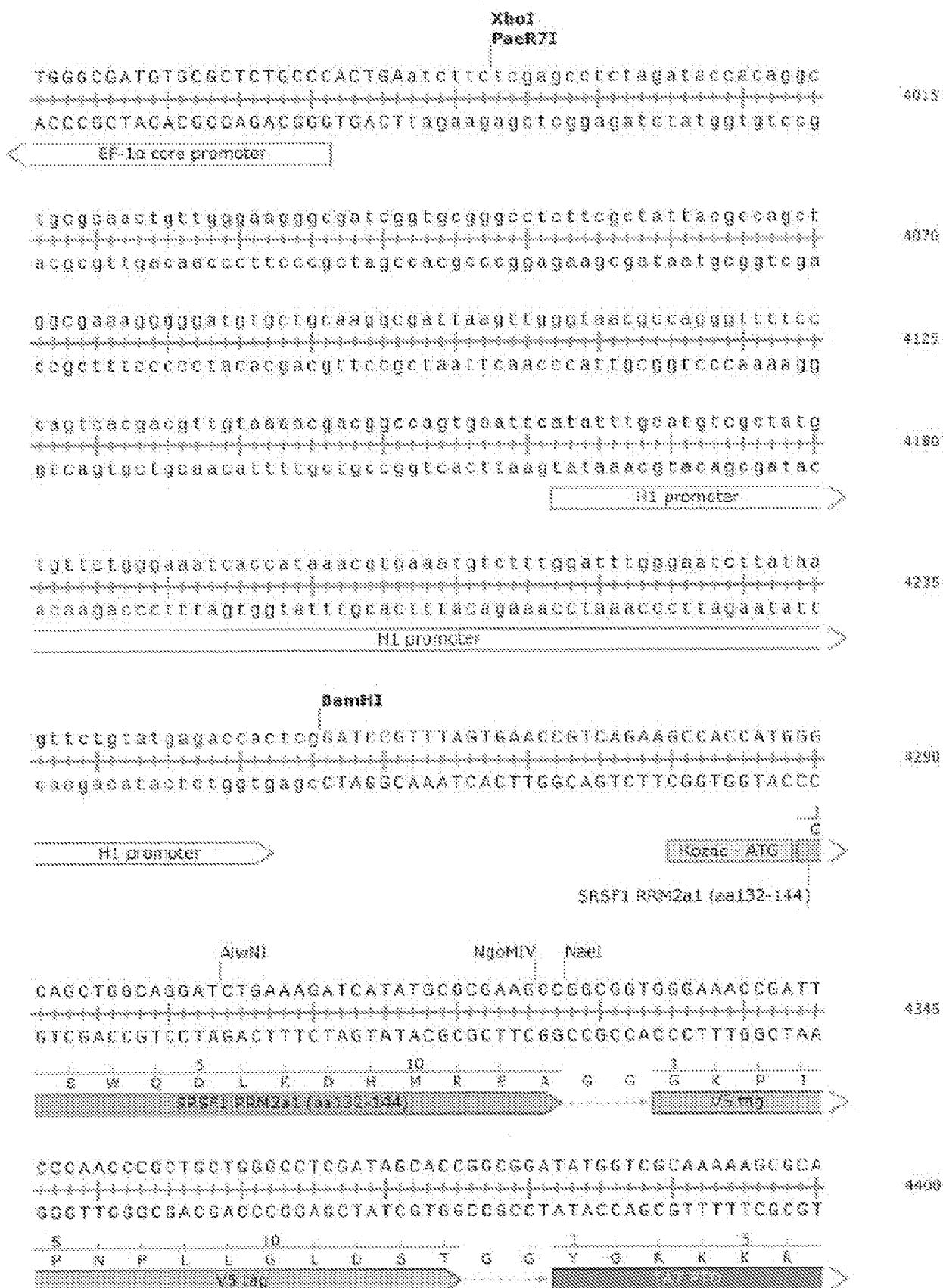


Figure 7 continued

HindIII

CACAGCGCCGGAGCTAATTTCCTTAAgctttggcgtaatccatggctatagctgttca
CTGTCCCGGCCCTCAATTAAAAATCggaaacggcatagtaccaggatcggacaaa
30
B O R R R ■■■
KNAPIII term.

cgttgtgaaattgttatccgcctcaaaatccccacacatccggggggggaaat
gacacacattaaacaataccggccatgtttaaagggtgtgttgtatgtccggcccttcata
4550

Acc65I KpnI

aaagtgtatcttagagccgttccacccgtgtatgtatccatccatccgggtgtat
tttccacatagatctccatgggtggccacttaatgtttaggtccggccatcc
5030
tccactcccttcgtgcgttcgtccgtactgaggccggggccatccaaaggctc
gggtgggggaaacggccggccggccggccggccgtggccgtggccgtggccggcc
5420

■■■

SmaI SmaI

ccggacggccgggtttggccggggccctcaatgtggccggccggccggccggcc
ggggccggccggccggccggccggccggccggccggccggccggccggccggcc
5875

■■■

gtgttttttttttttttttttttttttttttttttttttttttttttttttttttttt
cttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
6230

SwI

caaaaaaaaacggccatgttttttttttttttttttttttttttttttttttttt
tttttttttttttttttttttttttttttttttttttttttttttttttttttttt
6785

■■■

tttttttttttttttttttttttttttttttttttttttttttttttttttttttt
aatttttttttttttttttttttttttttttttttttttttttttttttttttt
7140

tttttttttttttttttttttttttttttttttttttttttttttttttttttt
aatttttttttttttttttttttttttttttttttttttttttttttttttt
7540

ccatggccggaaatccggccaaatccctataaaatccaaatccaaatccaaatcc
gggtttatccgggtttccgggttttttagggaaatatttagtttttttttttttt
7895

■■■

Figure 7 continued

gggttgcgttgttcagtttggaaacaaaggccactattaaagaaacggact
 cctcaacttcacaaaggcataaaccttttttcaggtgataatttctgeacctga
 f1 ori 4850

Draft

CCAACGTCAAAGGGCGAAAAACCGCTCTATCAAGGGCATGCCCACTACCGTGAACC
 5005
 66TTGCAGTTCCCGCTTTGGCAGATASTCCCCCTACCGGCTGATCGACTTGG

f1 ori

ATCACCTTAATCAAASTTTTGGCTCGAGGTGCCGTAAGCACTAAATCGAAC
 5060
 TASTGGGATTAGTTCAAAAAACCCGCTCCACGGCATTTCGTGATTAGCCTTG

f1 ori

NcoMIV Nael

CCTAAAGGGAGCCCCGATTTAGCTTGACGGGAAAGCCGGCGAACCGTGCSSA
 5115
 GGATTTCCCTCGGGGCTAAATCTCGAACCTGCCCTTTCGGCCGCTTGCACCGCT

f1 ori

GAAAGCAAGGGAAAGAACGAAAGGAGCGGGCGCTAGGGCGTGGCAAGTGTAGC
 5170
 GTTTCCTTCCCTTCTTCGCTTCCCGCCGATCCCCCGAACCGCTTCACATCG

f1 ori

GGTCAACGCTCGCGCTAACGACCCGCGCGCTTAATCGCCGCTACAGGCC
 5225
 CCAGTGCBAACCGCGATTCGTGCTGTCGGCCGCGAATTACGGCGCGATGTCCTG

f1 ori

CGCTACTATGGTTGCTTGACGAGCACGTTAAAGGTGCTTCCCTGGTTaaatca
 5280
 CGCATGATAACCAACGAAACTGCTGCTGATATTGACGAAAGGAGGAAAGGAA

f1 ori

GACCGGGAGCTAAACAGGAGGCCGATTAAGGGATTTAGACAGGAACCGCTACCG
 5335
 CTGCGCCCTCGATTTGCTTCCGCGCTAAATTGCGCTAAATCTGTCCTGCCATGCC

CABAATCCGTAGAAAGTGTAAAATAATCACTGAGGCCACCGAGTAAAGAGAGTCTG
 5390
 GTTTAGGACTCTTCACAAAAATAATTAGTCACTCCGCTGGCTCAATTCTGAGAC

TCCATCACGCAARTIAACCGTTGTCGCAATACTTCTTGTATTAGTAATAACATCA
 5445
 AGCTAGTGGCTTAAATGCGCAACAGGCTATGAAAGAAACTAAATCATATTGTAGT

Figure 7 continued

OTTGCCCTGAGTAGAAGAACTCAAACATATCGGCCTTGCTGGTAAATATCCAGAACAA
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 5580
 GAACGGACTCATCTTCTT6AGTTGATA6CCGGAACGACCAATTATAGGTCTTGT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TATTACCGCCAGCCATTGCAACGGAATCGCCATTGCCATTCAAGGCTGCGCAACT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 5585
 ATAATGGCGGTGGTAACGTTGCCCTTACGCGTAAGCGGTAAAGTCCGACGCGTTGA
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTTGGGAAGGGCGATCGGTGCCGGCCTTCGCTATTACGCCAGCTGCATTAATG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 5610
 CAACCCCTTCCCCTAGCCACGCCCCGGAGAACGCGATAATGCGGTGACGTAATTAC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 AATCGGCCAACGCGCGGGGAGAGGGCGGTTTGCCTATTGGGC 3'
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 5651
 TTAGCCGGTTGC6CGCCCCCTCTCCGCCAAACGCGATAACCCG 5'
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

Figure 8

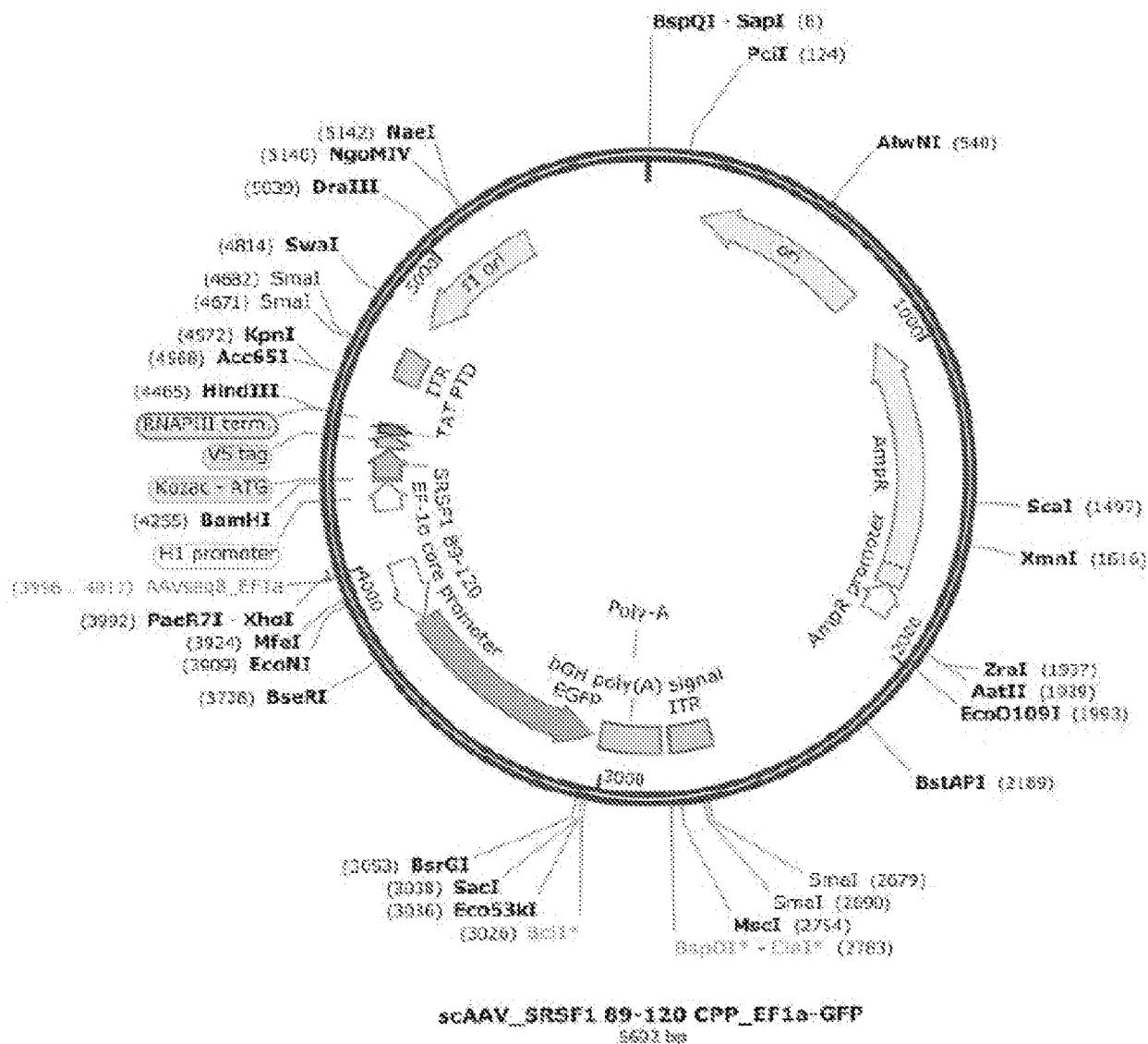


Figure 8 continued

SapI
BspQI

5' GCTCTTCCGCTTCGCTCACTGACTCGCTGCCTCGGTGTTGGCTGGCGCGCG
3' CGAGAAGGCCAAGGAGCGAGTGAATGACGCGACGCCAGCAASCCSACGCCG
AGC6GTATCAGCTCACTCAAASGCCGTAATAACGGTTATCCACAGAAATCAGGGAT
TGCGCATAGTCGAGTGGAGTTCCGCCATTATGCCATAGGTGTCTTAGTCCCCCTA
150

PclI

AACGCAGGAAGAACATG TGAGCARAAAGGCCAGCAAAASGCCAGGAACCGTAAGAA
TYCCGTCTTCTTCTTACACTCGTTTCCGGTCGTTTCCGGTCTTSGCATT
195

AGGCCGCCTTggcgggtttttccataggctccggcccccctgacggggatcata
TCCGGCSCAaggggggggggggggggggggggggggggggggggggggggggg
220

5' ← 150 ← 195 ← 220 ←

aaatcgacggccaaatcaggagggtggcgggggggggggggggggggggggggg
tttagtgcgggggttcagtctccacccgctttggggctgttgtatatttttttt
275

ccgtttccccccctggggaggctccctgtggctccctgttccggcccccggccctt
ccggcaaggggggggggcccttgggggggggggggggggggggggggggggggg
330

ccggataccgtgtccggcccttccctttccgggggggggggggggggggggggg
ggccatgggggggggggggggggggggggggggggggggggggggggggggggg
385

acgtgtgggtatctcaggccgggtgtgggggggggggggggggggggggggg
tcggggatggatggatggatggatggatggatggatggatggatggatggatgg
440

cacggaaaaaaaacgggtttgggggggggggggggggggggggggggggggg
gtgtttgggggggggggggggggggggggggggggggggggggggggggggg
495

5' ← 150 ← 195 ← 220 ←

Figure 8 continued

AmpR

550

605

660

715

770

825

880

935

AluHII

550 605 660 715 770 825 880 935

Tyr His Lys Ile Ile Ser Asp Gly Ile Glu Asp Ile Glu Asp

Figure 8 continued

ttcgttcatccatagttgcctgactccccgtcgtagataactacgataacggga
 1045
 aagcaagtaggttatcaacggactgaggggcagcacatctattgatgctatgccct
 279 286 293 299 306 313 320 327 334 341 348 355 362 369 376 383
 Arg Glu Asp Met Thr Ala Gln Ser Gly Thr Thr Tyr Ile Val Val Ile Arg Ser
 AmpR

gggcttaccatctggccccagtgcgtcaatgataccggcgagacccacgctcaccg
 1106
 cccgaatggtagaccggggtcacgacgttactatggcgctctgggtgcgactggc
 350 357 364 371 378 385 392 399 406 413 420 427 434 441 448 455
 Pro Lys Gly Asp Pro Gly Leu Ala Ala Ile Gly Arg Ser Gly Arg Gly Gly
 AmpR

gctccagatttatacgcaataaccaggccagccggaaaggcccggcagaatgt
 1155
 cgaggcttaaatagtcgttatttgtcggtcgccctccggctcgcgctttcac
 235 242 249 256 263 270 277 284 291 298 305 312 319 326 333 340
 Ala Gly Ser Lys Asp Ala Ile Phe Trp Gly Asp Pro Leu Ala Ser Arg Leu Ile Pro
 AmpR

gtcctgcaactttatccgcctccatccaggcttatttgcgttgcggaaagctag
 1210
 caggacgttgaaataggccggaggtaggtcagataattaacsacggcccttcgatc
 215 222 229 236 243 250 257 264 271 278 285 292 299 306 313 320
 Gly Ala Val Lys Asp Ala Glu Met Trp Asp Ile Leu Gln Gln Arg Ser Ala Leu
 AmpR

agtaagtagttcgccaggtaatagttgcgcacgttgcattgtacaggc
 1265
 tcattcatcaagcggtcaattatcaaacgcgttgcacaacaacgttaacgtatgtccg
 195 198 201 204 207 210 213 216 219 222 225 228 231 234 237 240
 Thr Leu Leu Glu Gly Thr Leu Leu Lys Arg Leu Thr Thr Ala Met Ala Val Pro
 AmpR

atcggtgtgtcacgctcgcttggatggcttattcagctccggttccaaac
 1320
 tagcaccacagtgcgagcagcaaccataccgaagtaagtcgaggccaagggttg
 180 187 194 198 201 208 215 222 229 236 243 250 257 264 271 278
 Met Thr Thr Asp Arg Glu Asp Asn Pro Ile Ala Glu Asn Leu Glu Pro Glu Trp Arg
 AmpR

Figure 8 continued

gatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagtcctt
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ctagttccgctcaatgtactaggggtacaacacgttttgccaaattcgaggaa
 . 160 155 150 145
 Asp Leu Arg Thr Val His Asp Gly Met Asn His Leu Phe Ala Thr Leu Glu Lys
 <-----
 AmpR

cggtcctccgategttgtcagaagaatgtggccgcagtgttatcactcatggtt
 +-----+-----+-----+-----+-----+-----+-----+-----+
 gccaggaggctagcaacagtcattcatticaaccggcgtcacaatagtgagttaccaa
 . 140 135 130
 Pro Gly Gly Ile Thr Thr Leu Leu Leu Asn Ala Ala Thr Asn Asp Ser Met Thr
 <-----
 AmpR

atggcagcactgcataattcttactgtcatgccatccgtaaagatgttttctg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 taccgtcggtgacgtttaagagaatgacagatcaggtaggcatttacgaaaagac
 . 125 120 115 110
 Ile Ala Ala Ser Cys Leu Glu Arg Val Thr Met Gly Asp Thr Leu His Lys Glu Thr
 <-----
 AmpR

Scal

tgactggtagtactcaaccaaggcatctgagaatagtgtatgcggcgaccgag
 +-----+-----+-----+-----+-----+-----+-----+-----+
 actgaccactcatgagttggttcagtaagactcttacatcacatcggccgtggct
 . 105 100 95 90
 Val Pro Ser Tyr Glu Val Leu Asp Asn Gln Ser Tyr His Ile Arg Arg Gly Leu
 <-----
 AmpR

ttgctttgccccgggtcaatacggataataccggccacatagcagaacttta
 +-----+-----+-----+-----+-----+-----+-----+-----+
 aacgagaacggggccgcagttatgccttattatggcgccgtgtatcgttttgaast
 . 85 80 75
 Gln Glu Gln Gly Ala Asp Ile Arg Ser Leu Val Asp Gly Cys Ile Leu Val Lys
 <-----
 AmpR

XmnI

aaagtgcacatcattggaaascgttcttggggccaaaaactctcaaggatttac
 +-----+-----+-----+-----+-----+-----+-----+-----+
 tttcacgagtagtaaaccttttgcagaagaagccccgcctttgagagttccataatg
 . 70 65 60 55
 Phe Thr Ser Met Met Pro Phe Arg Glu Glu Pro Arg Phe Ser Gln Leu Ile Lys Gly
 <-----
 AmpR

Figure 8 continued

cgctgtggagatccaggatcgatgtaaaccactcgatgcacccaaactgatcttcagc
 1705
 gcgacaactctaggtaagctacattgggtgagcacgtgggtgactagaagtcg
 50 45 40 35
 Ser Asn Leu Asp Leu Glu Ile Tyr Gly Val Arg Ala Gly Leu Gln Asp Glu Ala
 AmpR

atcttttactttcaccaggcggttctgggtgagcasaaacaggaaggcaaaatgcc
 1760
 tagaaaaatgaaagtggcgccaaagaccactcgttttgtccttcggtttacgg
 30 25 20
 Asp Lys Val Lys Val Leu Thr Glu Pro His Ala Phe Val Pro Leu Cys Phe Ala
 signal sequence
 AmpR

gcaaaaaaaaaggaaataagggcgacacggaaatgttaatactcataactcttcattt
 1815
 cgtttttcccttattcccgctgtgcctttacaacttatgagtatgagaaggaaa
 15 10 5 1
 Ala Phe Pro Ile Leu Ala Val Arg Phe His Gln Ile Ser Met
 signal sequence AmpR promoter

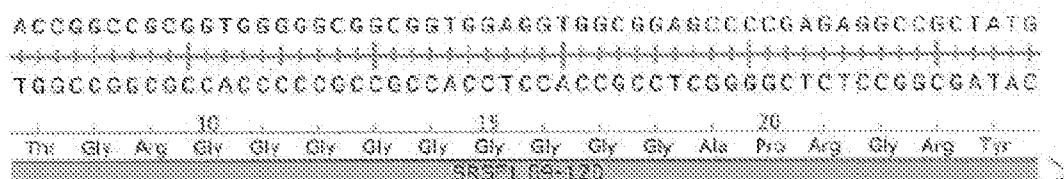
ttcatatattattgaangcatttatcagggttattgtctcatgagcggtatacatatt
 1870
 aaggatataataacttcgttaatagtcccaataacagagtactcgcttatgtataa
 AmpR promoter

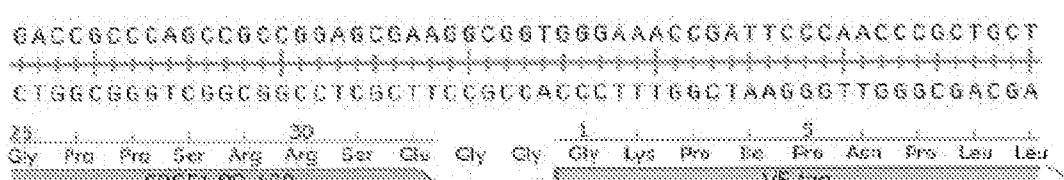
tgaatgtattttagaaaaataaaacaaaatagggttccgcacattccccgaaaaa
 1925
 acttacataaaatcttttattttgtttatccccaaaggcggtgtaaaggcccattt
 AmpR promoter

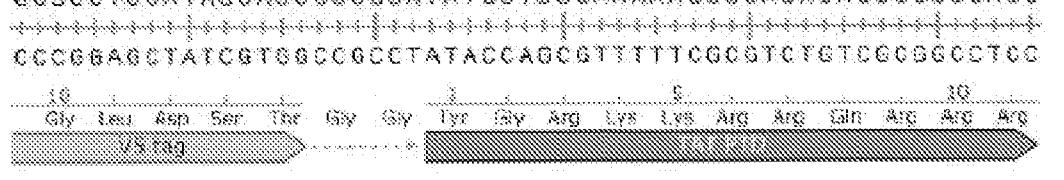
ZraI AatII
 gtgccacctgacgtctaagaaaccattattatcatgacattaaacctataaaaata
 1980
 cacgggtggactgcagatcttttgttaataatgtactgttaattggatattttat

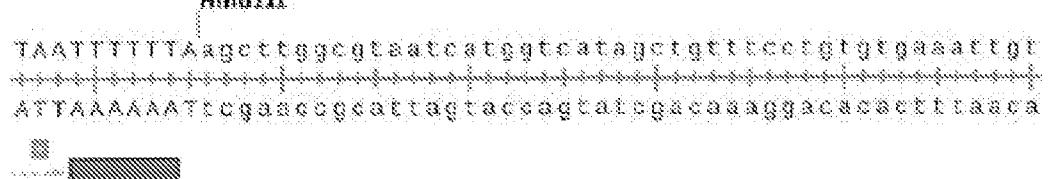
EcoO109I
 ggcgtatcacgaggcctttgcgtctcgccgtttcggtatgcgggtgaaaacct
 2035
 ccgcatagtgtccggaaagcagagcgcgcataactgcccacttttgg
 ctgacacatgcagctccggagacggtcacagcttgcgtgttaagcggtgcgggg
 2090
 gactgtgtacgtcgagggccctgtccagtgtaacagacattcgccctacggccc
 agcagacaaggccgtcagggcgccgtcagcggtgttggccgggtgtcgccccgtggc
 2145
 tcgtctgttcggccagtcggccgcagtcgcggccacaacgcggccacagccccgaccgg

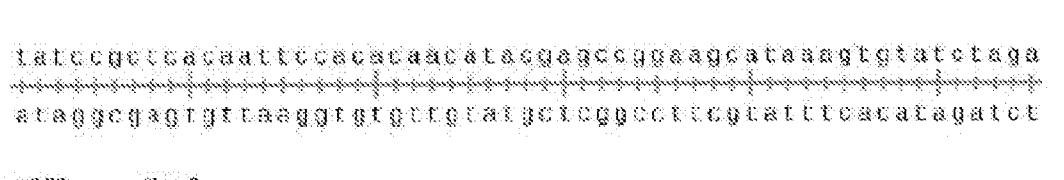
Figure 8 continued

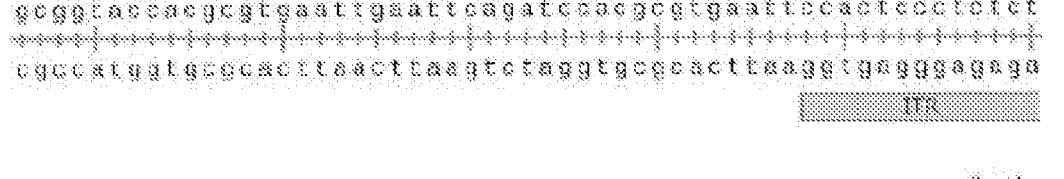
ACCGGCCCSCGGCTGGGGCGGCAGGTGGC GGAGCCCCGAGAAGGCCATATG
 TGGCCCGGCCACCCcccGCCCCACCTCCACCGCCCTCGGGCTCTCCGGCGATAAC
 10 21 26
 Thr Gly Arg Gly Gly Gly Gly Gly Gly Ala Pro Arg Gly Arg Tyr


GACCGCCCCGAGCCGCCGGAGCCAAGGGCGSTGGGAAGACCGATTCCCAACCCCGCTGCT
 CTGGCGGGTCGGCGGCCCTCGCTTCCGCTAACCCCTTTGGCTAAGGGTTGGGCGACGA
 28 33 3 5
 Gly Pro Pro Ser Arg Arg Ser Glu Gly Gly Lys Pro Ile Pro Asn Phe Leu Leu


GGCCCTCGATAGCAACGGCGATATGGTGGCAAAAGCGCAGACAGACAGCCGGAGG
 CCCGGAGCTATCGTGCGGCCCTATACCAAGCGTTTCCGCTCTCGCGGCCCTCC
 10 1 3 5 10
 Gly Ile Asp Ser Thr Gly Gly Tyr Asp Arg Lys Lys Arg Arg Glu Arg Arg Arg


HindIII
 TAATTTTTTAgtcttggcgtaatcatgtgtcatagctgtttccctgtgtgaaattgt
 ATTAAAAAAATTcggaaacggatttagtacccgtatcggasaaaaggacacattttaca


tatcccgctcacbattccacacaaactacggggccggaaaggataaagtgtatctaga
 *tagggcgagtgttaagggtgttgtatgtccggccatccgtatttcaatatacatct


Acc651 XbaI
 cgggtacccacggcggtggatttgaaattcagatccggccgggtggaaattccactcccttcet
 ccgcxtygtccccacttaacttaaggcttaggtggccacttggagggtgggggggg


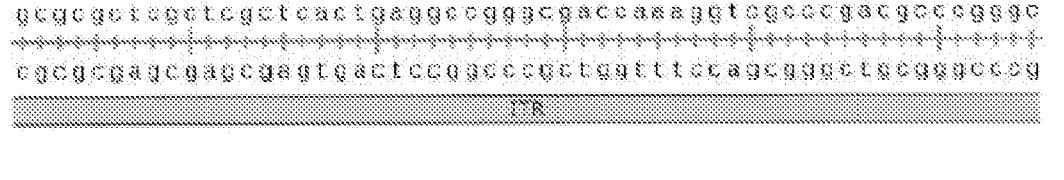
SmaI
 ggcgcgttcgttcgttcacttgaggccggccggccggccggccggccggccggccggccgg
 cggccggccggccggccggccggccggccggccggccggccggccggccggccggccggccgg


Figure 8 continued

Figure 8 continued

AAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCGCTGC
 TTTCGCCTTCTCCGGGGCGATCCCGCGACCGTTACADATCGCCAGTCGGACGCC
 ← f1.on | S225

TAACCACCCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGTACTATGGTGC
 ATTCGGTGGTGTGGCGGGCGGAATTACGCGCGATGTCGGCGCATGATAACCAAC
 ← f1.on | S280

CTTTGACGAGCACGTATAACGTGCTTCTCGTT>>CAGAGCGGGAGCTAAA
 6AAACTGCTCGTGCATATTGCAACGAAAGCAGCAATcttagTCTCGCCCTCGATTT
 ← S335

CAGGAGGCCSATTAAGGGATTTAGACAGGAACGGTACGCCAGAAATCCTGAGAA
 GTCCTCCGGCTAATTFCCTAAAACTGTCCTGCCATGCGTCTTAGGACTCTT
 ← S390

GTGTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCATCACGCAAAT
 CACAAAAATATTACTCAGTCGGGTGGCTCATTTCTCAGACAGGTAGTGCCTTTA
 ← S445

TAACCCTGTCGCAATACTTCTTGATTAGTAATACATCACTTGCCTGASTAGA
 ATTGGCAACAGCGTTATGAAAGAAACTAATCATTATTGTAAGTGAACGGACTCATCT
 ← S500

AGAAACTCAAACATCGGCCATTGCTGGTAATATCCAGAACAAATTACCCCCAGCC
 TCTTGAGTTGATGCCGGAACGACCATTATAGGTCTTGTATAATGCGCGTCCG
 ← S555

ATTGCAACGGAATCGCCATTGCGCATTCAAGGCTGCGCAACTGTTGGAAAGGGCGA
 TAACGTTGCCATTAGCGCTAAGCGGTAAGTCCGACCGCTTGACAACCCCTCCCGCT
 ← S610

TCGGTGCGGGCGCTCTTCGCTATTACGCCAGCGTGCATTAAATGAAATGCCAACCGCG
 AGCCACGCCGGAGAAGCGATAATGCGGTGACGTAATTACCTAGCCGGTTGCGC
 ← S665

CGGGGAGAGGGCGGTTGCGTATTGGGC 3'
 ... 5' S682
 GCCCCCTCTCCGCCAAACGCGATAACCGC 5'

Figure 9

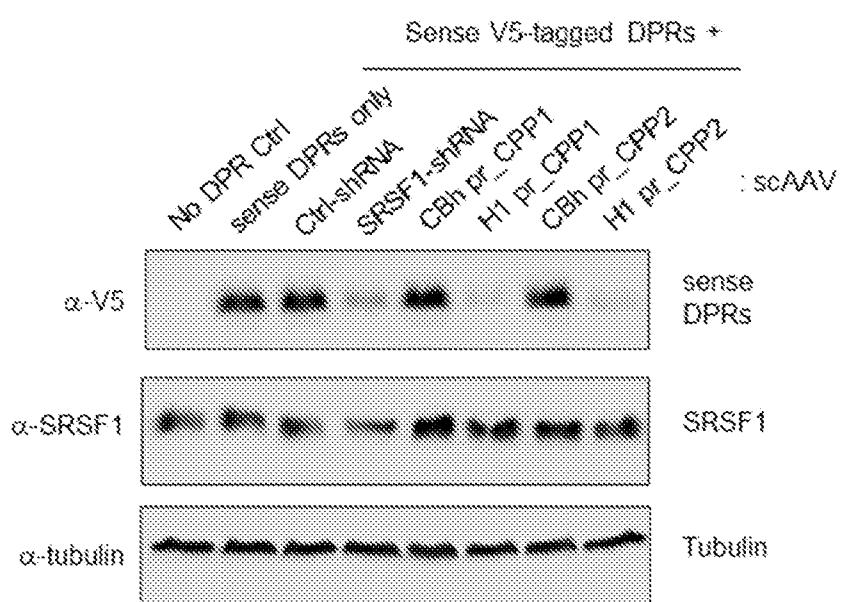
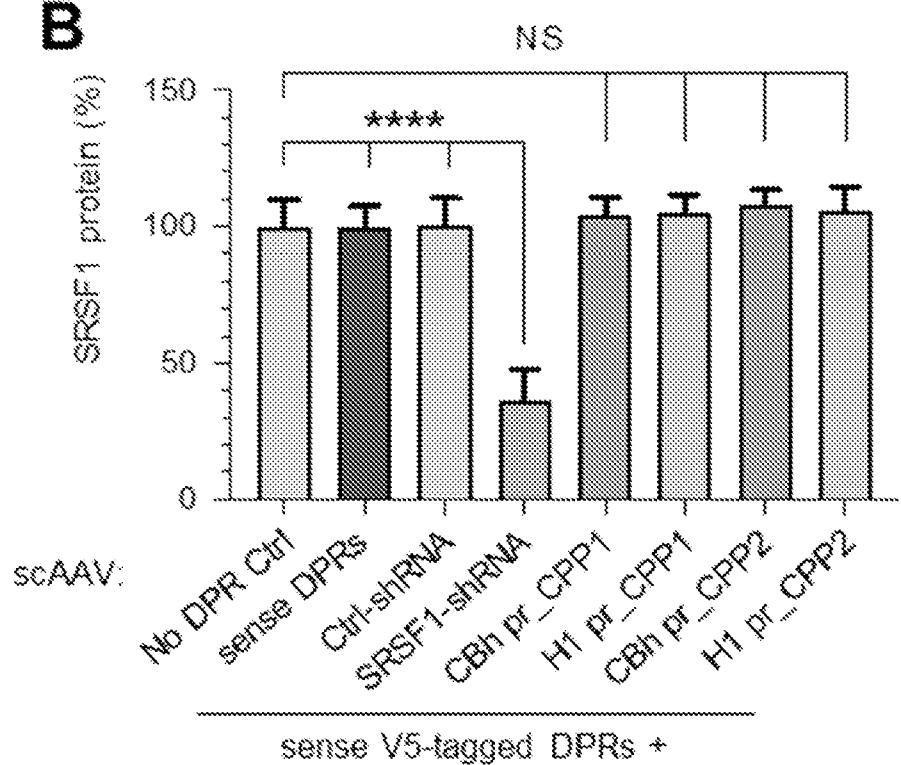
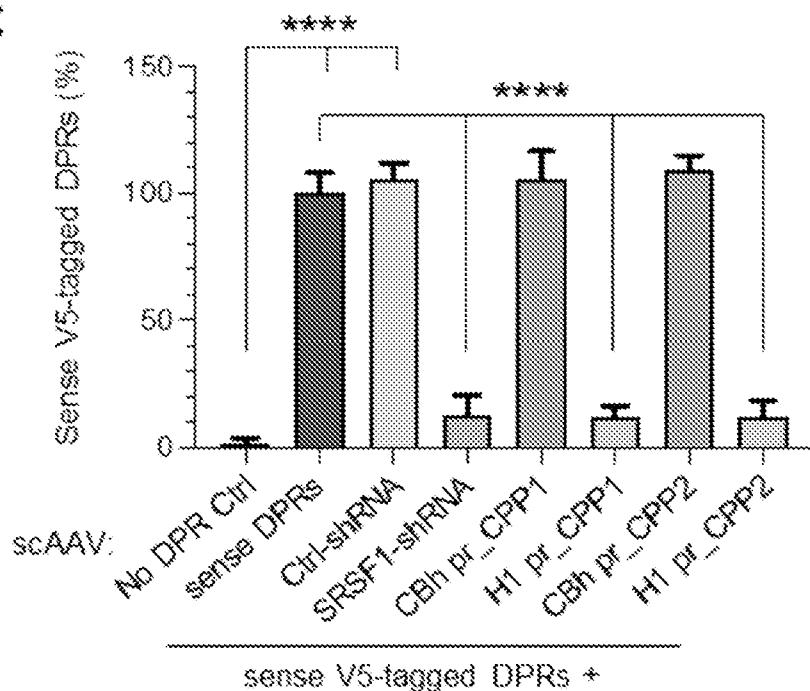
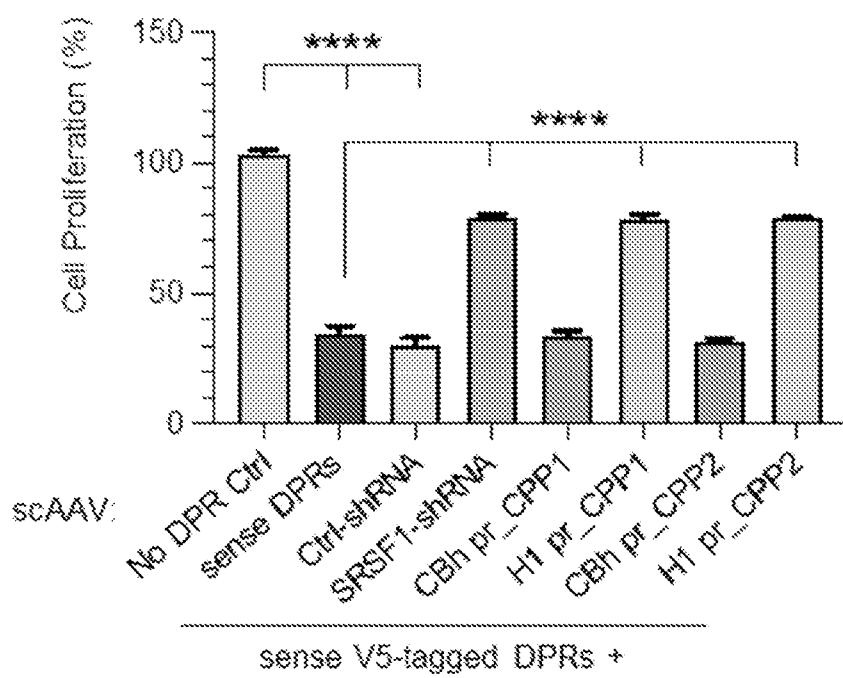
A**B**

Figure 9 continued

C

sense V5-tagged DPRs +

D

sense V5-tagged DPRs +

Figure 10

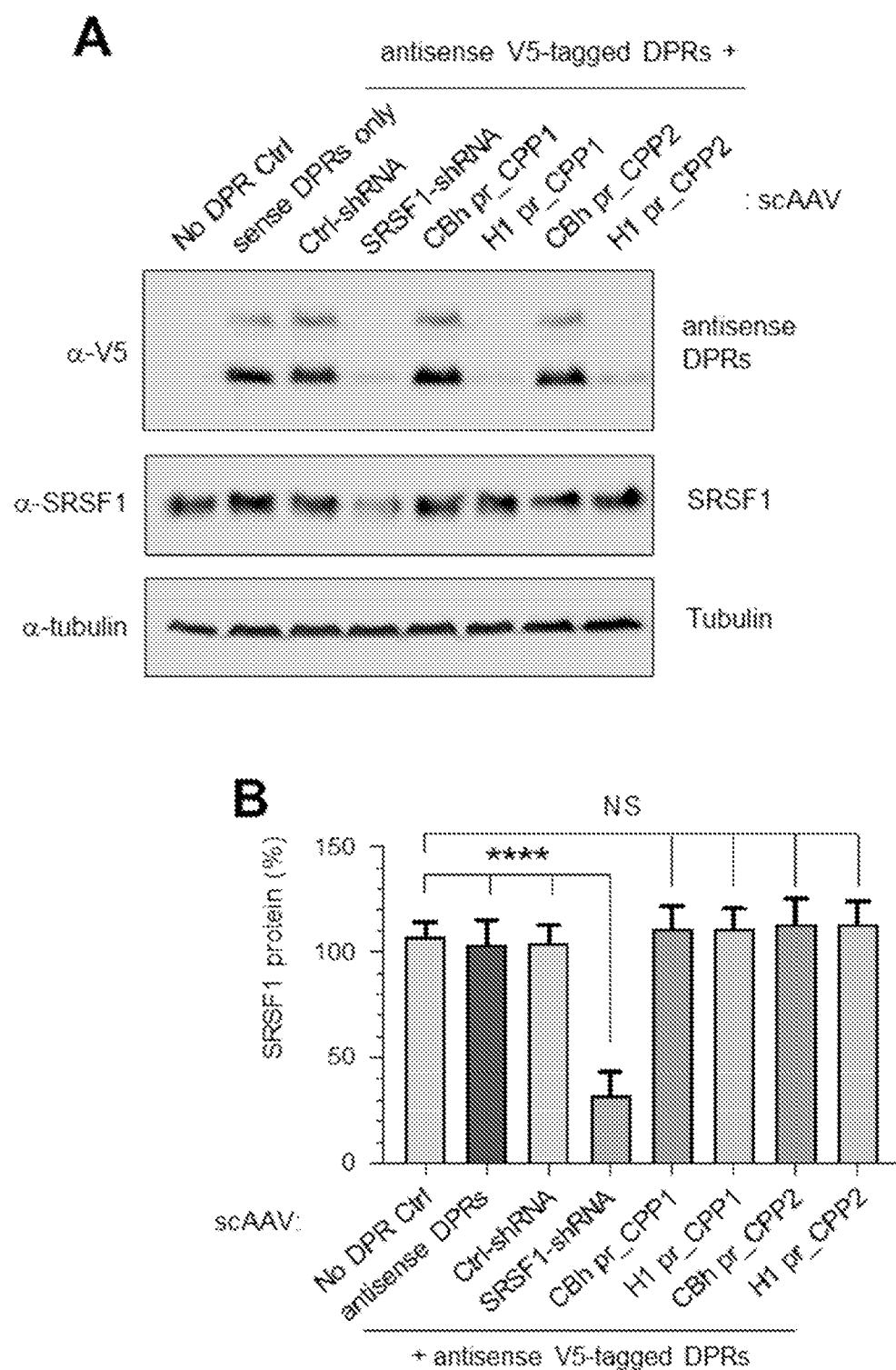


Figure 10 continued

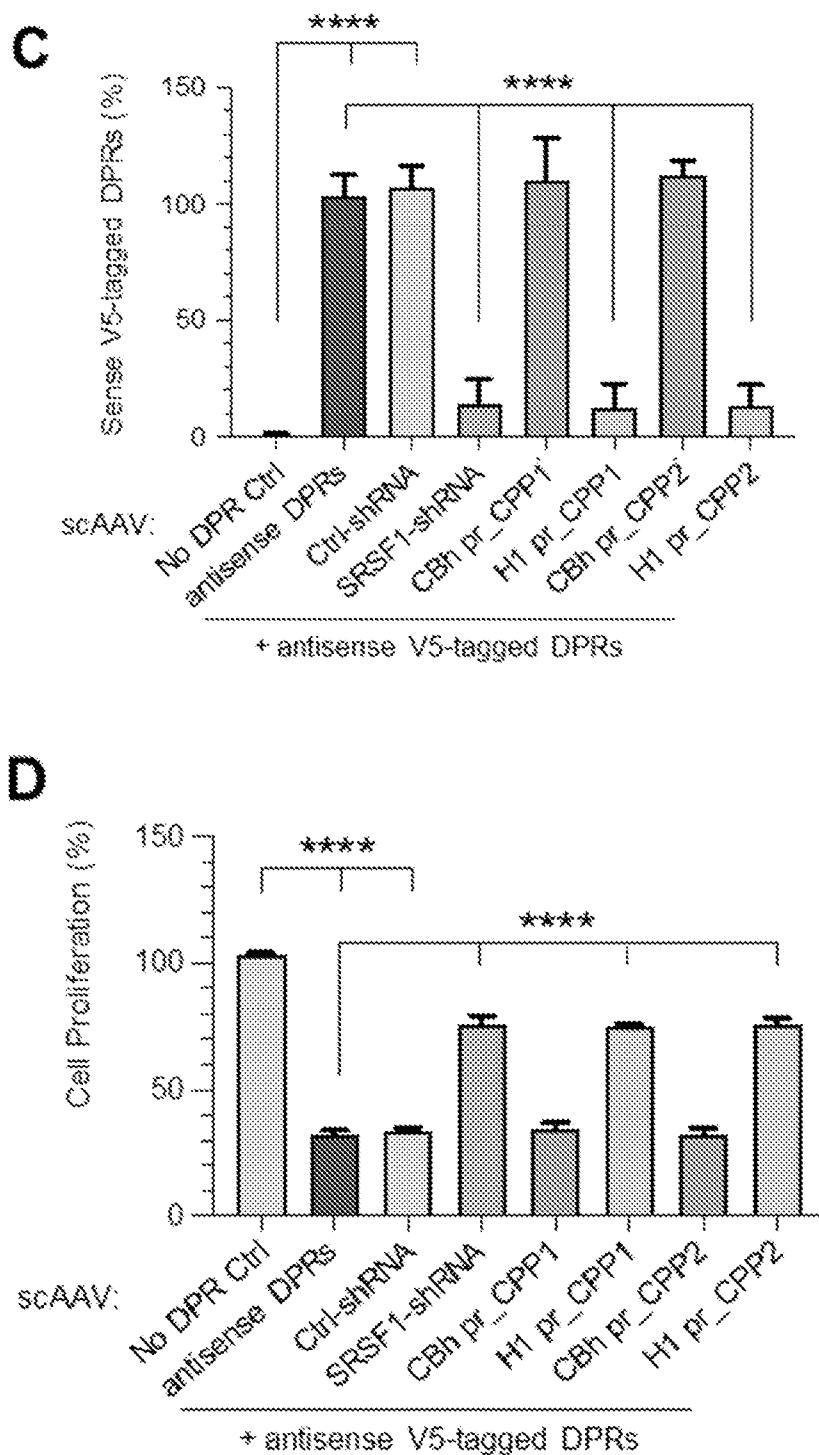


Figure 11

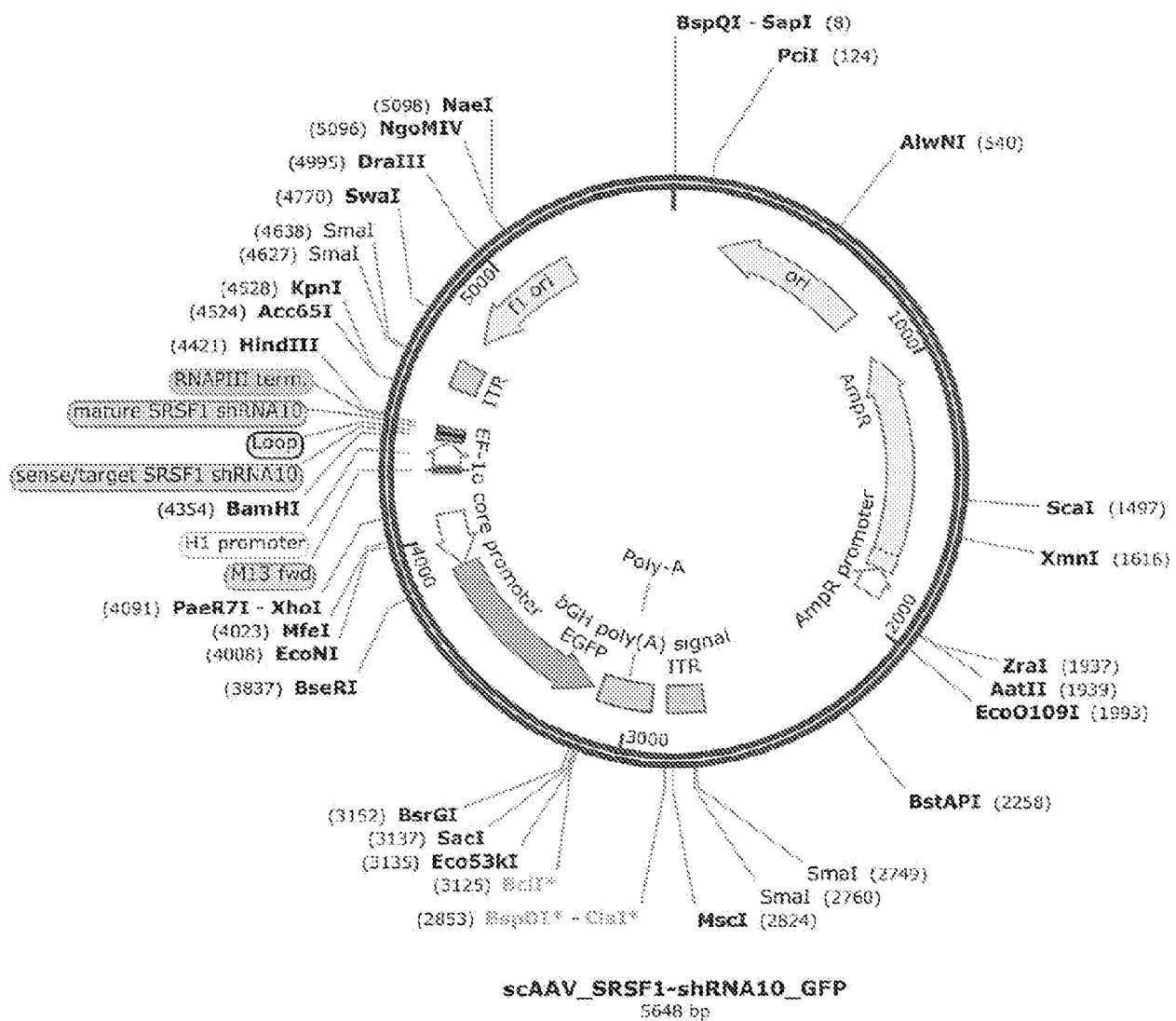


Figure 11 continued

Sapi
BspQI

5' ... CCTCTTCCGGTTCTC6CTCACTGACTCGCTGCCTCGGTGTTGGCTGC6CG
3' CGAGAAGGCCAAGGAGGGACTGACTGAGGCCAGCCAGGAAGCCGACGCCGC

55

AGCGGTATCAGCTGACTCAAAGGCCGTAATACGGTTATCCACAGAAATCAGGGAT
TCGCCATAGTCGAGTGAGTTCCGCCATTATGCCAATAGGTCTTAGTCCCCTA

110

PstI

AACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
TTGCGTCCTTCTTGTACACTCGTTTCCGCTCGTTTCCGGTCCCTGGCATTT

165

AGGCCGCCTgtggcgttttccataggctccgcggccctgacgagcatcaca
TCCGGCGCAargaaactgcaaaaaagtatccggaggcggggggactgtctgttgtt

220

ori

aaatcgacgttcaggctaggctagggtggcgaaaccgcacaggactataaagataccag
tttagctggaggttttagtctccaccgccttgggtgtctgtatattatcggttc

275

ori

gggtttcccccttggaaagtcctcggtggcggtctctgttccgaccctgcgcctt
ccgcacaaaggcgaccttccggggggcaccgcggagaggatgggacggcaat

330

ori

ccggataccctgtccggcctttcctccatccggaaagcggtggcgctttatcgctc
ggccatggacaggcgggasagagggasgcgccttgcgcaccgcggaaagatcgag

385

ori

acgttgtggatcttcgttccgggtgtgggtcggttcgttcggctgtgt
tgcgcacatccatagatgtcaagccacatccggcaagcgagggttcgcacccgacac

440

ori

cacggaccccccgttccggccggacccgtggcccttataccggtaactarcgcttg
gtgtttggggggcaagtcgggtgtggcgacgcggaaataggccattgtatcgagaac

495

ori

Figure 11 continued

A↓wNI

550

605

660

715

770

825

880

935

990

on

on

on

on

on

on

AmpR

288 Trp His Lys Ile Leu Ser Ala Gly Phe Glu Ala Ile Glu Arg Asp

Figure 11 continued

ttcgttcatccatagttgcctgactccccgtcgtagataactacgatacgggs
 +-----+-----+-----+-----+-----+-----+-----+-----+
 aagcaagttaggttatcaacggactgaggggcagcacatctattgatgctatgcct
 . 379 . 385 . 388 . 394 . 255 .
 Arg Gln Asp Met Thr Ala Gln Ser Gly Thr Thr Tyr Ile Val Val Ile Arg Ser
 ← AmpR

gggcttaccatctggccccagtgctgcaatgataaccggcgagacccacgtcaccg
 +-----+-----+-----+-----+-----+-----+-----+
 cccgaatggtagaccggggtaacggactatggcgctctgggtgcggagtgcc
 . 250 . 245 . 240 . 240 .
 Pro Lys Gly Asp Pro Gly Leu Ala Ile Ile Gly Arg Ser Gly Arg Glu Gly
 ← AmpR

gctccagatttatcagcaataaaccaggccggaaaggcccggcggcagaatg
 +-----+-----+-----+-----+-----+-----+-----+
 cgagggtctaaatagtcgttatttgcgtcgcccttccggctcgctgtttcac
 . 236 . 230 . 228 . 220 .
 Ala Gly Ser Lys Asp Ala Ile Phe Trp Gly Ala Pro Leu Ala Ser Arg Leu Leu Pro
 ← AmpR

gtcttgcactttatccgcctccatcccgatctattttgttgcggggasgtcg
 +-----+-----+-----+-----+-----+-----+-----+
 caggacgttcaaataataggcgaggtaggtcagataattaacaacggcccttcgtc
 . 218 . 210 . 208 . 200 .
 Gly Ala Val Lys Asp Ala Glu Met Trp Asp Ile Ile Gln Gln Arg Ser Ala Leu
 ← AmpR

agtaagttagtttgtccaggtaatagtttgcgcacgttgttgtccattgttacaggc
 +-----+-----+-----+-----+-----+-----+-----+
 tcattcatcaagcggtcaattatcaaacgcgttgcacacaacggtaacgttccg
 . 195 . 190 . 185 .
 Thr Ile Ile Gln Gly Thr Ile Ile Lys Arg Ile Thr Thr Ala Met Ala Val Pro
 ← AmpR

atcgtgggtgcacgtcgctgtttggatggcttcattcgatccgggtccccacc
 +-----+-----+-----+-----+-----+-----+-----+
 tagcaccacagtgcgagcaggcaaccataccgaagtaagtgcgaggccsagggttg
 . 180 . 175 . 170 . 165 .
 Met Thr Thr Asp Arg Glu Asp Asn Pro Ile Ala Glu Asn Ile Glu Pro Gln Trp Arg
 ← AmpR

gatcaaggcgagttacatgtatccccatgttgtgcaaaaaagcggttagtttt
 +-----+-----+-----+-----+-----+-----+-----+
 ctatgttccgtcaatgtacttaggggtacaacacgtttttcgccaaatcgaggaa
 . 160 . 155 . 150 . 145 .
 Asp Leu Arg Thr Val His Asp Gly Met Asn His Ile Phe Ala Thr Ile Gln Lys
 ← AmpR

Figure 11 continued

cggctccgatcggttcagaagtaagtggccgcgtgttatcactcatgggt
 +-----+-----+-----+-----+-----+-----+-----+-----+
 gccaggaggctaccaacagtcttcaitcaaccggcgatcacaatagttagtaccsa
 +-----+-----+-----+-----+-----+-----+-----+-----+
 Pro Gly Gly Ile Thr Thr Leu Leu Leu Asn Ala Ala Asp Thr Asn Asp Ser Met Thr
 ← AmpR

atggcagcactgcataattcttactgtcatgccatccgtaaagatgctttctg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 taccgttgtacgtattasgagaatgacagttacggtaggcatttacgaaaagac
 +-----+-----+-----+-----+-----+-----+-----+-----+
 His Ala Ala Ser Cys Leu Glu Arg Val Thr Met Gly Asp Thr Leu His Lys Glu Thr
 ← AmpR

Scal

tgactggtgagttactcaaccaagtcattctgagaatagtgtatgcggcgaccgag
 +-----+-----+-----+-----+-----+-----+-----+-----+
 actgaccactcatgagttggttcaagactttatcacatascggcgctggctc
 +-----+-----+-----+-----+-----+-----+-----+-----+
 Val Pro Ser Tyr Glu Val Leu Asp Asn Gln Ser Tyr His Ser Arg Arg Gly Leu
 ← AmpR

ttgtctttccccggcggtcaatacggataatacggcgccacataggagaacttta
 +-----+-----+-----+-----+-----+-----+-----+-----+
 aacgagaacggcccgcagttatgccctattatggcgccgtgtatcgctttgasat
 +-----+-----+-----+-----+-----+-----+-----+-----+
 Gln Glu Gln Gln Asp Asp Ile Arg Ser Leu Val Asp Gly Cys Leu Leu Val Lys
 ← AmpR

XmnI

aaagtgtcatcattggaaaaacgtttttggggcgaaaaactctcaaggatcttac
 +-----+-----+-----+-----+-----+-----+-----+-----+
 ttccacyatgttacccctttgtcaagaagccccgttttggagatgttccagaatg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 Phe Thr Ser Met Met Pro Phe Arg Glu Glu Pro Arg Phe Ser Glu Leu Ile Lys Gly
 ← AmpR

cgttgttggatccaggatcgatgtaacccactcgatgcacccactgtatcc
 +-----+-----+-----+-----+-----+-----+-----+-----+
 gcgacaaactctaggatcaagctacattgggtgaggcacgtgggttggacttagaaatcg
 +-----+-----+-----+-----+-----+-----+-----+-----+
 Ser Asn Leu Asp Leu Glu Ile Tyr Gly Val Arg Ala Gly Leu Gln Asp Glu Ala
 ← AmpR

Figure 11 continued

atcttttactttcaccagcggttctgggtgagccaaaacaggaaggcaaaatgcc
 tagaaaaatgaaagtggtcgcaagaccactcgttttgtccttccgtttacgg 1760

Asp Lys Val Lys Val Leu Thr Glu Pro His Ala Phe Val Pro Leu Cys Phe Ala
 signal sequence

← AmpR

gcaaaaaaggaaataagggcgacacggaaatgttgaatactcatactttcctt 1815

cgtttttcccttattccgcgtgtgcctttacaacttatgagtagatgagaaggaaa

Ala Phe Phe Pro Ser Leu Asp Val Arg Phe His Gln Ser Met
 signal sequence AmpR AmpR promoter

ttaaatattatttgaagcatttatcagggttattgtctcatgagcggtataatt 1870

sagttataataacttcgttaatagtcoccaataacagagtgacttcgttatgtataa

← AmpR promoter

tgaatgtatTTTtagaaaaatacacaatagggttccggcacattccccgaaaa 1925

acttacataaaatctttttttgttatacccaaggcgcgttaaaggggcttt

← AmpR promoter

ZraI AatII

gtgcacacctgacgtcteaagaaaccattattatcatgacattaaacctataaaaata 1980

cacggtgactgcagatttttgtaataatagtgactgttaattggatattttat

EcoO109I

ggcgtatcacgaggcccttcgtctcgccgtttcggtgtgtgacggtaaaaac 2035

ccgcatagtgtccggaaagcagagcgcgcaagccactactgcacttttga

ctgacacatgcagetccggagacggtcacagettgtctgttaagcggatgccgg 2090

gactgtgtacgtcgagggccctctgcctgtgtgcacagacattgcgttacggcc

sgcagacaagcccgtcaggcgccgtcagcggtgtgggggtgtggggctggc 2145

tcgttcgttcggcagtcggccacacccacagccccacagccccacgg

ttaactatgcggcatcagagcagattgtactgagagtgcaccataGTGTTGGCGG 2200

aattgataacgcgttgtctgttaacatgacttcacgttgtatCACAAACCGCC

GTGTCGGGCCTGGCTTAACTATGCCCATCAGAGCAGATTGTACTGAGAGTGCAC 2255

CACAGCCCCGACCGAATTGATAACGCCGTAGTCTCGTCTAACATGACTCTCACGTG

Figure 11 continued

BstAPI
 CATAcgccgtgtgaaatcccgccaggatgcgttaaggagaaaataccgcattcagge
 2310
 CTATacggccacactttatggcgtgtctacgcattcttttatggcgttagtcg
 gattccacatccaaataaatcatacaggccaggccaaagaatttagcasaattaaag
 2365
 ctaaggtttttagtttttttttttttttttttttttttttttttttttttttttt
 aataaaggccctcagagcatcagactaaatcggttgtactaaaaacatattatgaccc
 2420
 ttatttggagtttcgtatttcgattttagccaaatggttttgtaaatctggga
 gtaatcacittttgtgggagaaggccctttatitcaacgcacggataaaaaattttttag
 2475
 cattatgaaaaacggcccttttggaaatcaggatgttggttttttttttttttttt
 acccatatatatttaatgcataatgcctyagtaatgtgttaggttttttttttttt
 2530
 tgggagttatataaaatttacgtttacggactcattacacatccatttttttttttt
 gggtgagaaaggccggagacagttccaaatcaccatcaatatgtatatttttttttt
 2585
 cccacttttccggccctttgtcgttttttttttttttttttttttttttttttttt
 tagctgateaaatttttttttttttttttttttttttttttttttttttttttttttt
 atcgactattttttttttttttttttttttttttttttttttttttttttttttttt
 gcttttttttttttttttttttttttttttttttttttttttttttttttttttttt
 2640
 cgtatagtccaggtaacggacttcggcccttttttttttttttttttttttttttt
 smal
 gggggggggggggccacttttttttttttttttttttttttttttttttttttttt
 2695
 cccccccccccccccgggttttttttttttttttttttttttttttttttttttt
 smal
 gggggggggggggccacttttttttttttttttttttttttttttttttttttt
 2750
 cccccccccccccccgggttttttttttttttttttttttttttttttttttttt
 smal
 gggggggggggggccacttttttttttttttttttttttttttttttttttttt
 2805
 cccccccccccccccgggttttttttttttttttttttttttttttttttttttt

Figure 11 continued

Figure 11 continued

tcaacaaactccaggcaggaccatgtgcggcgcttcgttggggcttttgttcag
 agtgcgtgggtggteetggtaactaactagcggcgggggggggggggggggggggg
 3298
 770 213 219
 Val The Ile Leu Lys Val Met His Asp Arg Lys Glu Asn Pro Asp Lys Ser Leu

 ggggggactgggtgctcaggtagtggttgtccccccggcgggggggggggggggggg
 cggccctgaccacacggatccatcaccaaacggccgggtgtcgccgggggggggg
 3300
 205 296 189 196
 Ala Ser Gln Thr Ser Ieu Tyr His Asn Asp Pro Leu Leu Val Pro Gly Asp Gly

 atgggggtttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 tacccccacaaagacgacatcaccaacggccgtcgacgtgggggggggggggggg
 3302
 383 180 178 179
 His Pro Thr Asn Gln Gln Tyr His Asp Ala Leu Glu Val Ser Gly Asp Glu Ile Asn

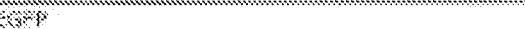
 tggggggatcttggggatggggatggggatggggatggggatggggatggggatgg
 acaccggcttggggatggggatggggatggggatggggatggggatggggatgg
 3410
 165 166 155
 His Arg Ile Lys Pro Asn Val Lys He Ser Asn Lys Glu Lys Asp Ala Met Ile

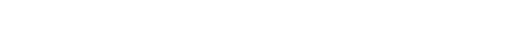
 atagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 tataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 3465
 182 185 180 185
 Tyr Val Asn His Ser Asn Tyr Asp Ile Glu Leu Lys His Gly Ile Leu Asn Gly

 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
 aggaggaggaggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg
 3500
 130 125 120 115
 Asn Glu Lys The Asp Ile Gly Ile Leu Glu Ile Asp Asn Val Leu Thr Asp Glu Glu

 cggatccatccatccatccatccatccatccatccatccatccatccatccatccatcc
 gtttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
 3575
 113 105 100
 Phe Lys Val Asn Arg Thr Ile Tyr Asp Asn Glu Asp Asp Ile Phe The Ile Thr

Figure 11 continued

ge gct cct ggacgt agect tgggc atggggactt gaaga agt cgt gct gtc
 cggaggaccatgcata gga a gcccgt acce gca tttc a cgc a cgc a a g
 35 50 65 80
 Arg Gln Glu Val Tyr Gly Glu Pro Met Ala Ser Lys Phe Phe Asp His Glu Lys

 at ggg tgg ggg ggt a ggg gct gaa gca ctg c a c g c g t a g g t c a g g t g g t c a c g a
 c a c a c c a g c c c c a t g c c g a c t t g t g a e g t g g g c a t e a g t c c c a c c a g t p o t
 75 79 65 60
 Met His Asp Pro Tyr Arg Ser Phe Cys Glu Val Gly Tyr Thr Leu Thr Thr Val Leu

 g g t g g g c c a g g g c a c g g g c a g t t g c c g g t g g t c a g a t g a a c t t c a g g g t x a g
 c c c a c c c g g t c c c g t g c c c g t c p a a c g g c c a c c a g t c t a c t t p a a g t c c c a g c c
 35 50 45
 Thr Pro Ile Pro Val Pro Leu Lys Gly Thr Thr Cys Ile Phe Lys Leu Thr Leu

 c t t g c c g t a g g t g g c a t c g c c c t c g c c g g a c a c q c t g a a c t t r g t g g c c
 g a a c g g c a t c c a c c g t a g g t g g g a g c g g g a g c g g c c t g t g c g a c t t g a a c a c c g c
 40 25 30 25
 Lys Gly Tyr Thr Ala Asp Gly Glu Gly Glu Ser Val Ser Phe Lys His Gly

BseRI
 t t t a c g t e g c c g t c c a g o t c g a c c a g g a t g g g c a c c a c c c g g t g a a c a g c t e a t
 s a a t g c a g c c c g c a g g t c g a g c t g g t c t a c c e g t g g t g g g c a c t t g t c g a g g a
 30 15 10 5
 Asn Val Asp Gly Asp Leu Glu Val Leu Ile Pro Val Val Gly Thr Phe Leu Glu Glu

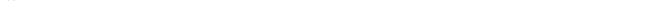
 c g c c c t t g c t c a c c a t t c c c t t c t t c c c c a a a c c c c t t c c g a a a a a g a a a c c
 g c g g g a a c g a g t g g t c c c g a c a c a a g a c c c c c t t c c c a a c g g t t t t t c t t g c
 1
 Gly Lys Ser Val Met

 T T C A C C G G C A C T A C T G C A C T T A T A T A C G G T T C T C C C C C A C C C T C G G G A A A R A G G C
 A A G T G C C C G T G A T G A C G T G A A T A T A T G C C A A G A G G G G G T G G G A G C C C T T T T C C G
 3360
 EF-1α core promoter

Figure 11 continued

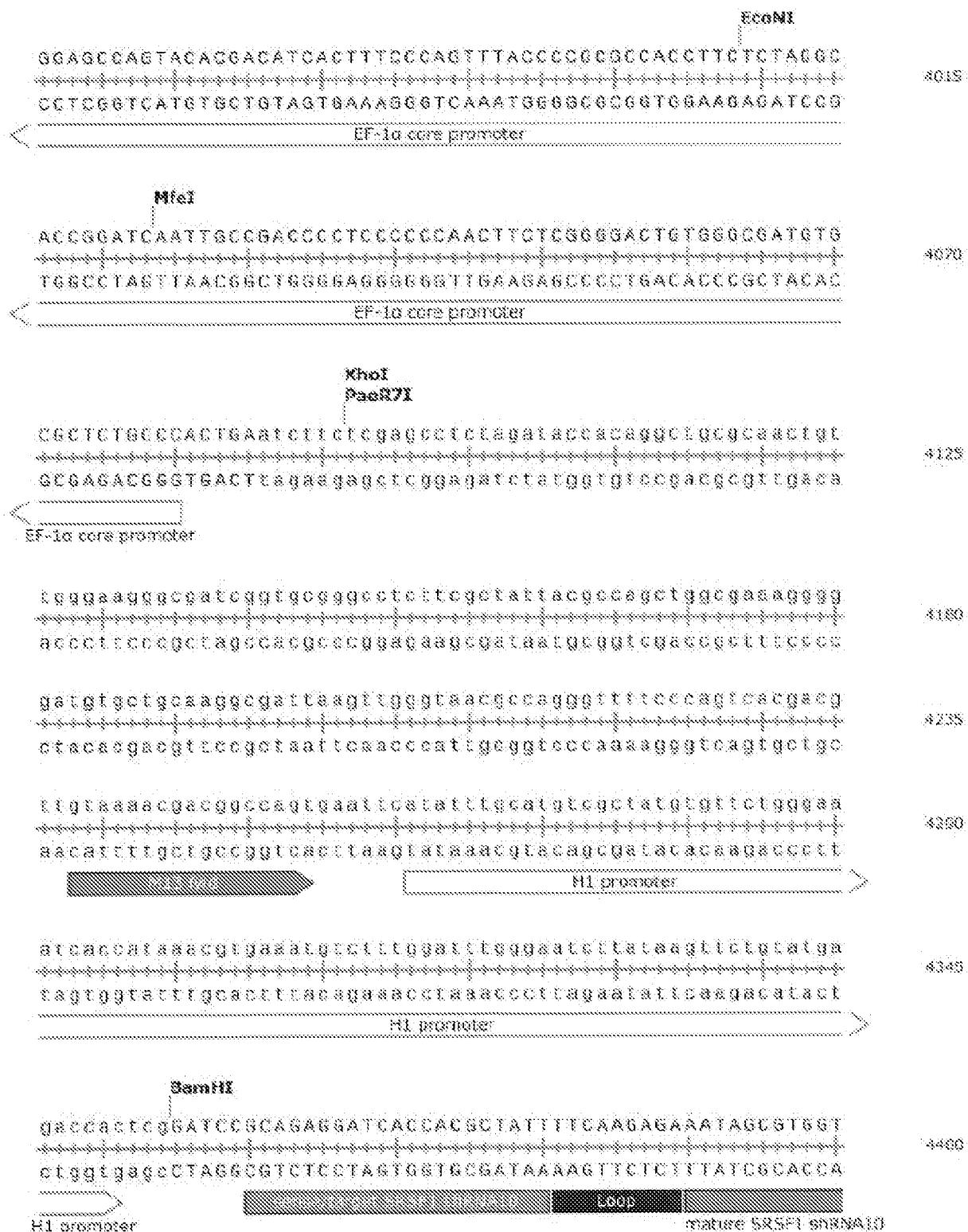


Figure 11 continued

Figure 11 continued

tttggatgttgttcaggtttgaaacaaagagtccactatcataaagaaccgtggactcca
 naccaacaacaaggtaaaaaacccttgcattcagggtgtataatttttgtcaacctgaaGTT
 ← f1.on

495C

DraIII

ACGTCAAAAGGGCGAAAAAACCGCTCTATCAGGGCGATGGCCCACTACGTAAACCATC
 TGCAGTTTCCCCTTTTGCAAGATAGTCCCCCTACCCGGTGTACTGACTTGGTAG
 ← f1.on

500C

ACCTTAATCAAGTTTTTGGGGTCAGGTGCCGTAACGCACTAAATCGCAACCC
 TGGGATTAGTTCAAAAAACCCCGAGCTCCACGGGATTTGGTGTAGTTAGCCTTGGGA
 ← f1.on

506C

NgoMIV NaeI

AAAGGGAGCCCCGATTTAGAGCTTGTACGGGGAAAGCCGGAAACGTGGCGAGAA
 TTTCCTCGGGGCTAAATCTCGAAGTGCCTTGGCCGCTTGTACCGCTCTT
 ← f1.on

511S

ACGAAGGGAAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGCT
 TCCCTCCCTTCTTGGCTTCTCGCCCGCGATCCCCGACCGTTCACATCGCCA
 ← f1.on

517S

CACCGCTGCCTAAACCACACACCCGCCGCCTTARTGGCCGCTACAGGGCGC
 GTBCGACCGCGCATTCGTGTGGCGGGCGAATTACGCAGCGATGGCCCGCC
 ← f1.on

522S

TACTATGGTTGCTTGTACGGCACGTATAACSTGCTTGTGTAGGTAGAG
 ATGATACCAACGAAACTGCTGTGCAATTGGCACGAAAGGGAGCAATCTTGTCTC
 ← f1.on

528S

CGGGAGCTAAACAGGAGGCCGATTAGGGATTGGTACACAGAAACGGTACGCCAG
 CCCCTCGATTTGTCTCCGGCTAAATTGCGCTAAATCTGTGTCTTGGCATGCCGCT
 ← f1.on

532S

AATGCTGAGAAGTGTGTTATAATCACTGAGGCCACCGAGTAAAAGAGSTCTGTC
 TTAGGACTCTTCAACAAAAATATTAGTCACTCCGGTGTGCTCATTTCTCAAGACAGA
 ← f1.on

535S

ATCACGCCAAATTACCGSTGTGCGCAATAACTTCTTGTATTACTAATAACATCACTT
 TAGTGGCTTTAATTGGCAACAGCGTTATGAAGAAACTAATCATTATTGTAGTGAA

544S

Figure 11 continued

GCCTGAGTAGAAGAACTCAAACATATCGGCCTTGCTGGTAATATCCAGAACAAATAT
 CGGACTCATCTTCTTGAGTTGATAAGCCCGAACGACCATTATAAGGTCTTGTATA 5500

 TACCGCCAGCCATTGCAACGGAATGCCATTGCCATTAGGCTGCGCAACTGTT 5555
 ATGGCGGTGGTAACGTTGCCCTAGCGGTAAAGCGGTAAAGTCCGACGCGTTGACAA

 GGGAAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGCATTAATGAAT 5610
 CCCTTCCCCGCTAGCCACGCCCCGGAGAACCGATAATGCCGTGACGTAATTACTTA

 CGGCCAACGGCGGGGAGAGGGGGTTGCGTATTGGGC 3'
 CGCGGTTGCGGCCCCCTCCGCCAACCGCATAACCGS 5'
 *** 5648

Figure 12

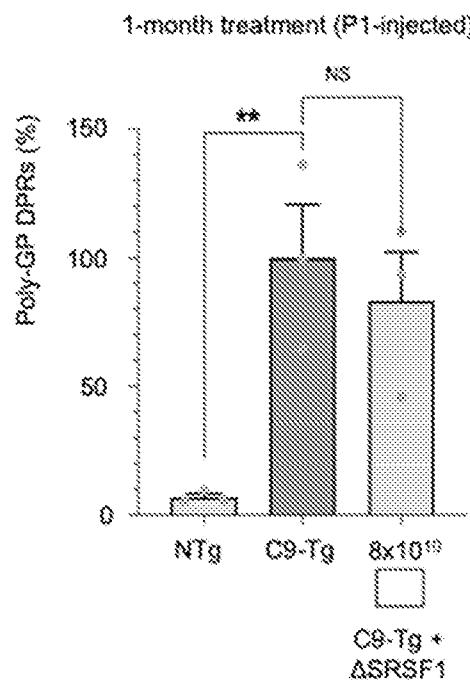
A

Figure 12 continued

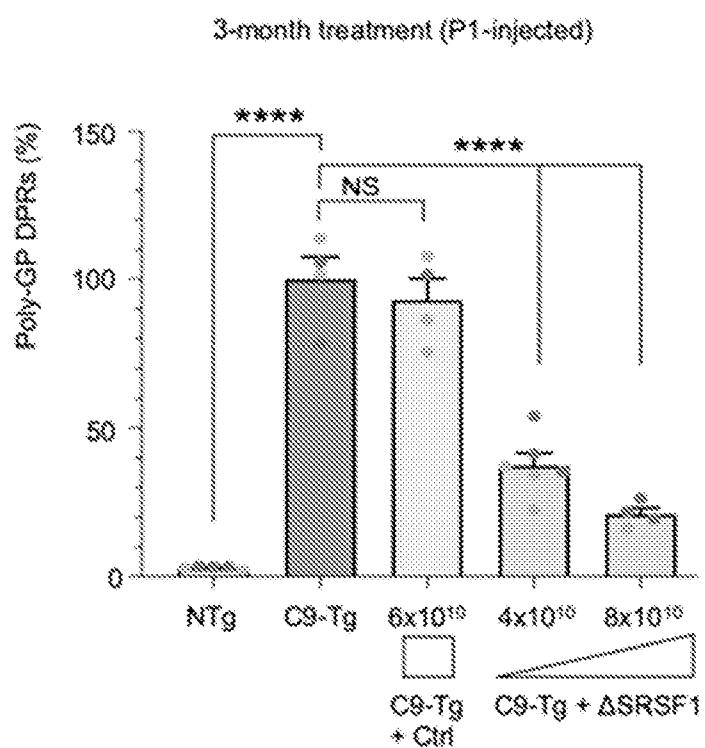
B

Figure 13

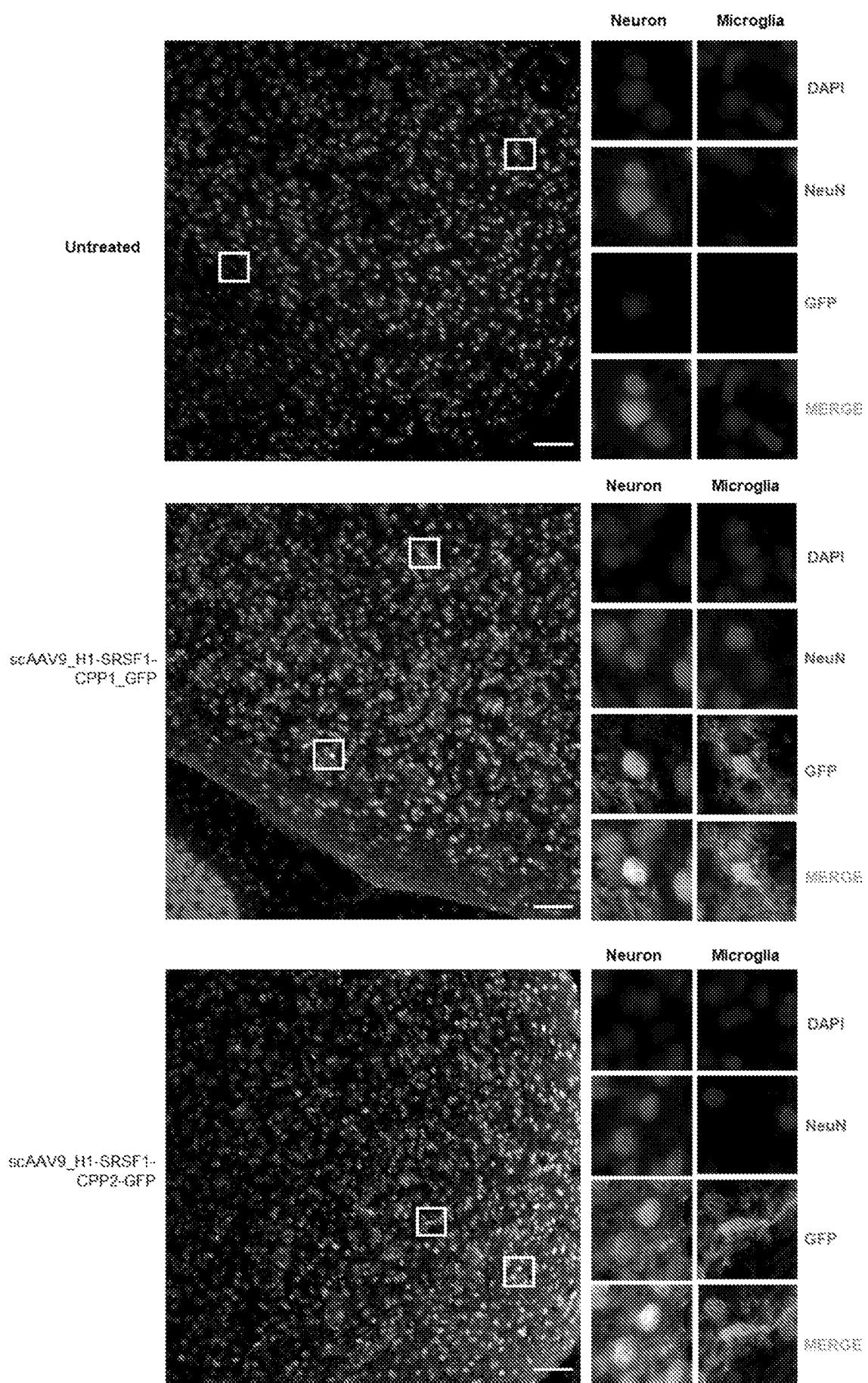
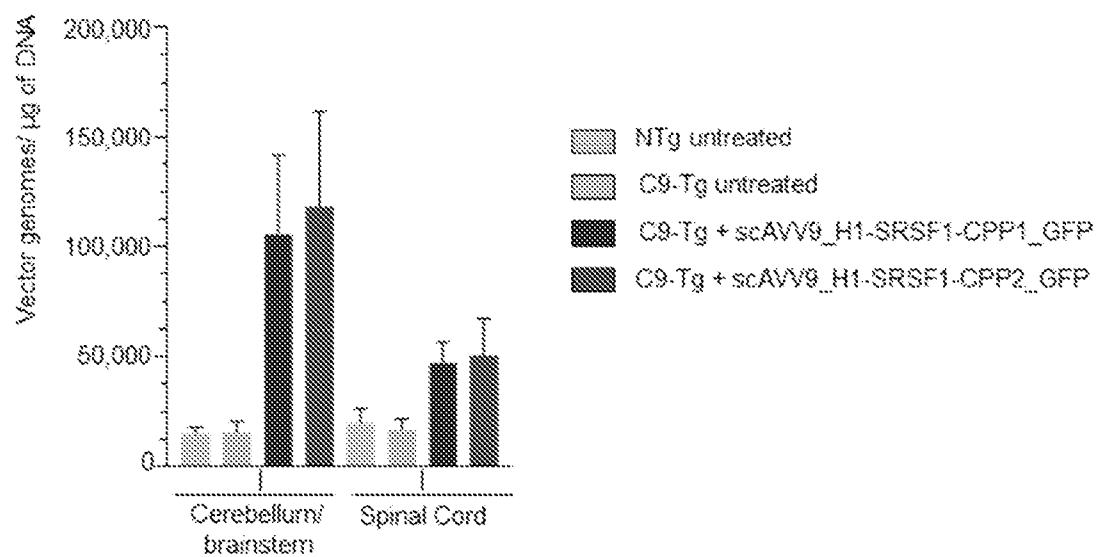
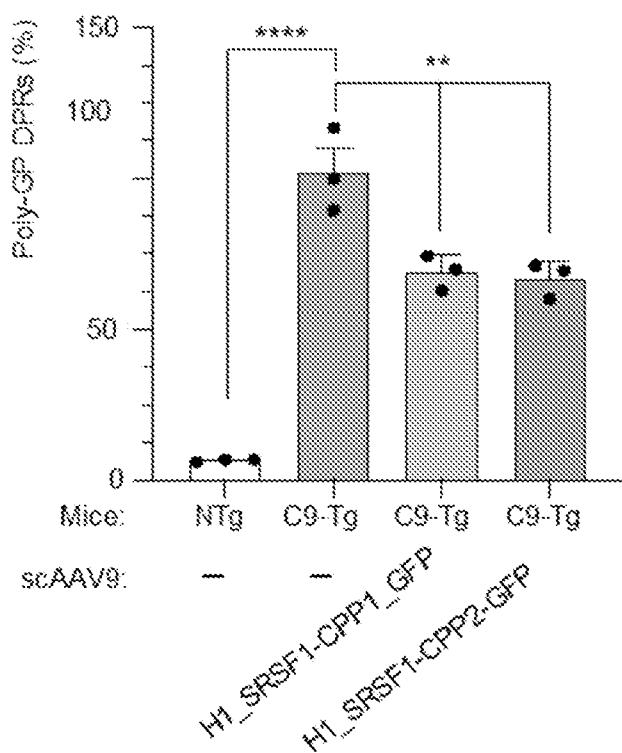


Figure 14

A**B**



Intellectual
Property
Office

Application No. GB2312083.5

RTM

Date : 5 February 2024

The following terms are registered trade marks and should be read as such wherever they occur in this document:

Glutamax
Opera
Phenix

Viral Vector

Field of the Disclosure

The present disclosure relates to antagonists that target, directly or indirectly, Serine/Arginine Rich Splicing Factor 1 (SRSF1); viral vectors comprising a nucleic acid sequence encoding SRSF1 antagonists. The use of said vector in gene therapy for the treatment of neurodegenerative diseases such as for example Amyotrophic Lateral Sclerosis (ALS) or sporadic Amyotrophic Lateral Sclerosis which is not caused by a pathological C9ORF72 hexanucleotide repeat expansion and methods thereof are also disclosed.

10 Background the Disclosure

Gene therapy aims to treat diseases long-term by the introduction of genetic material which alters cell function. Several gene therapy approaches exist such as the delivery of a functional gene to replace a faulty one, inactivation of toxic genes through gene silencing or antisense, introduction or overexpression of genes absent in the host and gene editing approaches. The 15 genetic material is most commonly delivered using viral based vectors such as adenoviruses (Ads), adeno-associated virus (AAVs), self-complementary AAVs and retroviruses i.e. lentiviruses.

The safety of gene therapy vectors requires particular attention as gene therapy vectors 20 persist in the patient's body over a long time and gene therapy vectors must be designed to reduce genotoxic effects, immune reactions or prevent activation of adjacent genes close to the integration site. The backbone of viral vectors typically comprises the protein capsid for packaging the expressed nucleic acid, the genetic information describing the expressed nucleic acid placed between inverted terminal repeats and elements such as promoter 25 elements which allow efficient expression in the host. When delivering genetic material of small size such as short hairpin RNA (shRNA) or antisense oligonucleotides, non-expressed "stuffer" nucleotide sequences are often required to increase the efficiency of shRNA or oligonucleotide nucleic acid targeting, expression and reach optimal packaging capacity.

30 Neurodegenerative diseases are typically caused by neuronal dysfunction or neuronal loss and affects millions of people worldwide. Neurodegenerative diseases are more prevalent in the aging populations and include but are not limited to amyotrophic lateral sclerosis (ALS), multiple sclerosis, Parkinson's disease, Alzheimer disease, motor neuron and Huntington's disease. ALS and frontotemporal dementia (FTD) are adult-onset neurodegenerative diseases 35 with no effective treatment. ALS is the most common form motor neuron disease (MND), a collective term for a group of neurological disorders characterised by degeneration and loss

of motor neurons. ALS is characterised by selective degeneration of the upper and lower motor neurons, leading to muscle wasting and premature death usually due to respiratory failure and paralysis. Around 90% of ALS cases are classified as sporadic, with approximately 10% showing a genetic component and familial inheritance. FTD is the second most-common form
5 of early-onset dementia characterised by a progressive loss of neuronal cells in frontal and temporal lobe leading to alterations in cognitive function and personality.

The most common genetic cause of ALS and FTD is a hexanucleotide repeat expansion of GGGGCC in the first intron of the chromosome 9 open reading frame 72 (C9orf72) gene,
10 termed C9ALS/FTD.

Antisense oligonucleotide therapies targeting C9ORF72 are in clinical trials and are aimed at reducing the expression of the repeat expansion, thus reducing RNA and DPR toxicity, without affecting the normal expression of C9orf72. Patent US10,801,027 demonstrates that depletion
15 of the export adaptor serine/arginine-rich splicing factor 1 (SRSF1) inhibits the nuclear export of pathological C9ORF72 repeat transcripts retaining hexanucleotide repeat expansions and is hereby incorporated by reference.

However, although depletion of SRSF1 works in patients with ALS caused by hexanucleotide
20 repeat expansions, the present disclosure identified that depletion of SRSF1 also confers neuroprotection in sporadic ALS cases which are not caused by a pathological C9ORF72 hexanucleotide repeat expansion.

Statement of the Invention

25 According to an aspect of the invention there is provided a viral vector comprising a transcription cassette for the expression of a nucleic acid molecule in a mammalian host cell wherein said nucleic acid molecule is operably linked to a promoter adapted to express said nucleic acid molecule in said mammalian host cell characterised in that said vector comprises
30 a non-expressed nucleotide sequence and wherein said nucleic acid molecule encodes an antagonistic agent that targets Serin/Arginine Rich Splice Factor (SRSF1) or an SRSF1 peptide sequence.

The non-expressed nucleotide sequence is typically referred to a "Stuffer" sequence. Stuffer
35 nucleotide sequences are known in the art and are non-expressed nucleotide sequences that provide optimal viral packaging of viral based vectors. Stuffer sequences are disclosed in PCT/US2013/031644 and is hereby incorporated by reference in its entirety. Stuffer nucleotide

sequences can be placed between the viral inverted terminal repeat sequences, either side of the transgene of interest or two stuffer sequences could be added on each side of the transgene of interest.

- 5 In a preferred embodiment of the invention said antagonistic agent is a polypeptide or peptide.

In a preferred embodiment of the invention said antagonistic agent is a nucleic acid-based agent.

- 10 In a preferred embodiment of the invention said nucleic acid-based agent is an antisense nucleic acid, an inhibitory RNA or shRNA or miRNA molecule that is complementary to and inhibits the expression of a nucleic acid encoding a Serin/Arginine Rich Splice Factor (SRSF1).

Preferably said SRSF1 comprises or consist of a sequence set forth in SEQ ID NO 67.

15

Alternatively, said SRSF1 comprises or consist of a sequence set forth in SEQ ID NO 76.

The nucleic acid-based agent is designed with reference to the sequence set forth in SEQ ID NO 67, or alternatively with reference to the sequence set forth in SEQ ID NO 76.

20

In a preferred embodiment of the invention said nucleic acid-based agent is an inhibitory RNA.

In a preferred embodiment of the invention said nucleic acid-based agent is an antisense RNA.

- 25 In a further preferred embodiment of the invention said inhibitory RNA is a shRNA or miRNA molecule.

A technique to specifically ablate gene function is through the introduction of double stranded RNA, also referred to as small inhibitory or interfering RNA (siRNA, shRNA and miRNA), into

- 30 a cell which results in the destruction of mRNA complementary to the sequence included in the siRNA molecule. The siRNA molecule comprises two complementary strands of RNA (a sense strand and an antisense strand) annealed to each other to form a double stranded RNA molecule. The siRNA molecule is typically derived from exons of the gene which is to be ablated. The mechanism of RNA interference is being elucidated. Many organisms respond
35 to the presence of double stranded RNA by activating a cascade that leads to the formation of siRNA. The presence of double stranded RNA activates a protein complex comprising RNase III which processes the double stranded RNA into smaller fragments (siRNAs,

approximately 21-29 nucleotides in length) which become part of a ribonucleoprotein complex. The siRNA acts as a guide for the RNase complex to cleave mRNA complementary to the antisense strand of the siRNA thereby resulting in destruction of the mRNA.

- 5 In a preferred embodiment of the invention said inhibitory RNA molecule is between 19 nucleotides [nt] and 29nt in length. More preferably still said inhibitory RNA molecule is between 21nt and 27nt in length. Preferably said inhibitory RNA molecule is about 21nt in length.
- 10 In a preferred embodiment of the invention said inhibitory RNA comprises or consists of a nucleotide sequence as set forth in SEQ ID NO: 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57 or 58.
- 15 In a preferred embodiment of the invention said shRNA comprises or consist of a nucleotide sequence selected from the group consisting of SEQ ID NO 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11.
In a preferred embodiment of the invention said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 7.
- 20 In a preferred embodiment of the invention said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 10.
In a preferred embodiment of the invention said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 11.
In a preferred embodiment of the invention said peptide comprises an amino acid sequence that is at least 10 amino acids in length and comprises all or part of the amino acid sequence set forth in SEQ ID NO: 59.
- 25 In a preferred embodiment of the invention said peptide comprises an amino acid sequence that is at least 32 amino acids in length and comprises the amino acid sequence set forth in SEQ ID NO: 59.
- 30 In a preferred embodiment of the invention said peptide is at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 29, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or at least 100 amino

acids in length but less than the full-length amino acid sequence set forth in SEQ ID NO: 60 or 61.

In a preferred embodiment of the invention said peptide consists of an amino sequence as set forth in SEQ ID NO: 59.

In an alternative embodiment of the invention said peptide is a dominant negative protein comprising a modification of the amino acid sequence set forth in SEQ ID NO: 60 or 61.

10 In a preferred embodiment of the invention said dominant negative protein comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 60 or 61 wherein said amino acid sequence is modified by addition, deletion or substitution of one or more amino acid residues.

15 In a preferred embodiment of the invention said modified protein comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 62 or 63.

In a preferred embodiment of the invention said nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide or peptide is set forth set forth in SEQ ID NO: 89, or a 20 sequence which is to 90% identical to the sequence set forth in SEQ ID NO 89.

In a further preferred embodiment of the invention said nucleic acid sequence is at least 36 nucleic acids in length.

25 In a preferred embodiment of the invention said peptide comprises an amino acid sequence that is at least 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40 or 42 amino acids in length and set forth in SEQ ID NO: 90.

30 In a preferred embodiment of the invention said peptide comprises an amino acid sequence that is set forth in SEQ ID NO: 75 (GSWQDLKDHMREA).

In a preferred embodiment of the invention said viral vector comprises a RNA Pol III terminator.

Preferably said terminator comprises the nucleic acid sequence 5' TTTTTT 3'.

35 In a preferred embodiment of the invention said vector comprises inverted terminal repeat nucleotide sequences.

Inverted terminal repeat sequences (ITR) are typically positioned upstream and downstream of a transcription cassette. Alternatively, the ITRs are upstream and downstream of the transcription cassette, the non-expressed nucleotide sequence and any optional regulatory elements.

5 In a preferred embodiment of the invention said ITR sequence is set forth in SEQ ID NO 64.

10 In a preferred embodiment of the invention said ITR sequence is set forth in SEQ ID NO 88.

In a preferred embodiment of the invention said promoter is selected from the group consisting of H1 Polymerase III promoter, U6 promoter, U7 promoter or the mammalian 7SK promoter.

15 In a further preferred embodiment of the invention said promoter is a H1 Polymerase III promoter.

In a preferred embodiment said H1 Polymerase III promoter is set forth in SEQ ID NO 65.

Viruses are commonly used as vectors for the delivery of exogenous genes. Commonly 20 employed vectors include recombinantly modified enveloped or non-enveloped DNA and RNA viruses, for example baculoviridae, parvoviridae, picornoviridae, herpesviridae, poxviridae, adenoviridae, picornnaviridae or retroviridae e.g. lentivirus. Chimeric vectors may also be employed which exploit advantageous elements of each of the parent vector properties (See 25 e.g., Feng, et al (1997) Nature Biotechnology 15:866-870). Such viral vectors may be wild-type or may be modified by recombinant DNA techniques to be replication deficient, conditionally replicating or replication competent. Conditionally replicating viral vectors are used to achieve selective expression in particular cell types while avoiding untoward broad-spectrum infection. Examples of conditionally replicating vectors are described in Pennisi, E. (1996) Science 274:342-343; Russell, and S.J. (1994) Eur. J. of Cancer 30A(8):1165-1171.

30 Preferred viral vectors are derived from the adenoviral, adeno-associated viral or retroviral genomes.

35 In a preferred embodiment of the invention said viral based vector is an adeno-associated virus [AAV].

In a preferred embodiment of the invention said adeno-associated virus is a self-complementary adeno-associated virus (scAAV).

In a preferred embodiment said viral based vector is selected from the group consisting of:

- 5 AAV2, AAV3, AAV6, AAV13; AAV1, AAV4, AAV5, AAV6, AAV9 and AAVrh10.

In a preferred embodiment said scAAV is selected from the group consisting of: scAAV2, scAAV3, scAAV6, scAAV13; scAAV1, scAAV4, scAAV5, scAAV6, scAAV9 and scAAVrh10.

- 10 In a preferred embodiment of the invention said viral based vector is scAAV9 or scAAVrh10.

In an alternative preferred embodiment of the invention said viral based vector is a lentiviral vector.

- 15 According to a further aspect of the invention there is provided a pharmaceutical composition comprising a viral vector according to the invention and an excipient or carrier.

The viral vector compositions of the present invention are administered in pharmaceutically acceptable preparations. Such preparations may routinely contain pharmaceutically

- 20 acceptable concentrations of salt, buffering agents, preservatives, compatible carriers and supplementary therapeutic agents. The expression vector compositions of the invention can be administered by any conventional route, including injection or by gradual infusion over time and in particular intrathecal (e.g., lumbar puncture) and/or intracerebral.

- 25 The viral vector compositions of the invention are administered in effective amounts. An “effective amount” is that amount of the expression vector that alone, or together with further doses, produces the desired response. In the case of treating a disease, the desired response is inhibiting the progression of the disease. This may involve only slowing the progression of the disease temporarily, although more preferably, it involves halting the progression of the
30 disease permanently. This can be monitored by routine methods. Such amounts will depend, of course, on the particular condition being treated, the severity of the condition, the individual patient parameters including age, physical condition, size and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are
35 well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical

judgment. It will be understood by those of ordinary skill in the art, however, that a patient may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons.

- 5 The viral vector compositions used in the foregoing methods preferably are sterile and contain an effective amount of expression vector according to the invention for producing the desired response in a unit of weight or volume suitable for administration to a patient. The doses of vector administered to a subject can be chosen in accordance with different parameters, in particular in accordance with the mode of administration used and the state of the subject.
- 10 Other factors include the desired period of treatment. If a response in a subject is insufficient at the initial doses applied, higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that patient tolerance permits. Other protocols for the administration of vector compositions will be known to one of ordinary skill in the art, in which the dose amount, schedule of injections, sites of injections, mode of
- 15 administration and the like vary from the foregoing. The administration of compositions to mammals other than humans, (e.g. for testing purposes or veterinary therapeutic purposes), is carried out under substantially the same conditions as described above. A subject, as used herein, is a mammal, preferably a human, and including a non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent.

20 When administered, the viral vector compositions of the invention are applied in pharmaceutically acceptable amounts and in pharmaceutically acceptable compositions. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active agent. Such preparations may routinely contain salts, buffering agents, preservatives, compatible carriers, and optionally other therapeutic agents' (e.g. those typically used in the treatment of the specific disease indication). When used in medicine, the salts should be pharmaceutically acceptable, but non-pharmaceutically acceptable salts may conveniently be used to prepare pharmaceutically acceptable salts thereof and are not excluded from the scope of the invention. Such

25 pharmacologically and pharmaceutically acceptable salts include, but are not limited to, those prepared from the following acids: hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, maleic, acetic, salicylic, citric, formic, malonic, succinic, and the like. Also, pharmaceutically acceptable salts can be prepared as alkaline metal or alkaline earth salts, such as sodium, potassium or calcium salts.

30 The pharmaceutical compositions containing the viral vectors according to the invention may contain suitable buffering agents, including acetic acid in a salt; citric acid in a salt; boric acid

in a salt; and phosphoric acid in a salt. The pharmaceutical compositions also may contain, optionally, suitable preservatives, such as: benzalkonium chloride; chlorobutanol; parabens and thimerosal.

The viral vector compositions may conveniently be presented in unit dosage form and may be

- 5 prepared by any of the methods well-known in the art of pharmacy. All methods include the step of bringing the active agent into association with a vector which constitutes one or more accessory ingredients. The preparation may be formulated according to known methods using suitable dispersing or wetting agents and suspending agents. The sterile injectable preparation also may be a sterile injectable solution or suspension in a non-toxic parenterally acceptable diluent or solvent, for example, as a solution in 1, 3-butanediol. Among the acceptable solvents that may be employed are water, Ringer's solution, and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any bland fixed oil may be employed including synthetic mono-or di-glycerides. In addition, fatty acids such as oleic acid may be used in the
- 10 preparation of injectables. Carrier formulation suitable for oral, subcutaneous, intravenous, intramuscular, etc. administrations can be found in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, PA.
- 15

According to a further aspect of the invention there is provided a viral vector according to the

- 20 invention for use as a medicament.

According to a further aspect of the invention there is provided a viral vector according to the invention for use in the treatment of a neurodegenerative disease.

- 25 In a preferred embodiment of the invention said neurodegenerative disease is selected from the group consisting of: amyotrophic lateral sclerosis (ALS) sporadic amyotrophic lateral sclerosis, familial ALS caused by a mutation other than a pathological C9ORF72-repeat expansion, frontotemporal dementia (FTD) motor neurone disease, frontotemporal lobar dementia (FTLD), Huntington's like disorder, and Fragile X-associated tremor/ataxia syndrome (FXTAS).
- 30

In a preferred embodiment of the invention said neurodegenerative disease is amyotrophic lateral sclerosis (ALS).

- 35 In a preferred embodiment of the invention said neurodegenerative disease is sporadic and/or familial amyotrophic lateral sclerosis.

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansions

- 5 In a preferred embodiment of the invention said neurodegenerative disease is sporadic frontotemporal dementia (FTD).

In a preferred embodiment of the invention said neurodegenerative disease is Fragile X-associated tremor/ataxia syndrome (FXTAS).

10

According to a further aspect of the invention there is provided a cell transfected with a viral vector according to the invention.

In a preferred embodiment of the invention said cell is a neurone and/or an astrocyte.

15

In a preferred embodiment of the invention said neurone is a motor neurone and/or an astrocyte.

20

According to a further aspect of the invention there is provided a method to treat or prevent a neurodegenerative disease comprising administering a therapeutically effective amount of a viral vector according to the invention to prevent and/or treat said neurodegenerative disease.

In a preferred method of the invention said neurodegenerative disease is sporadic amyotrophic lateral sclerosis and familial amyotrophic lateral sclerosis.

25

In a preferred method of the invention said neurodegenerative disease is amyotrophic lateral sclerosis.

30

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansions.

In a preferred method of the invention said neurodegenerative disease is sporadic frontotemporal dementia (FTD).

35

In a preferred method of the invention said neurodegenerative disease is Fragile X-associated tremor/ataxia syndrome (FXTAS).

According to a further aspect of the invention there is provided an isolated nucleic acid molecule encoding an shRNA molecule comprising or consisting of a nucleotide sequence selected from the group consisting of SEQ ID NO 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11.

- 5 The invention includes sequence variants corresponding to the recited SEQ ID. A sequence variant is one that varies from a reference sequence by 1, 2, 3, 4 or 5 nucleotide base changes.

In a preferred embodiment of the invention said nucleic acid molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 7.

10

In a preferred embodiment of the invention said nucleic acid molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 10.

15

In a preferred embodiment of the invention said nucleic acid molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 11.

According to a further aspect of the invention there is provided shRNA molecules comprising a nucleotide sequence, or variant thereof, selected from the group consisting of:

SRSF1-shRNA1 (SEQ ID NO 91):

20 GCUGAUGUUUACCGAGAUGGC UUCAAGAGA GCCAUCUCGGUAAACAUCAGC;
SRSF1-shRNA2 (SEQ ID NO 92):

GGAGUUUGUACGGAAAGAAGA UUCAAGAGA UCUUCUUUCCGUACAAACUCC;

SRSF1-shRNA3 (SEQ ID NO 93):

GGAAAGAAGAU AUGACCUAUG UUCAAGAGA CAUAGGUCAUAUCUUCUUUCC;

25 SRSF1-shRNA4 (SEQ ID NO 94):

GAAAGAAGAU AUGACCUAUGC UUCAAGAGA GCAUAGGUCAUAUCUUCUUUC;

SRSF1-shRNA5 (SEQ ID NO 95):

GCCUACAUCCGGGUAAAAGUU UUCAAGAGA AACUUUAACCCGGAUGUAGGC;

SRSF1-shRNA6 (SEQ ID NO 96):

30 GGGCCCAGAAGUCCAAGUUUAU UUCAAGAGA AUAAACUUGGACUUCUGGGCCC;
SRSF1-shRNA7 (SEQ ID NO 97):

GGCCCAGAAGUCCAAGUUUAUG UUCAAGAGA CAUAACUUGGACUUCUGGGCC;

SRSF1-shRNA8 (SEQ ID NO 98):

GCCCAGAAGUCCAAGUUUAGG UUCAAGAGA CCAUAACUUGGACUUCUGGGC;

35 SRSF1-shRNA9 (SEQ ID NO 99):

GGAAGAUCUCGAUCUCGAAGC UUCAAGAGA GCUUCGAGAUCGAGAUCUUCC; and

SRSF1-shRNA10 (SEQ ID NO 100):

GCAGAGGAUCACCACGUUU UUCAAGAGA AAUAGCGUGGUGAUCCUCUGC.

In a preferred embodiment of the invention said shRNA molecule comprises or consists of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 96.

5

In a preferred embodiment of the invention said shRNA molecule comprises or consists of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 99.

10 In a preferred embodiment of the invention said shRNA molecule comprises or consists of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 100.

According to an aspect of the invention there is provided an isolated nucleic acid molecule or shRNA according to the invention for use as a medicament.

15 According to a further aspect of the invention there is provided an isolated nucleic acid molecule or shRNA according to the invention for use in the treatment of a neurodegenerative disease.

20 In a preferred embodiment of the invention said neurodegenerative disease is selected from the group consisting of: amyotrophic lateral sclerosis (ALS) sporadic amyotrophic lateral sclerosis, familial ALS caused by a mutation other than a pathological C9ORF72-repeat expansion, frontotemporal dementia (FTD) motor neurone disease, frontotemporal lobar dementia (FTLD), Huntington's like disorder, and Fragile X-associated tremor/ataxia syndrome (FXTAS).

25

In a preferred embodiment of the invention said neurodegenerative disease is amyotrophic lateral sclerosis (ALS).

30 In a preferred embodiment of the invention said neurodegenerative disease is sporadic and/or familial amyotrophic lateral sclerosis.

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansions

35 In a preferred embodiment of the invention said neurodegenerative disease is sporadic frontotemporal dementia (FTD).

In a preferred embodiment of the invention said neurodegenerative disease is Fragile X-associated tremor/ataxia syndrome (FXTAS).

According to an aspect of the invention there is provided an siRNA molecule comprising or
5 consisting of a nucleic acid sequence designed with reference to the shRNA set forth in SEQ
ID NO 77-86.

According to an aspect of the invention there is provided an siRNA molecule according to the
invention for use as a medicament.

10 According to a further aspect of the invention there is provided an siRNA molecule according
to the invention for use in the treatment of a neurodegenerative disease.

In a preferred embodiment of the invention said neurodegenerative disease is selected from
15 the group consisting of: amyotrophic lateral sclerosis (ALS) sporadic amyotrophic lateral
sclerosis, familial ALS caused by a mutation other than a pathological C9ORF72-repeat
expansion, frontotemporal dementia (FTD) motor neurone disease, frontotemporal lobar
dementia (FTLD), Huntington's like disorder, and Fragile X-associated tremor/ataxia
syndrome (FXTAS).

20 In a preferred embodiment of the invention said neurodegenerative disease is amyotrophic
lateral sclerosis (ALS).

25 In a preferred embodiment of the invention said neurodegenerative disease is sporadic and/or
familial amyotrophic lateral sclerosis.

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused
by pathological C9ORF72-repeat expansions

30 In a preferred embodiment of the invention said neurodegenerative disease is sporadic
frontotemporal dementia (FTD).

In a preferred embodiment of the invention said neurodegenerative disease is Fragile X-
associated tremor/ataxia syndrome (FXTAS).

35 According to an aspect of the invention there is provided a cell penetrating polypeptide
comprising or consisting of an amino acid sequence set forth in SEQ ID NO 90.

In a preferred embodiment of the invention said polypeptide is between 12-42 or preferably between 13-42 amino acids in length.

- 5 In a further preferred embodiment of the invention said polypeptide comprises or consist of an amino acid sequence set forth in SEQ ID NO 75.

According to an aspect of the invention there is provided a polypeptide according to the invention for use as a medicament.

10

According to a further aspect of the invention there is provided a polypeptide according to the invention for use in the treatment of a neurodegenerative disease.

- In a preferred embodiment of the invention said neurodegenerative disease is selected from
15 the group consisting of: amyotrophic lateral sclerosis (ALS) sporadic amyotrophic lateral sclerosis, familial ALS caused by a mutation other than a pathological C9ORF72-repeat expansion, frontotemporal dementia (FTD) motor neurone disease, frontotemporal lobar dementia (FTLD), Huntington's like disorder, and Fragile X-associated tremor/ataxia syndrome (FXTAS).

20

In a preferred embodiment of the invention said neurodegenerative disease is amyotrophic lateral sclerosis (ALS).

- In a preferred embodiment of the invention said neurodegenerative disease is sporadic and/or
25 familial amyotrophic lateral sclerosis.

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansions

- 30 In a preferred embodiment of the invention said neurodegenerative disease is sporadic frontotemporal dementia (FTD).

In a preferred embodiment of the invention said neurodegenerative disease is Fragile X-associated tremor/ataxia syndrome (FXTAS).

35

According to a further aspect of the invention there is provided an antagonistic agent comprising a nucleic acid molecule wherein said nucleic acid molecule comprises a

nucleotide sequence designed with reference to human Serine/Arginine Rich Splice Factor (SRSF1) and wherein said nucleic acid molecule inhibits expression of SRSF1.

In a preferred embodiment of the invention said nucleic acid molecule is a double stranded
5 nucleic acid molecule comprises a sense strand and an antisense strand comprising a nucleotide sequence wherein said antisense nucleotide strand is adapted to anneal by complementary base pairing to a nucleic acid molecule encoding human SRSF1.

In a preferred embodiment of the invention said double stranded nucleic acid molecule is
10 RNA. Preferably, said RNA is siRNA or miRNA.

In an alternative embodiment of the invention said nucleic acid molecule is a single stranded nucleotide sequence comprising an antisense nucleotide sequence wherein said antisense nucleotide sequence is adapted to anneal by complementary base pairing to a nucleic acid
15 molecule encoding SRSF1.

In a preferred embodiment of the invention said single stranded nucleic acid is DNA.

In a further preferred embodiment of the invention said single stranded nucleic acid is DNA
20 and/or RNA.

Preferably, said DNA and/or RNA is a therapeutic antisense oligonucleotide such as an antisense oligonucleotide, a splice-switching oligonucleotide, a gapmer or similar.

25 Preferably said DNA is an antisense oligonucleotide.

In a preferred embodiment of the invention said nucleic acid molecule encoding human SRSF1 is set forth in SEQ ID NO: 67.

30 In a preferred embodiment of the invention said antagonistic agent comprises a nucleic acid molecule that is at least 15 nucleotides in length.

In a preferred embodiment of the invention said antagonistic agent comprises a nucleic acid molecule comprising a nucleotide sequence set forth in SEQ ID NO: 67 wherein said nucleic
35 acid molecule is a double stranded inhibitory RNA and is 19-23 nucleotides in length.

In a preferred embodiment of the invention said antagonistic agent comprises a nucleic acid molecule comprises modified nucleotides.

In a preferred embodiment of the invention said double stranded nucleic acid molecule comprising sense and antisense nucleic acid molecules comprise modified nucleotides.

- 5 In a preferred embodiment of the invention said modified nucleotides/sugars are selected from the group: a 3'-terminal deoxy-thymine (dT) nucleotide, a 2'-0-methyl modified nucleotide, a 2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, a locked nucleotide, an unlocked nucleotide, a conformationally restricted nucleotide, a constrained ethyl nucleotide, an abasic nucleotide, a 2'-amino-modified nucleotide, a 2'-0-allyl-modified nucleotide, 2'-C-alkyl-modified nucleotide, 2'-hydroxyl-modified nucleotide, a 2'-methoxyethyl modified nucleotide, a 2'-0-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, a non-natural base comprising nucleotide, a tetrahydropyran modified nucleotide, a 1,5-anhydrohexitol modified nucleotide, a cyclohexenyl modified nucleotide, a nucleotide comprising a phosphorothioate group, a nucleotide comprising phosphorodithioate (PS2), a nucleotide comprising a methylphosphonate group, a nucleotide comprising a 5'-phosphate, and a nucleotide comprising a 5'-phosphate mimic, for example a 5'-vinyl phosphate, a nucleotide comprising a 2'-deoxy-2'-fluro and a 2' methyl sugar base.
- 10
- 15

In a preferred embodiment of the invention said double stranded nucleic acid molecule comprising sense and antisense nucleic acid molecules comprise modified sugar(s).

- 20 In a preferred embodiment of the invention said modified sugar is selected from the group: a modified version of the ribosyl moiety, such as -O- modified RNA such as 2'-O-alkyl or 2'-O-(substituted)alkyl e.g. 2'-0-methyl, T-0-(2- cyanoethyl), 2'-0-(2-methoxy)ethyl (2'-MOE), 2'-0-(2-thiomethyl)ethyl, 2'-O-butyryl, -O- propargyl, 2'-O-allyl, 2'-O-(2-amino)propyl, 2'-O-(2-(dimethylamino)propyl), 2'-O-(2- amino)ethyl, 2'-O-(2-(dimethylamino)ethyl); 2'-deoxy (DNA);
- 25 2'-O-(haloalkoxy)methyl, e.g. 2'-0-(2-chloroethoxy)methyl (MCEM), -O- (2,2-dichloroethoxy)methyl (DCEM); 2'-<3-alkoxycarbonyl e.g. T-0-[2- (methoxycarbonyl)ethyl] (MOCE), 2'-O-[2-(N-methylcarbamoyl)ethyl] (MCE), T-0-[2-(N,N- dimethylcarbamoyl)ethyl] (DCME); 2'-halo e.g. 2'-F, FANA (2'-F arabinosyl nucleic acid); carbasugar and azasugar modifications; 3'-O-alkyl e.g. 3'-0-methyl, 3'-0-butyryl, V-O- propargyl and their derivatives.
- 30 In a preferred embodiment of the invention said antagonistic agent comprises or consists of a nucleotide sequence designed with reference to the target nucleic acid sequences selected from the group:

- TGGCACTGGTGTGTCGTGGAGTTGTA (SEQ ID NO 110);
TGGTGTGTCGTGGAGTTGTACGGAAA (SEQ ID NO 111);
TCGTGGAGTTGTACGGAAAGAAGA (SEQ ID NO 112);
AAGATATGACCTATGCAGTCGAAA (SEQ ID NO 113);
5 GAGAAAAGTGCCTACATCCGGGTTAA (SEQ ID NO 114);
CGGGTTAAAGTTGATGGGCCAGAA (SEQ ID NO 115);
TGATGGGCCAGAACAGTCCAAGTTAT (SEQ ID NO 116);
CAGAAGTCCAAGTTATGGAAGATCT (SEQ ID NO 117);
GAGAAGCAGAGGATACCACGCTAT (SEQ ID NO 118); and
10 CGTCATAGCAGATCTCGCTCTCGTA (SEQ ID NO 119).

In a preferred embodiment of the invention said antagonistic agent comprises a nucleic acid molecule comprising a nucleotide sequence wherein said nucleic acid molecule is a double stranded inhibitory RNA and is 19-23 nucleotides in length.

15

According to a further aspect of the invention there is provided a pharmaceutical composition comprising an antagonist agent according to the invention according and including an excipient or carrier.

20

According to a further aspect of the invention there is provided an antagonistic agent according to the invention for use as a medicament.

25

According to a further aspect of the invention there is provided an antagonistic agent to the invention for use in the treatment of a neurodegenerative disease.

In a preferred embodiment of the invention said neurodegenerative disease is amyotrophic lateral sclerosis (ALS).

30

In a preferred embodiment of the invention said neurodegenerative disease is sporadic and/or familial amyotrophic lateral sclerosis.

In a preferred embodiment of the invention said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansion.

35

In an alternative preferred embodiment of the invention said neurodegenerative disease is sporadic frontotemporal dementia (FTD).

In an alternative preferred embodiment of the invention said neurodegenerative disease is

- 5 Fragile X-associated tremor/ataxia syndrome (FXTAS).

Throughout the description and claims of this specification, the words "comprise" and "contain" and variations of the words, for example "comprising" and "comprises", means "including but not limited to", and is not intended to (and does not) exclude other moieties, additives,

- 10 components, integers or steps. "Consisting essentially" means having the essential integers but including integers which do not materially affect the function of the essential integers.

Throughout the description and claims of this specification, the singular encompasses the plural unless the context otherwise requires. In particular, where the indefinite article is used,

- 15 the specification is to be understood as contemplating plurality as well as singularity, unless the context requires otherwise.

Features, integers, characteristics, compounds, chemical moieties or groups described in conjunction with a aspect, embodiment or example of the invention are to be understood to

- 20 be applicable to any other aspect, embodiment or example described herein unless incompatible therewith.

An embodiment of the invention will now be described by example only and with reference to the following figures:

Figure 1. Timeline for differentiation and co-culture of motor neurons and astrocytes derived from healthy control and sporadic ALS (sALS) patients;

Figure 2. Images show that MNs treated with lentivirus expressing SRSF1-miRNA retain processes/axons characteristic of neurons compared to MN treated with LV_Ctrl-miRNA which degenerate and die;

- 30 Figure 3. Bar charts show MN survival expressed as a ratio of MNs quantified at counting day 3 over day 1 (%). 2-way ANOVA with Tukey's multiple comparison test; NS: non-significant;
: p<0.01; *: p<0.001; ****: p<0.0001;

Figure 4 Western immunoblotting shows that all 3 shRNAs lead to efficient depletion of SRSF1 and inhibition of the RAN translation of V5-tagged DPRs;

Figure 5. Bar charts represents mean±sem (2-way ANOVA with Tukey's multiple comparison test; NS: non-significant; ****: p<0.0001; n=3 biological replicates). Quantification in 3 independent triplicate experiments;

Figure 6. C9ORF72-ALS/FTD mice were injected via cisterna magna at post-natal day 1 (P1) with either 8×10^{10} scAAV9_Ctrl-shRNA_GFP vector genomes (vg) or 6×10^{10} scAAV9_SRSF1-shRNA10_GFP vg. Animals were sacrificed 1 month and 3 months post injections. Western blots show that the scAAV9_SRSF1-shRNA10_GFP virus leads to specific depletion of SRSF1 in C9ORF72-ALS/FTD mice as well as in wild type C57BL/6 mice (not shown) while the Ctrl-shRNA has no effect. GAPDH is used as a loading control;

Figure 7: map of scAAV_SRSF1 132-144 CPP_GFP (SEQ ID NO 1 and 101);

Figure 8: map of scAAV_SRSF1 89-120 CPP_GFP (SEQ ID NO 74 and 102);

Figure 9: (A) Western blots show depletion of SRSF1 and inhibition of the RAN translation of sense DPRs upon co-transfection with scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP, but not when CPPs transcription is driven the RNAPII promoter. SRSF1 and DPRs expression levels are quantified in triplicate biological experiments in panels B and C respectively. (D) MTT cell proliferation assays in biological triplicates showing that scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP alleviates the cytotoxicity mediated by the expression of DPRs, but not when CPPs transcription is driven the RNAPII promoter;

Figure 10: (A) Western blots show depletion of SRSF1 and inhibition of the RAN translation of antisense DPRs upon co-transfection with scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP, but not when CPPs transcription is driven the RNAPII promoter. SRSF1 and DPRs expression levels are quantified in triplicate biological experiments in panels B and C respectively. (D) MTT cell proliferation assays in biological triplicates showing that scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP alleviates the cytotoxicity mediated by the expression of DPRs, but not when CPPs transcription is driven the RNAPII promoter;

Figure 11 map of scAAV_SRSF1 -shRNA10_GFP (SEQ ID NO 66 and 103);

Figure 12 DPR quantification in mouse brains. C9ORF72-ALS/FTD (C9-Tg) mice were injected intrathecally (via cisterna magna) with 6×10^{10} vector genome (vg) of scAAV9_Ctrl-shRNA_GFP or 2 doses of scAAV9_SRSF1-shRNA10_GFP (4×10^{10} and 8×10^{10} vg) at post-natal day 1-2 (P1-2). Non-transgenic (NTg) mice are used as a control. Animals were sacrificed 1 month (A) or 3 months (B) post injection prior to MSD- ELISA quantification of poly-GP DPRs in the cerebellum/brainstem (mean \pm SEM; one-way ANOVA with Tukey's correction for multiple comparisons; NS: non-significant, **: p<0.01, ****: p<0.0001; n=4-6 mice/group). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and levels normalised to 100 % for the untreated C9-Tg mice;

Figure 13 Viral transduction in mouse brains. Immunohistochemical analysis of C9ORF72-ALS/FTD mice injected intrathecally (via cisterna magna) with 5×10^{10} vector genome (vg) of scAAV9_H1-SRSF1-CPP1_GFP or scAAV9_H1-CPP2_GFP at post-natal day 1-2 (P1-2). Animals were sacrificed one month post injection prior to anti-GFP immunofluorescence microscopy in the brain. Representative images are shown on the sections of midbrain. GFP co-expression is displayed in the green channel. DAPI (blue channel) and NeuN (red channel) stain nuclei and mature neurons respectively. Side panels: Enlarged immunofluorescence images showing transduction and scAAV9-mediated co-expression of GFP expression in both neuronal and microglial cells. Scale bars represent 500 μ m; and

Figure 14 Viral biodistribution and DPR quantification in mouse brains. C9ORF72-ALS/FTD (C9-Tg) mice were injected intrathecally (via cisterna magna) with 5×10^{10} vector genome (vg) of scAAV9_H1-SRSF1-CPP1_GFP or scAAV9_H1-CPP2_GFP at post-natal day 1-2 (P1-2). Non-transgenic (NTg) mice are used as a control. Animals were sacrificed one month post injection. (A) qPCR quantification of viral DNA extracted from the brain (cerebellum) and spinal cords (n=3), showing efficient transduction. (B) MSD- ELISA quantification of poly-GP DPRs in the cerebellum/brainstem (mean \pm SEM; one-way ANOVA with Tukey's correction for multiple comparisons; **: p<0.01, ****: p<0.0001; N=3 mice/group). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and levels normalised to 100 % for the untreated C9-Tg mice.

Materials and Methods

PART 1: SRSF1 depletion promotes the survival of sALS patient-derived motor neurons co-cultured with astrocytes

1/ Timeline for differentiation and co-culture of motor neurons and astrocytes derived from healthy control and sporadic ALS (sALS) patients:

Summary: Both iMotor Neurons (iMNs) and iAstrocytes are treated with either 5 MOI (Multiplicity of Infection) of lentivirus (LV) expressing a Ctrl-miRNA or 2 chained miRNAs directed against SRSF1 (constructs described in Hautbergue et al, Nature Communications 2017; 8:16063 and in our patent WO2017207979A1) at day 18 and 3 of the differentiation respectively, prior to establishing co-culture from day 20 (iMN) / 5 (iA). High content automated live imaging quantify iMN survival at day 22, 23, 24. scAAV9 does not efficiently transduce cells *in vitro*, in contrast to lentivirus which have been used here in this system.

Detailed protocol: Co-cultures of patient-derived astrocytes and motor neurons

Differentiation of iMotor Neurons (iMNs). Human patient and control-derived neurons (iNeurons) were differentiated from induced neural progenitor cells (iNPCs) using a modified version of protocol (Meyer K et al. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:829–832) as previously described (Hautbergue GM et al, Nature Communications 2017; 8:16063). In brief, 100,000 iNPCs were plated in a 6-well plate coated with fibronectin (Millipore) and expanded to 70-80% confluence. Once they reached this confluence, iNPC medium was replaced with neuron differentiation medium (DMEM/F-12 with glutamax supplemented with 1% N2, 2% B27 (Gibco) containing 2.5 µM of DAPT (Tocris) to determine differentiation towards neuronal lineage on day 1. On day 3, the neuron differentiation medium was supplemented with 1 µM retinoic acid (Sigma), 0.5 µM smoothened agonist (SAG) (Millipore) and 2.5 µM forskolin (Sigma) for 7 days until Day 10. This protocol leads to typical yields of 70% β-III tubulin (Tuj1) positive cells. To obtain iMotor Neurons (iMN), ~ 5,000 iNeurons per well were re-plated on 96-well plates coated with fibronectin and maintained in iNeuron differentiation medium (containing retinoic acid, SAG and forskolin) supplemented with BDNF, CNTF and GDNF (all at 20 ng/ml) for the last 14 days of differentiation.

Differentiation of iAstrocytes. Human patient-derived astrocytes (iAstrocytes) were differentiated from iNPCs as previously described (Meyer K et al. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:829–832; Hautbergue GM et al, Nature Communications 2017; 8:16063) and cultured in DMEM glutamax (Gibco) with 10% FBS (Sigma) and 0.02% N2 (Invitrogen) for 5 days. Cells were maintained in a 37°C incubator with 5% CO2.

Co-cultures of patient-derived iMNs and iAstrocytes. iAstrocytes were lifted at day 5 of differentiation and ~5,000 iAstrocytes were re-plated on iMNs at day 20 of differentiation. Co-cultured iMNs and iAstrocytes were maintained in neuron differentiation medium with BDNF,

GDNF and CTNF (all at 20 ng/ml) for 4 days. 12 h after the start of co-cultures (on day 21), 1 or 10 µM CPP was added to the medium and iMNs/ iAstrocytes were imaged for 72 h at days 22, 23, 24. For SRSF1 knockdown, iMNs and iAstrocytes were separately transduced 48h prior to co-culture with lentivirus (LV) expressing control or SRSF1-RNAi co-expressing GFP (Hautbergue GM et al, Nature Communications 2017; 8:16063) at a MOI of 5 at day 18 of iMN differentiation and at day 3 of iAstrocyte differentiation.

PART 2: scAAV9-driven expression of SRSF1-shRNA

Pre-clinical vector design: scAAV_SRSF1-shRNA_GFP

10

1/ SRSF1-shRNA cassette targeting mouse, rat, non-human primate and human SRSF1 Take region human SRSF1 448-750 (3' end of open reading frame) which is highly conserved with mouse SRSF1.

15 Human SRSF1 (NM_006924.4) SEQ ID NO 67

gctgatgttaccgagatggcactgggtcggtggagttgtacggaaagaagatatgacctatgcagttcgaaaactggataacac
taagtttagatctcatgaggagaaaactgcctacatccgggtaaagtgtatgggcccagaagtcaagttatgaaagatctcgat
ctcgaagccgtagtcgtacagaagccgtacgagaagcaacacgcaggagtcgcagttactccccaggagaaggcagagga
tcaccacgctattctcccgtagcagatctcgctctcgatcataaa

20

ttaaagtgtatgggcccagaa miRNA (SEQ ID NO 87) used to target human and mouse SRSF1 in the lentivirus construct (Hautbergue et al. Nature Communications 2017; 8:16063 and in our patent WO2017207979A1)

25 Design shRNA using the following website:

Block-iT RNAi Designer tool: <http://rnaidesigner.lifetechnologies.com/rnaiexpress/>

Table 1: SEQ ID NO 2-11

No.	Start (nt)	Target sequence (DNA)	Region	GC%	Rank (predicted efficacy 0-5)
30	2	GCTGATGTTACCGAGATGGC	52.39	3.5	
	3	GGAGTTGTACGGAAAGAAGA	42.86	4.5	
	4	GGAAAGAAGATATGACCTATG	38.1	3.5	
		not fully conserved human/mouse			
35	5	GAAAGAAGATATGACCTATGC	38.1	3.5	
		not fully conserved human/mouse			
	6	GCCTACATCCGGTAAAGTT	47.62	3.5	

7	139	GGGCCAGAAGTCCAAGTTAT	52.39	4.5
8	140	GGCCCAGAAGTCCAAGTTATG	52.39	3.5
5	9	GCCCAGAAGTCCAAGTTATGG	52.39	4.0
10	160	GGAAGATCTCGATCTCGAACGC	52.39	4.5
11	245	GCAGAGGATCACCAACGCTATT	52.39	5.0

10

Table 2: Use siSPOTR (Boudreau RL et al. Nucleic Acids Res. 2013;41(1):e9) to predict the off target of the common human/mouse sequences targeting SRSF1 and predicted most efficient

15	shRNA sequence	antisense/ mature shRNA sequence	POTS	POTS	Seed
77	gccaucucgguaaacaucaucagc	355.351 (mouse)	463.716 (human)	CCATCTC	
78	ucuuucuuuccguacaaacucc	501.411 (mouse)	588.488 (human)	CTTCTTT	
79	cauaggucauauaucuuuuucc	96.137 (mouse)	149.126 (human)	ATAGGTC	
20	80	gcauaggucauauaucuuuuuc	140.373 (mouse)	167.649 (human)	CATAGGT
81	aacuuuaaccggauaguaggc	390.256 (mouse)	526.984 (human)	ACTTTAA	
82	auaacuuggacuucugggccc	221.397 (mouse)	339.052 (human)	TAACTTG	
83	cauaacuuggacuucugggcc	237.138 (mouse)	351.458 (human)	ATAACTT	
84	ccauaacuuggacuucugggc	159.58 (mouse)	215.339 (human)	CATAACT	
25	85	gcuucgagagaucgagacuucc	41.326 (mouse)	41.3938 (human)	CTTCGAG
86	aauagcguggugauccucugc	21.324 (mouse)	22.5396 (human)	ATAGCGT	

SRSF1 target shRNA6 sequence: 5'- GGGCCCAGAAGTCCAAGTTAT -3' (SEQ ID NO 7)

Antisense/mature shRNA6 sequence: 5'- AUUACUUGGACUUUCUGGGCCC -3' (SEQ ID NO

30 82)

SRSF1 target shRNA9 sequence: 5'- GGAAGATCTCGATCTCGAACGC -3' (SEQ ID NO 10)

Antisense/mature shRNA9 sequence: 5'- GCUUCGAGAUCGAGAUCUUCC -3'(SEQ ID NO

85)

35

SRSF1 target shRNA10 sequence: 5'- GCAGAGGATCACCAACGCTATT -3' (SEQ ID NO 11)

Antisense/mature shRNA10 sequence: 5'- AAUAGCGUGGUGAUCCUCUGC -3' (SEQ ID NO

86)

2/ Alignment human (NM_006924.4; SEQ ID NO 104) and mouse (NM_173374.4; SEQ ID NO 105) SRSF1

Sequences corresponding to the shRNAs 7, 10 and 11 (predicted the most efficient with the less predicted off-target effects) are highlighted on the aligned human and mouse SRSF1

5 open reading frames.

		hSRF1	ATGTCGGGAGGTGGTGTGATTCTGGCCCGCAGGAACAACGATTGCCGCATCTACGTG	60
		mSRF1	ATGTCGGGAGGTGGTGTGATCGTGGCCCGGGGGAAACAACGACTGCCGCATCTACGTG	60
10	hSRF1	GGTAACCTACCTCCAGACATCGAACCAAGGACATTGAGGACGTGTTCTACAAATAACGGC	120	
	mSRF1	GGTAACCTACCTCCGGATATCGAACCAAGGACATCGAGGACGTGTTTACAAATAACGGC	120	
15	hSRF1	GCTATCCCGCACATCGACCTCAAGAATGCCCGGGGGACCGCCCTCGCCTCGTTGAG	180	
	mSRF1	GCCATCCCGCACATCGACCTGAAGAACCGCCGCCGGGGACCGCCCTCGCCTCGTTGAG	180	
20	hSRF1	TTCGAGGACCCCGCGAGACCGGAAAGACCGGGTGTATGGTCGCGACGGCTATGATTACGAT	240	
	mSRF1	TTCGAGGACCCCGCGAGACCGGAAAGATGCGGTGTACGGTCGCGACGGCTACGACTACGAC	240	
25	hSRF1	GGGTACCGTCTCGGGTGGAGTTCTCGAACCGGCCGTGGAACAGGCCGAGGCCGGCG	300	
	mSRF1	GGCTACCGGCTCGGGTAGAGTTCCCCGAAGCGGCCGCCGGACCGCCGAGGCCGGCG	300	
30	hSRF1	GGGGGTGGAGGTGGCGAGCTCCCGAGGTGCGTATGGCCCCCATCCAGGCGGTCTGAA	360	
	mSRF1	GGGGGTGGAGGCGGCGCCCGAGAGGCCGCTATGGCCCGCCGTCCAGGCCGGTCCGAG	360	
35	hSRF1	AACAGAGTGGTTGTCTGGACTGCCTCCAAGTGGAAAGTGGCAGGATTAAAGGATCAC	420	
	mSRF1	AACAGAGTGGTTGTCTGGACTGCCTCCGAGTGGAAAGCTGGCAGGACTAAAGGATCAC	420	
40	hSRF1	ATGCGTGAAGCAGGTGATGTATGCTATGCTGATGTTACCGAGATGGCACTGGTCTCGTG	480	
	mSRF1	ATGCGTGAAGCAGGTGATGTATGTTACGCTGATGTTACCGAGATGGCACTGGTCTCGTG	480	
45	hSRF1	GAGTTTGTACGGAAAGAAGATATGACCTATGCAGTCGAAACTGGATAACACTAACGTT	540	
	mSRF1	GAGTTTGTACGGAAAGAAGATATGACGTATGCAGTCGAAACTGGATAACACTAACGTT	540	
50	hSRF1	AGATCTCATGAGGGAGAAACTGCCTACATCCGGTTAAAGTTGATGGCCCCAGAACGTC	600	
	mSRF1	AGATCTCACGAGGGAGAAACTGCCTACATCCGGTTAAAGTTGATGGCCCCAGAACGTC	600	

hSRSF1	AGTTAT GGGAGTCAAGTGTATTTTGAGTC CGTAGTCGTAGCAGAAGCCGTAGCAGAAC	660
mSRSF1	AGTTAT GGGAGTCAAGTGTATTTTGAGTC CGTAGTCGTAGCAGAAGCCGTAGCAGAAC	660
5		
hSRSF1	AGCAGGAGTCGAGTTACTCCCCAAGGAGAAC AGACAGGATCACCAAGCTAT CTCCCCGT	720
mSRSF1	AGCAGGAGTCGAGTTACTCCCCAAGGAGAAC AGACAGGATCACCAAGCTAT CTCCCCGT	720
10		
hSRSF1	CATAGCAGATCTCGCTCTCGTACATAA	747
mSRSF1	CATAGCAGATCTCGCTCTCGTACATAA	747

15	shRNA	antisense/ mature shRNA sequence	POTS	POTS	Seed sequence
	82	auaacuuggacuuucuggggccc	221.397 (mouse)	339.052 (human)	TAAC TTG
	86	gcuucgagauccgagaucuuucc	41.326 (mouse)	41.3938 (human)	CTTC GAG
	86	aauagcguggugauccucugc	21.324 (mouse)	22.5396 (human)	ATAG CGT

20

3/ Cloning of SRSF1-targeting shRNAs into the scAAV_GFP vector

We then designed and custom synthesised the following oligonucleotides for cloning shRNAs 7, 10 and 11 into our scAAV_H1promoter_GFP vector (SEQ ID NO 68-73):

25 Cut BamHi / cut HindIII Red sequences corresponds to SRSF1 targeted region Blue sequences correspond to antisense/mature shRNA Black sequence corresponds to hairpin loop

SRSF1 shRNA_6_fwd (SEQ ID NO 68)

GATCC GGGCCCAGAAGTCCAAGTTAT TTCAAGAGA ATAACTTGGACTTCTGGGCC C TTTTTT GGA A

30

SRSF1 shRNA_6_rev (SEQ ID NO 69):

AGCTT TCC AAAAAA G GGGCCCAGAAGTCCAAGTTAT TCTCTTGAA ATAACTTGGACTTCTGGGCC G

35 SRSF1 shRNA_9_fwd (SEQ ID NO 70):

GATCC GGAAGATCTCGATCTCGAACG TTCAAGAGA GCTTCGAGATCGAGATCTTCC C TTTTTT GGA A

SRSF1 shRNA_9_rev (SEQ ID NO 71):

AGCTT TCC AAAAAA G GGAAGATCTCGATCTCGAACG TCTCTTGAA GCTTCGAGATCGAGATCTTCC G

40

SRSF1 shRNA_10_fwd (SEQ ID NO 72):

GATCC GCAGAGGATCACCACGCTATT TTCAAGAGA AATAGCGTGGTGATCCTCTGC C TTTTTT GGA A

SRSF1 shRNA_10_rev (SEQ ID NO 73):

AGCTT TCC AAAAAA G GCAGAGGATCACCAACGCTATT TCTCTGAA AATAGCGTGGTATCCTCTGC G

5/ Full sequence of the pre-clinical scAAV vector co-expressing the SRSF1-shRNA

5 cassette (under constitutively-expressed RNAPIII H1 promoter) and eGFP (under a weak RNAPII eF-1alpha core promoter to avoid potential GFP-induced toxicity)

scAAV_SRSF1-shRNA10_GFP circular sequence (5,648 bp) SEQ ID NO 66

5'...

10 GCTCTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGTCGGCTCGGGAGCGGTATCAGCTCAC
TCAAAGGCCGTAATACGGTTATCCACAGAACATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCA
GCAAAGGCCAGGAACCGTAAAAGGCCGCGTGGCTGGCTTTCCATAGGCTCCGCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCAAACCCGACAGGACTATAAGATAACCAGGCCTTCC
CCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCCGCTACCGATACTGTCCGCCCTTCTCC
15 CTTCGGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTCGGTAGGTCTCGCTCC
AAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCGACCGCTGCCCTATCCGTAACACTCGTCTTG
GTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTCGGAAAAGAGTTGGTAGCTCTGATCCGCAAACAAACCCG
20 CTGGTAGCGGTGGTTTTGTTGCAAGCAGCAGATTACGCCAGAAAAAGGATCTCAAGAACATCCT
TGATCTTTCTACGGGCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTGGTCTGAGATTAT
CAAAAAGGATCTCACCTAGATCCTTAAATTAAAATGAAGTTAAATCAATCTAAAGTATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAACTCAGTGAGGCACCTATCTCAGCGATCTGCTATTCGTTCTACCAT
AGTTGCCTGACTCCCCGTCGTAGATAACTACGATAACGGAGGGCTTACCATCTGCCAGTGCTGCAA
25 TGATACCGCGAGACCCACGCTACCGGCCAGATTACGCAATAAACCAAGCCAGCCGGAAAGGGCCGA
GCGCAGAAGTGGCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAGTAAG
TAGTCGCCAGTTAATAGTTGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTT
TGGTATGGCTTCTAGCTCCGGTCCACGATCGTGTAGAAGTAAGTGGCCGAGTGTATCACTCATGGTT
AGCGGTTAGCTCCTCGGTCCCGATCGTGTAGAAGTAAGTGGCCGAGTGTATCACTCATGGTT
30 GGCAGCACTGCATAATTCTCTTACTGTCTGCCATCCGTAAAGATGCTTCTGTACTGGTAGTACTCAACC
AAAGTCATTCTGAGAATAGTGTATGCCGACCGAGTTGCTCTGCCGGCGTCAATACGGATAATACCGC
GCCACATAGCAGAACTTAAAGTGCTCATCATTGGAAAACGTTCTCGGGCGAAAACCTCAAGGATCTT
ACCGCTGTTGAGATCCAGTTGATGTAACCCACTCGCACCCACTGATCTCAGCATCTTACTTCACC
AGCGTTCTGGGTGAGCAAAACAGGAAGGCAGGAAATGCCGAAAAAGGGATAAGGGCGACACGGAAATG
35 TTGAATACTCATACTCTCCTTTCAATATTATTGAAGCATTATCAGGGTATTGCTCATGAGCGGATACA
TATTGAATGTATTAGAAAATAACAAATAGGGTTCCGCGCACATTCCCCGAAAAGTGCACCTGACGT
CTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCAGGAGGCCCTTCGCTCGCGCT
TTCGGTGATGACGGTGGAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTGTAAGCGGA
TGCCGGGAGCAGACAAGCCGTCAGGGCGCGTCAGGGTGTGGCGGGTGTGGGGCTGGCTTAAC
40 TGCAGGATCAGAGCAGATTGTACTGAGAGTGCACCATAGTGTGGCGGGTGTGGGGCTGGCTTAAC
CGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAGTGTGGCGGGTGTGGGGCTGGCTTAAC
AAATACCGCAGGCGATTCCAACATCCAATAATCATAACAGGCAAGGCAAAGAATTAGCAAAATTAAAGCAA
TAAAGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAACATTATGACCCGTAAACTTTGCGGGAGAA
GCCTTATTCAACGCAAGGATAAAAATTAGAACCCCTCATATATTAAATGCAATGCCTGAGTAATGTGT

AGGTAAAGATTCAAACGGGTGAGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCT
AGCTGATAAATTATGCCGGAGAGGGTAGCTATTTGAGAGGTCTCTACAAAGGCTATCAGGTATTGCCT
GAGAGTCTGGAGCAAACAAGAGAACGCCCCGGGCAAGGCCGGCGTGGCGACCTTGGTCGCCGGCTCAGTG
5 AGCGAGCGAGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGTTCTCAGATCGATCTCTCCCCA
GCATGCAGGCCTCTGCAGTCACGGGCCGGCATCGTTACTCCCCAGCATGCCTGCTATTCTCTTCCC
AATCCTCCCCCTGCTGCTGCCCAACCCCCACCCCCAGAACATAGAACGACACCTACTCAGACAATGCGATG
CAATTCTCATTTATTAGGAAGGACAGTGGAGTGGCACCTCCAGGGTAAGGAAGGCACGGGAG
GGGCAAACACAGATGGCTGGCAACTAGAACGGCACAGTCGAGGCTGATCAGCGAGCTCTAGGAATTACT
10 TGACAGCTCGTCCATGCCGAGAGTGTACCGCGGGCGGTACGAACTCCAGCAGGACCATGTGATCGCG
CTTCTCGTGGGTCTTGCTCAGGGCGACTGGTGCTCAGGTAGTGGTAGTGGTCGGCGAGCTGCACGCTGCCG
CCGTCGCCGATGGGGTGTCTGCTGGTAGTGGTAGTGGCGAGCTGCACGCTGCCGCTCGATGTTGGC
GGATCTGAAGTTCACCTGATGCCGTTCTCTGCTTGTGCCATGATAGAACGTTGGCTGTTAGTT
GTACTCCAGCTTGTGCCCAAGGATGTTGCCGTCCCTGAAGTCGATGCCCTCAGCTCGATGCCGTTCA
15 CCAGGGTGTGCCCTCGAACCTCACCTCGCGGGCTTGAGTTGCCGTGTCCTGAAGAACGATGGT
CGCTCCTGGACGTAGCCTCGGGCATGGCGACTTGAAGAACGTCGTCGCTCATGTTGGCGGGTAGC
GGCTGAAGCACTGCACGCCGTAGGTAGGGTGGCACGAGGGTGGCCAGGGCACGGCAGCTGCCGG
TGGTAGCAGATGAACCTCAGGGTCAGCTGCCGTAGGCTAGGCTGCGCAACTGTTGGGAAGGGCGATCG
CTTGTGGCGTTACGTCGCCAGCTGACCAGGATGGCACCAACCCGGTAACAGCTCTGCC
20 TTGCTCACCATGCCGTGTTCTGGCGCAAACCCGGTGCAGAAAAGAACGTTCACGGCGACTACTGACTTA
TATACGGTTCTCCCCACCCCTGGAAAAAGGCGGAGCCAGTACACGACATCACTTCCCAGTTACCCG
CGCCACCTCTCTAGGCACCGGATCAATTGCCGACCCCTCCCCCAACTCTCGGGGACTGTTGGCGATGT
GGCTCTGCCACTGAATCTCTCGAGCCTCTAGATACCACAGGCTGCCACTGTTGGGAAGGGCGATCG
GTGCGGGCCTCTCGTATTACGCCAGCTGGCAAAGGGGATGTGCTGCAAGGGGATTAAGTTGGTAA
25 CGCCAGGGTTTCCAGTCAGCACGTTGAAAACGACGCCAGTGAATTATTCATATTGCATGTCGCTATGTT
TCTGGAAATCACCATAAACGTGAAATGTTGGATTTGGAAATCTTATAAGTCTGTATGAGACCACCTGG
ATCCGCAGAGGATCACCAAGCTATTTCAGAGAAATAGCGTGGTATCCTCTGCCCTTTGGAAAGCTTG
GCGTAATCATGGTCTAGCTGTTCTGTGAAATTGTTATCCGCTACAATTCCACACAACATACGAGCCG
GAAGCATAAAAGTGTATCTAGAGCGGTACACCGCTGAATTGAAATTCAAGTCCACGCCGTAATTCCACTCC
30 CTCTGCCGCTCGCTCGACTGAGGCCGGGACCAAAGGTGCCGACGCCGGCTTGCC
CGGCCCTAGTGAGCGAGCGAGCGCAGGGCGATGAAAGGTCAATCGAAAAGTACGATGTCATCATATGT
ACCCCGGTTGATAATCAGAAAAGCCCCAAAACAGGAAGATTGTATAAGCAAATATTAAATTGTAAGCGTT
ATATTGTTAAAATCGCTTAAATTGTTAAATCAGCTCTTAAACCAATAGGCCGAAATCGGCAA
ATCCCTTATAAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTGGAACAAGAGTCCACTAT
35 TAAAGAACGTGGACTCCAAACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGCCCAACTACGTGAACCA
TCACCCCTAACGTTGGGAAAGCCGGCAACGTGGCAGAGAAAGGAAGGGAAAGAACGCAAGGGAGCCCG
ATTAGAGCTTGACGGGAAAGCCGGCAACGTGGCAGAGAAAGGAAGGGAAAGAACGCAAGGGAGCGGG
CGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCCGTAACCACCCGCCGCTTAATGCCCG
CTACAGGGCGCTACTATGGTTGCTTGACGAGCACGTATAACGTGTTCTCGTTAGAATCAGAGCGGG
40 AGCTAACACAGGAGGCCGATTAAGGGATTAGACAGGAACGGTACGCCAGAACCTGAGAACGTT
AATCAGTGAGGCCACCGAGTAAAGAGTCTGTCATCACGCAAATTACCGTTGCGCAATACCTTTGATT
AGTAATAACATCACTGCCTGAGTAGAACGAACTCAAACATCGGCCATTGCCATTAGGCTGCCACTGTTGGGAAGGGCGATCGGTG
CGGGCCTTCGCTATTACGCCAGCTGCATTAATGAATCGGCCAACGCCGGGGAGAGGCAGGTTGCGTAT
45 TGGC... 3'

6/ Functionality of scAAV9-driven expression of SRSF1-shRNA10 in mouse brains.

C9ORF72-ALS/FTD mice were injected via cisterna magna at post-natal day 1 (P1) with either
5 8×10^{10} scAAV9_Ctrl-shRNA_GFP vector genomes (vg) or 6×10^{10} scAAV9_SRSF1-
shRNA10_GFP vg. Animals were sacrificed 1 month and 3 months post injections. Western
blots show that the scAAV9_SRSF1-shRNA10_GFP virus leads to specific depletion of
SRSF1 in C9ORF72-ALS/FTD mice as well as in wild type C57Bl6 mice (not shown) while
the Ctrl-shRNA has no effect. GAPDH is used as a loading control.

10

4/ ITR sequences

SEQ ID NO 64:

ITR1:

5'-

ccactccctctcgcgctcgctcactgaggccgcgggcaaagccggcgacccgggtcgccgcctcagtgagcgagcg

15

cgcagagagggagtggcaactccatcactaggggtct -3'

SEQ ID NO 88:

20

ITR2:

5'-

ccactccctctcgcgctcgctcactgaggccgggaccaaaggcgccgcggcttgcggccgcctcagtgagcgagcg
cccg -3'

Sequence of Cell Permeable Peptide

25

SRSF1 132-144 CPP nucleotide sequence:

5'-

GGCAGCTGGCAGGATCTGAAAGATCATATGCGCGAAGCCGGCGTGGAAACCGATTCCAACCCGCTGC

TGGGCCTCGATAGCACCGCGGATATGGTCGCAAAAAGCGCAGACAGCGCCGGAGG

-3' (SEQ ID NO 89)

30

SRSF1 132-144 CPP sequence which corresponds to SRSF1 amino acids 132-144, a V5 tag
and the protein transduction domain TAT amino acids 47-57: Nt-
GSWQDLKDHMREAGGGKIPNPLLGLDSTGGYGRKKRRQRRR – Ct (SEQ ID NO 90)

35

5/ Sequence of the scAAV_SRSF1 89-120 CPP_GFP circular sequence (5,692 bp) (SEQ ID

NO 74)

5'...

GCTCTCCGCTTCCTCGCTCACTGACTCGCTGCGCTGGTCGGCTGGCGAGCGGTATCAGCTCAC

TCAAAGGGGTAATACGGTTATCCACAGAACATCAGGGATAACGCAGGAAGAACATGTGAGCAAAGGCCA

GCAAAGGCCAGGAACCGTAAAAGGCCGCGTGGCTGGCTTCCATAGGCTCCGCCCCCTGACGAG

40

CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCAAACCCGACAGGACTATAAGATAACCAGGCCTTCC
CCCTGGAAGCTCCCTCGTGCCTCCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCCCTTCTCC

CTCGGGAAGCGTGGCGTTCTCATAGCTACGCTGTAGGTATCTCAGTTGGTAGGTCGTCGCTCC
 AAGCTGGGCTGTGCACGAACCCCCCGTTAGCCGACCGCTGCCCTATCCGTAACTATCGTCTGA
 GTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGCGAGGT
 ATGTAGGGCGGTGCTACAGAGTTCTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTA
 5 TCTGCGCTCTGCTGAAGCCAGTTACCTCGGAAAAAGAGTTGGTAGCTCTGATCCGCAAACAAACCACCG
 CTGGTAGCGGTGGTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
 TGATCTTTCTACGGGTCTGACGCTCAGTGGAACAAAACACGTTAAGGGATTGGTATGAGATTAT
 CAAAAAGGATCTCACCTAGATCCTTTAAATTAAAATGAAGTTAAATCAATCTAAAGTATATGAGTAAA
 CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTCGTTATCCAT
 10 AGTTGCCTGACTCCCCGTCGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA
 TGATACCGCGAGACCCACGCTACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAGGGCCGA
 GCGCAGAAGTGGTCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAGCTAGAGTAAG
 TAGTTGCCAGTTAATAGTTGCGCAACGTTGCTTGCATTGCTACAGGCATCGTGGTGTACGCTCGTCTT
 TGGTATGGCTTCATTAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGCAAAAAA
 15 AGCGGTTAGCTCCTCGGTCCCGATCGTGTAGAAGTAAGTTGGCCAGTGTATCACTCATGGTTAT
 GGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTGTGACTGGTGAGTACTCAACC
 AAGTCATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTGCCGGCGTAATACGGGATAATACCGC
 GCCACATAGCAGAACTTTAAAGTGTCTCATGGAAAACGTTCTCGGGCGAAAACCTCAAGGATCTT
 ACCGCTGTTGAGATCCAGTCGATGTAACCCACTCGTGCACCCACTGATCTCAGCATTTACTTCACC
 20 AGCGTTCTGGGTAGCAAAACAGGAAGGCAAAATGCCGAAAAAAGGGATAAAGGGCACACGGAAATG
 TTGAATACTCATACTCTCCTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCATGAGCGGATACA
 TATTGAATGTATTAGAAAATAACAAATAGGGTTCCGCGCACATTCCCCGAAAAGTGCACCTGACGT
 CTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGCGTACCGAGGCCCTTCGTCGCGGT
 TTCGGTGATGACGGTAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTGCTGTAAGCGGA
 25 TGCCGGGAGCAGACAAGCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGCTGGCTTAACCA
 TGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATCGGTGTGAAATACCGCACAGATCGTAAGGA
 GAAAATACCGCATCAGGCATTCCAACATCCAATAATCATAAGGCAAGGAAAGAATTAGCAAATAAG
 CAATAAAGCCTCAGAGCATAAAGCTAAATCGTTGACCAAAACATTATGACCTGTAATACTTGCAGGG
 GAAGCCTTATTCAACGCAAGGATAAAAATTAGAACCCCTCATATATTAAATGCAATGCTGAGTAATG
 30 TGTAGGTAAAGATTCAAACGGGTAGAAAGGCCGGAGACAGTCACCATCAATATGATATTCAACCGT
 TCTAGCTGATAAAATTCTGCGGAGAGGGTAGCTATTGAGAGGTCTCTACAAAGGCTATCAGGTATTG
 CCTGAGAGTCTGGAGCAAACAAGAGAATGCCGGGGGGGGGGGGGGGGCACTCCCTCTGCGCG
 CTCGCTCGCTCACTGAGGCCGGGCAAAGCCGGCGTCGGCGACCTTGGTCGCCGGCCTCAG
 TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCAACTCCATCACTAGGGTTCTCAGATCGATCTCCC
 35 CAGCATGCGTTTACCTCCCCAGCATGCCGTATTCTCTTCCAACTCCTCCCCCTGCTGTCCTGCCAAC
 CCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTCTCTATTAGGAAAGGAC
 AGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGAGGGGCAAACAAACAGATGGCTGGCAACTA
 GAAGGCACAGTCGAGGCTGATCAGCGAGCTAGGAATTACTGTACAGCTGTCATGCCGAGAGTGA
 TCCCGCGCGGGTACGAACCTCCAGCAGGACCATGTGATCGCCTCTCGTTGGGTCTTGCTCAGGGC
 40 GGACTGGGTGCTCAGGTAGTGGTGTGGCAGCAGCACGGGCCGTCGCCGATGGGGTGTCTGCTG
 GTAGTGGTGGCGAGCTGCACGCTGCCCTCGATGTTGGCGGATCTGAAGTTCACCTGATGCCGT
 TCTTCTGCTGTCGGCCATGATATAGACGTTGTGGCTGTTAGTTACTCCAGCTGTGCCCAAGGATGT
 TGCCGTCTCCTGAGTCGATGCCCTCAGCTCGATGCCGTTACCAAGGGTGTGCCCTCGAACCTCACC
 TCGCGCGGGTCTGTAGTTGCCGTCGTCCTGAAGAAGATGGTGCCTGGACGTAGCCTCGGGCAT
 45 GGCAGACTGAGAAGTCGTCGCTCATGTTGGTGGGGTAGCGGCTGAAGCAGTCACGCCGTAGGTC

AGGGTGGTCACGAGGGTGGGCCAGGGCACGGCAGCTGCCGTGGTCAGATGAACCTCAGGGTCAGC
 TTGCCGTAGGTGGCATGCCCTGCCCGAACGCTGAACCTGTGGCGTTACGTCGCCGTCCA
 GCTGACCAGGATGGGCACCACCCGGTAACAGCTCCTGCCCTGCTACCATGCCGTGTTCTGGCG
 GCAAACCCGGTGCAAAAAGAACGTTACGGCGACTACTGCACTTATATACGGTTCTCCCCCACCCTCGGG
 5 AAAAAGGCGGAGCCAGTACACGACATCACTTCCCAGTTACCCCGCCACCTCTCTAGGCACCGGATC
 AATTGCCGACCCCTCCCCCAACTCTCGGGACTGTGGCGATGTGCGCTTGCCACTGAATCTCTCG
 AGCCTCTAGATACCACAGGCTGCGCACTGTTGGAAAGGGCGATCGGTGCGCTTGCCCTTCGCTATTACGCC
 AGCTGGCGAAAGGGGATGTGCTGCAAGGCGATTAAGTTGGTAACGCCAGGGTTCCAGTCACGACG
 TTGTAACGACGGCAGTGAATTCATATTGCATGTCGCTATGTGTTCTGGAAATCACCATAAACGTGAAA
 10 TGCTTTGGATTGGAAATCTTATAAGTTCTGTATGAGACCACCTGGATCCGAAGGCCACCATGCCGCGCAGC
 GGCGCGGCACCGGCCGCGTGGGGCGGTGGAGGTGGCGGAGCCCCGAGAGGGCGCTATGGACC
 GCCCAGCCGCCGGAGCGAAGGGCGTGGAAACCGATTCCCAACCCGCTGCTGGGCCTCGATAGCACCGG
 CGGATATGGTCGCAAAAGCGCAGACAGCGCCGGAGGTAACTTTAAGCTGGCGTAATCATGGTCATAG
 CTGTTCTGTGTGAAATTGTTATCCGCTACAATTCCACACAATACGAGGCCAGCATAAAAGTGTATCT
 15 AGAGCGGTACCACCGGTGAATTGAATTCAAGATCCACCGCGTGAATTCCACTCCCTCTGCCGCGCTCGCTCG
 CTCACTGAGGCCGGCGACCAAAGGTCGCCGACGCCCGGGCTTGCCCGGGCGCTCAGTGAGCGAG
 CGAGCGCGCAGGGCGATGAACGGAATCGTAAACTAGCATGTCAATCATATGTACCCCGTTGATAATCA
 GAAAAGCCCCAAAACAGGAAGGATTGTATAAGCAAATATTAAATTGTAAGCGTTAATATTGTTAAAATTG
 CGTTAAATTGTTAAATCAGCTCATTTTAACCAATAGGCCAAATGGCAAAATCCCTATAAAATCAAAA
 20 GAATAGACCGAGATAGGGTGGTGGTCCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCC
 AACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCTAATCAAGTTT
 TTGGGGTCAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCCCCGATTAGAGCTTGACGGG
 GAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGGGAGCGAAAGGAGCGGGCGTAGGGCGCTGGCAA
 GTGTAGCGGTACGCTGCGCGTAACCACACCACCCGCCGCGCTTAATGCGCCGCTACAGGGCGTACTA
 25 TGGTTGCTTGACGAGCACGTATAACGTGCTTCCCTCGTTAGAATCAGAGCGGGAGCTAACACAGGAGGCC
 ATAAAGGGATTTAGACAGGAACGGTACGCCAGAACCTGAGAAGTGTGTTTATAATCAGTGAGGCCACCG
 AGTAAAGAGTCTGCCATACGCAAATTAAACGTTGCGCAACTTCTTGATTAGTAATAACATCACTTG
 CCTGAGTAGAGAACTCAAACATCGGCCCTGCTGTAATATCCAGAACAAATTACCGCCAGCCATTGCAA
 CGGAATGCCATTGCCATTCAAGCTGCGCAACTGTTGGAAAGGGCGATCGGTGCGGGCTTCGCTATT
 30 ACGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGGCGGTTGCGTATTGGC -3'

6/ 5' Sequence of the scAAV_SRSF1 132-144 CPP_GFP circular sequence (5,651 bp) (SEQ ID NO 1)

5'...

35 GCTCTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGTCGGCTGGCTGCCGAGCGGTATCAGCTCAC
 TCAAAGGCGGTAAACGGTTATCCACAGAACGTTGGGATAACGCAAGGAAAGAACATGTGAGCAAAAGGCCA
 GCAAAAGGCCAGGAACCGTAAAAGGCCCGTTGCTGGCTTTCCATAGGCTCCGCCCTGACGAG
 CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGAACATCAGGCCGTTCC
 CCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCCCTACCGGATACCTGTCCGCCCTTCTCC
 40 CTTCGGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGTGTAGGTCGTTGCTCC
 AAGCTGGGCTGTGCAAGAACCCCCCGTTCAAGCCGACCGCTGCCCTATCCGTAACACTCGTCTTGA
 GTCCAACCCGGTAAGAACACGACTTATGCCACTGGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
 ATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGGCTAACTACGGCTACACTAGAACGGACAGTATTGGTA
 TCTGCGCTGCTGAAGCCAGTTACCTCGGAAAAGAGTGGTAGCTTGTATCCGGCAAACAAACACCACCG

CTGGTAGCGGTGGTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTAAGAAGATCCTT
TGATCTTTCTACGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTGGTCATGAGATTAT
CAAAAAGGATCTCACCTAGATCCTTTAAATTAAAATGAAGTTAAATCAATCTAAAGTATATGAGTAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTGTTCATCCAT
5 AGTTGCCTGACTCCCCGTCGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA
TGATACCGCGAGACCCACGCTACCAGCTCCAGTTATCAGCAATAAACAGCCAGCCGGAAAGGCCGA
GCGCAGAAGTGGTCTGCAACTTATCCGCCTCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTGCCAGTTAATAGTTGCGAACGTTGCTACAGGCATCGTGGTGTACGCTCGTGT
TGGTATGGCTCATTAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGCAAAAA
10 AGCGGTTAGCTCCTCGGTCCGATCGTTGTCAGAAGTAAGTTGCCGCAGTGTATCACTCATGGTTAT
GGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTCTGTGACTGGTAGTACTCAACC
AAGTCATTGAGAATAGTGTATGCGGCACCGAGTTGCTCTGCCCAGCTCAATACGGGATAATACCGC
GCCACATAGCAGAACTTAAAGTGCATCATTGGAAAACGTTCTCGGGCGAAAACCTCAAGGATCTT
ACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACGATCTCAGCATCTTACTTCA
15 AGCGTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAAAAAGGGAATAAGGGCGACACGGAAATG
TTGAATACTCATACTCTCCTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCATGAGCGGATACA
TATTGAAATGTATTTAGAAAATAACAAATAGGGTTCCGCGCACATTCCCCGAAAAGTGCACCTGACGT
CTAAGAAACCATTATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCTTGTCTCGCGT
TTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTGTCTGTAAGCGGA
20 TGCCGGGAGCAGACAAGCCCCTCAGGGCGCGTCAGCGGGTGTGGCGGGGTGCGGGGCTGGCTTAACTA
TGC GG CATCAGAGCAGATTGACTGAGAGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGA
GAAAATACCGCATCAGGCATTCCAACATCCAATAACATACAGGCAAGGCAAAGAATTAGCAAATTAAAG
CAATAAAGCCTCAGAGCATAAAGCTAAATCGTTGACCAAAACATTATGACCTGTAATACTTGCGGGA
GAAGCCTTATTCACGCAAGGATAAAAATTAGAACCCCTCATATATTAAATGCAATGCTGAGTAATG
25 TGTAGGTAAAGATTCAAACGGGTGAGAAAGGCCGGAGACAGTCAAATCACCATAATGATATTCAACCGT
TCTAGCTGATAAATTGATGCCGGAGAGGGTAGCTATTGAGAGGTCTCTACAAAGGCTATCAGGTATTG
CCTGAGAGTCTGGAGCAAACAAGAGAATGCCGGGGGGGGGGGGGGGGGGCCACTCCCTCTGCGCG
CTCGCTCGCTCACTGAGGCCGGGCAAAGCCGGCGTCGGCGACCTTGGTCGCCGGCCTCAG
TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCAACTCCATCACTAGGGTTCTCAGATCGATCTCCC
30 CAGCATGCGTTTACCTCCCCAGCATGCCGTCTATTCTCTTCAATCCTCCCCCTGCTGTCCTGCCAAC
CCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTCTCATTATTAGGAAAGGAC
AGTGGGAGTGGCACCTCCAGGGTCAAGGAAGGCACGGGGAGGGCAAACACAGATGGCTGGCAACTA
GAAGGCACAGTCGAGGCTGATCGCGAGCTCTAGGAATTACTTGTACAGCTCGCCATGCCGAGAGTGA
TCCCGGGCGGGTCACTGAGGACCATGTGATCGCGCTTCTGTTGGGTCTTGCTCAGGGC
35 GGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCGTCGCCGATGGGGTGTCTGCTG
GTAGTGGTCGGCGAGCTGCACGCTGCCGCTCGATGTTGGCGGATCTGAAGTTCACCTGATGCCGT
TCTTCTGCTGTCGGCATGATATAGACGTTGTCGGTGTAGTTGTACTCCAGCTGTGCCCGAGGATGT
TGCGTCCTCTGAAGTCGATGCCCTCAGCTCGATGCCGTTCACTGGGTGCGCCTGGACGTAGCCTCGGGCAT
40 GGC GG ACTTGAAGAAGTCGTGCTGTTCATGTTGGTCGGGTAGCGGCTGAAGCAGTCACGCCGTAGGTC
AGGGTGGTACGAGGGTGGCCAGGGCACGGGCAGCTGCCGGTGGTCAGATGAACCTCAGGGTCAGC
TTGCCGTAGGTGGCATGCCCTGCCCTGCCGGACACGCTGAACCTGTTGCCGTTACGTCGCCGTCCA
GCTCGACCAAGGATGGCACCACCCGGTGAACAGCTCTGCCCTGCTCACCATGCCGTGTTCTGGCG
GCAAACCCGTTGCAAAAAGAACGTTCACGGCAGTACTGCACTTATACGGTTCTCCCCACCCCTCGGG
45 AAAAAGGCGGAGCCAGTACACGACATCATTCCAGTTACCCGCCACCTCTAGGCACCGGATC

AATTGCCGACCCCTCCCCCAACTTCTGGGGACTGTGGGCATGTGCGCTTGCCCCTGAATCTTCG
 AGCCTCTAGATACCACAGGCTGCGCACTGTTGGGAGGGCGATCGGTGCGGCCTTCGCTATTACGCC
 AGCTGGCGAAAGGGGGATGTGCTGCAAGGCATAAGTTGGTAACGCCAGGGTTTCCCAGTCACGACG
 TTGTAAAACGACGGCAGTGAATTCATATTGCATGTCGCTATGTGTTCTGGAAATCACCATAAACGTGAAA
 5 TGTCTTGGATTGGGAATCTTATAAGTTCTGTATGAGACCACTCGGATCCGTTAGTGAACCGTCAGAACGCC
 ACCATGGGCAGCTGGCAGGATCTGAAAGATCATATGCGCGAAGCCGGCGTGGAAACCGATTCCAACC
 CGCTGCTGGGCCTCGATAGCACC GGCGATATGGTCGCAAAAAGCGCAGACAGCGCCGGAGGTAAATTTTT
 AAGCTGGCGTAATCATGGTCATAGCTGTTCTGTGAAATTGTTATCCGCTACAATTCCACACAACATA
 CGAGCCGGAAGCATAAAGTGTATCTAGAGCGGTACCACCGCGTAATTGAATTCAAGATCCACCGCGTGAATT
 10 CACTCCCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCAAAGGTGCGCCGACGCCCGGGCTT
 GCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCAGGGCGATGAACGTAATCGTAAACTAGCATGTCAA
 TCATATGTACCCCCGGTTGATAATCAGAAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAATATTAAATTGT
 AAGCGTTAATATTGTTAAAATTGCGTTAAATTTGTTAAATCAGCTCATTAAACCAATAGGCCGAAAT
 CGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTGGAACAAGAG
 15 TCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGCCCACTAC
 GTGAACCATCACCTAACTCAAGTTTTGGGTCGAGGTGCCGTAAGCACTAAATCGGAACCCCTAAAGGG
 GCCCCCGATTAGAGCTTGACGGGGAAAGCCGGCAACGTGGCGAGAAAGGAAGGGAAAGAACGCAAAG
 GAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACACCCGCCGCGCTTA
 ATGCGCCGCTACAGGGCGCTACTATGGTGCTTGACGAGCACGTATAACGTGCTTCCTCGTTAGAATCA
 20 GAGCGGGAGCTAACACAGGAGGCCGATTAAAGGGATTAGACAGGAACGGTACGCCAGAACCTTGAGAAGT
 GTTTTATAATCAGTGAGGCCACCGAGTAAAGAGTCTGTCATCACGCAAATTAAACCGTTGCGCAATACTT
 CTTGATTAGTAATAACATCACTGCGCTGAGTAGAAGAAACTCAAACATCGGCCCTGCTGGTAATATCCAGAA
 CAATATTACCGCCAGCCATTGCAACGGAATGCCATTGCCATTAGGCTGCGCAACTGTTGGGAAGGGCG
 ATCGGTGCGGGCCTTCGCTATTACGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGAGAGGCCG
 25 TTGCGTATTGGGC -3'

Functionality of scAAV9-driven expression of SRSF1-shRNA10 in mouse brains.

C9ORF72-ALS/FTD mice were injected via cisterna magna at post-natal day 1 (P1) with either
 8 × 10¹⁰ scAAV9_Ctrl-shRNA_GFP vector genomes (vg) or 6 × 10¹⁰ scAAV9_SRSF1-
 30 shRNA10_GFP vg. Animals were sacrificed 1 month and 3 months post injections. Western
 blots show that the scAAV9_SRSF1-shRNA10_GFP virus leads to specific depletion of
 SRSF1 in C9ORF72-ALS/FTD mice as well as in wild type C57Bl6 mice (not shown) while the
 Ctrl-shRNA has no effect. GAPDH is used as a loading control.

35 Example 1

2/ iMN imaging examples at day 24

High content automated imaging (Opera Phenix) was used to quantify surviving MNs at Day 1,
 2 and 3 of imaging. Images (Figure 2) show that MNs treated with lentivirus expressing

SRSF1-miRNA retain processes/axons characteristic of neurons compared to MN treated with LV_Ctrl-RNAi which generate and die.

3/ iMN quantification

- 5 Co-culture of healthy control and sALS patient-derived MN and astrocytes show that LV_SRSF1-RNAi specifically promotes sALS MN survival in levels comparable to the depletion of SRSF1 in C9ORF72-ALD patient-derived MNs (Hautbergue GM et al, Nature Communications 2017; 8:16063; Castelli et al. bioRxiv 2021.05.23.445325v2) Bar charts show MN survival expressed as a ratio of MNs quantified at counting day 3 over day 1 (%). 2-way
10 ANOVA with Tukey's multiple comparison test; NS: non-significant; **: p<0.01; ***: p<0.001;
****: p<0.0001 (Figure 3)

Example 2

4/ Testing the functionality of SRSF1-shRNAs in human cells and mouse brains

- 15 scAAV plasmids co-expressing GFP and SRSF1 shRNA 6, 9 or 10 were co-transfected with either sense or antisense C9ORF72-repeat reporter constructs expressing V5-tagged sense or antisense dipeptide repeat proteins (DPRs) in all frames in a repeat-associated non-AUG (RAN) translation manner. Western immunoblotting shows that all 3 shRNAs lead to efficient depletion of SRSF1 and inhibition of the RAN translation of V5-tagged DPRs. SRSF1-shRNA10 was selected for viral production and further experiments in mice as it has the lowest POTS score and predicted genome-wide off-target effect in both mouse and human.

Example 3

- scAAV SRSF1-shRNA, CPP1 and CPP2 inhibits the production of sense DPRs and rescue 25 the DPR-associated cytotoxicity in a human cell model of C9ORF72-ALS/FTD. Human HEK293T cells were co-transfected with sense G4C2x45 C9ORF72-repeat plasmid expressing sense V5-tagged dipeptide-repeat protein (DPRs) in a RAN translation manner and scAAV plasmids expressing SRSF1-shRNA10 or 2 different cell permeable peptides (CPP1: SRSF1 aa89-120 CPP (SEQ ID NO 59) and CPP2: SRSF1 aa132-144 CPP (SEQ ID 30 NO 75)). We tested potential expression under mammalian ubiquitous RNA polymerase II (CBh) or RNA polymerase III (H1) promoters. As shown in Figure 9 (A) Western blots show depletion of SRSF1 and inhibition of the RAN translation of sense DPRs upon co-transfection with scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP, but not when CPPs transcription is driven by the RNAPII promoter. SRSF1 and DPRs expression levels are 35 quantified in triplicate biological experiments in panels B and C respectively (Bar charts show mean ± SEM; 1-way ANOVA; NS: non-significant, ****: p < 0.0001). (D) MTT cell proliferation assays in biological triplicates showing that scAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP

and H1-CPP2_GFP alleviates the cytotoxicity mediated by the expression of DPRs, but not when CPPs transcription is driven the RNAPII promoter. Bar charts shows mean ± SEM; 1-way ANOVA; NS: non-significant, ****: p < 0.0001.

5 **Example 4**

sCAAV SRSF1-shRNA, CPP1 and CPP2 inhibits the production of antisense DPRs and rescue the DPR-associated cytotoxicity in a human cell model of C9ORF72-ALS/FTD. Human HEK293T cells were co-transfected with antisense G2C4x43 C9ORF72-repeat plasmid expressing antisense V5-tagged dipeptide-repeat protein (DPRs) in a RAN translation manner 10 and sCAAV plasmids expressing SRSF1-shRNA10 or 2 different cell permeable peptides (CPP1: SRSF1 aa89-120 CPP (SEQ ID NO 59) and CPP2: SRSF1 aa132-144 CPP (SEQ ID NO 75)). We tested potential expression under mammalian ubiquitous RNA polymerase II (CBh) or RNA polymerase III (H1) promoters. Figure 10 (A) Western blots show depletion of 15 SRSF1 and inhibition of the RAN translation of antisense DPRs upon co-transfection with sCAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP, but not when CPPs transcription is driven the RNAPII promoter. SRSF1 and DPRs expression levels are quantified in triplicate biological experiments in panels B and C respectively (Bar charts shows mean ± SEM; 1-way ANOVA; NS: non-significant, ****: p < 0.0001). Note that there is no 20 expression of CPP1 or CPP2 from the protein-coding CBh promoter. (D) MTT cell proliferation assays in biological triplicates showing that sCAAV SRSF1-shRNA10_GFP, H1-CPP1_GFP and H1-CPP2_GFP alleviates the cytotoxicity mediated by the expression of DPRs, but not when CPPs transcription is driven the RNAPII promoter. Bar charts shows mean ± SEM; 1-way ANOVA; NS: non-significant, ****: p < 0.0001.

25 **Example 5**

The data for SRSF1-shRNA (Figure 12) shows that the svAAV9-SRSF1-shRNA10 virus leads to inhibition of the DPRs in mouse brains and complements the data showing that it leads to SRSF1 depletion in mouse brains.

30 **Example 6**

The data shows that the sCAAV9 virus expressing CPP1 or CPP2 and co-expressing GFP efficiently transduced neuronal and glial cells in mouse brains (Figure 13) and leads to DPR inhibition (Figure 14).

CLAIMS

1. A viral vector comprising a transcription cassette for the expression of a nucleic acid molecule in a mammalian host cell wherein said nucleic acid molecule is operably linked to a promoter adapted to express said nucleic acid molecule in said mammalian host cell characterised in that said vector comprises a non-expressed nucleotide sequence and wherein said nucleic acid molecule encodes an antagonistic agent that targets Serine/Arginine Rich Splice Factor (SRSF1).
10
2. The viral vector according to claim 1 wherein said SRSF1 comprises or consist of a sequence set forth in SEQ ID NO 67.
3. The viral vector according to claims 1 or 2 wherein said antagonistic agent is a nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide or peptide.
15
4. The viral vector according to claims 1 or 2 wherein said antagonistic agent is a nucleic acid-based agent.
5. The viral vector according to any one of claims 1-2 or 4 wherein said nucleic acid-based agent is an antisense nucleic acid, an inhibitory RNA, a shRNA or miRNA molecule that is complementary to and inhibits the expression of a nucleic acid encoding a Serin/Arginine Rich Splice Factor (SRSF1).
20
6. The viral vector according to claim 5 wherein said inhibitory RNA comprises or consists of a nucleotide sequence as set forth in SEQ ID NO: 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57 or 58.
25
7. The viral vector according to claim 5 wherein said shRNA comprises or consist of a nucleotide sequence selected from the group consisting of SEQ ID NO 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11.
30
8. The viral vector according to claim 7 wherein said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 7.
35
9. The viral vector according to claim 7 wherein said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 10.

10. The viral vector according to claim 7 wherein said shRNA comprises or consist of a nucleotide sequence set forth in SEQ ID NO 11.
- 5 11. The viral vector according to claim 3 wherein said peptide comprises an amino acid sequence that is at least 32 amino acids in length and comprises the amino acid sequence set forth in SEQ ID NO: 59.
- 10 12. The viral vector according to claim 3 wherein said peptide is a dominant negative protein comprising a modification of the amino acid sequence set forth in SEQ ID NO: 60 or 61.
- 15 13. The viral vector according to claim 3 wherein said modified protein comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 62 or 63.
- 20 14. The viral vector according to claim 3 wherein said peptide comprises an amino acid sequence that is at least 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40 or 42 amino acids in length and set forth in SEQ ID NO: 90.
- 25 15. The viral vector according to claims 3 or 14 wherein said peptide comprises an amino acid sequence that is set forth in SEQ ID NO: 75 (GSWQDLKDHMREA).
- 30 16. The viral vector according to any one of claims 1-3 and 11-15 wherein said viral vector comprises a RNA Pol III terminator.
- 35 17. The viral vector according to any one of claims 1-16 wherein said vector comprises inverted terminal repeat nucleotide sequences, and optionally wherein said ITR sequences are set forth in SEQ ID NO 64 or in SEQ ID NO 88.
- 40 18. The viral vector according to any one of claims 1-17 wherein said promoter is selected from the group consisting of H1 Polymerase III promoter, U6 promoter, U7 promoter or the mammalian 7SK promoter.
- 35 19. The viral vector according to claim 18 wherein said vector is a H1 Polymerase III promoter, and optionally is set forth in SEQ ID NO 65.
- 40 20. The viral vector according to any one of claims 1-19 wherein said viral based vector is an adeno-associated virus [AAV], and optionally a self-complementary adeno-associated virus (scAAV).

21. The viral vector according to claim 20 wherein said viral based vector is scAAV9 or
scAAVrh10.
22. A pharmaceutical composition comprising a viral vector according to any one of
5 claims 1-21 and an excipient or carrier.
23. A viral vector according to any one of claims 1-21 for use as a medicament.
24. A viral vector according to any one of claims 1-21 for use in the treatment of a
10 neurodegenerative disease.
25. The viral vector for use according to claim 24 wherein said neurodegenerative disease
is amyotrophic lateral sclerosis (ALS).
- 15 26. The viral vector for use according to claim 24 wherein said neurodegenerative disease
is sporadic and/or familial amyotrophic lateral sclerosis .
27. The viral vector for use according to claim 24 wherein said neurodegenerative disease
is ALS not caused by pathological C9ORF72-repeat expansions
- 20 28. The viral vector for use according to claim 24 wherein said neurodegenerative disease
is sporadic frontotemporal dementia (FTD).
29. The viral vector for use according to claim 24 wherein said neurodegenerative disease
25 is Fragile X-associated tremor/ataxia syndrome (FXTAS).
30. A cell transfected with a viral vector according to the invention.
31. The cell according to claim 30 wherein said cell is a neurone and/or an astrocyte, and
30 optionally wherein said neurone is a motor neurone and/or an astrocyte.
32. An isolated nucleic acid molecule encoding an shRNA molecule comprising or
consisting of a nucleotide sequence selected from the group consisting of SEQ ID NO 2, 3,
35 4, 5, 6, 7, 8, 9,10 and 11.
33. The isolated nucleic acid molecule according to claim 32 wherein said nucleic acid
molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 7.
34. The isolated nucleic acid molecule according to claim 32 wherein said nucleic acid
40 molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 10.

35. The isolated nucleic acid molecule according to claims 32 wherein said nucleic acid molecule comprises or consist of a nucleotide sequence set forth in SEQ ID NO 11.
36. A cell penetrating polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NO 90.
- 5
37. The cell penetrating peptide according to claim 36 wherein said polypeptide is between 13-42 amino acids in length.
- 10 38. The cell penetrating peptide according to claim 37 wherein said polypeptide comprises or consist of an amino acid sequence set forth in SEQ ID NO 75.
- 15 39. An antagonistic agent comprising a nucleic acid molecule wherein said nucleic acid molecule comprises a nucleotide sequence designed with reference to human Serine/Arginine Rich Splice Factor (SRSF1) and wherein said nucleic acid molecule inhibits expression of SRSF1.
- 20 40. The agent according to claim 39 wherein said nucleic acid molecule is a double stranded nucleic acid molecule comprising a sense strand and an antisense strand comprising nucleotide sequences wherein said antisense nucleotide strand is adapted to anneal by complementary base pairing to a nucleic acid molecule encoding human SRSF1.
- 25 41. The agent according to claim 40 wherein said double stranded nucleic acid molecule is RNA.
42. The agent according to claim 41 wherein said RNA is siRNA or miRNA.
- 30 43. The agent according to claim 39 wherein said nucleic acid molecule is a single stranded nucleotide sequence comprising an antisense nucleotide sequence wherein said antisense nucleotide sequence is adapted to anneal by complementary base pairing to a nucleic acid molecule encoding SRSF1.
- 35 44. The agent according to claim 43 wherein said single stranded nucleic acid is DNA and/or RNA, and optionally, said DNA and/or RNA is a therapeutic antisense oligonucleotide such as an antisense oligonucleotide, a splice-switching oligonucleotide, a gapmer or similar.
45. The agent according to claim 44 wherein said DNA is an antisense oligonucleotide.

46. The agent according to any one of claims 39 to 45 wherein said nucleic acid molecule encoding human SRSF1 is set forth in SEQ ID NO: 67.
- 5 47. The agent according to claim 46 wherein antagonistic agent comprises a nucleic acid molecule that is at least 15 nucleotides in length.
- 10 48. The agent according to claim 47 wherein said antagonistic agent comprises a nucleic acid molecule comprising a nucleotide sequence set forth in SEQ ID NO: 67 wherein said nucleic acid molecule is a double stranded inhibitory RNA and is 19-23 nucleotides in length.
- 15 49 The agent according to any one of claims 39 to 48 wherein said antagonistic agent comprises a nucleic acid molecule comprising modified nucleotides and/or modified sugars.
50. The agent according to claim 49 wherein said double stranded nucleic acid molecule comprising sense and antisense nucleic acid molecules comprise modified nucleotides.
- 20 51. The agent according to claim 50 wherein said modified nucleotides are selected from the group: a 3'-terminal deoxy-thymine (dT) nucleotide, a 2'-0-methyl modified nucleotide, a 2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, a locked nucleotide, an unlocked nucleotide, a conformationally restricted nucleotide, a constrained ethyl nucleotide, an abasic nucleotide, a 2'-amino-modified nucleotide, a 2'-0-allyl-modified nucleotide, 2'-C-alkyl-modified nucleotide, 2' -hydroxyl- modified nucleotide, a 2'-methoxyethyl modified nucleotide, a 2'-0- alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, a non-natural base comprising nucleotide, a tetrahydropyran modified nucleotide, a 1,5-anhydrohexitol modified nucleotide, a cyclohexenyl modified nucleotide, a nucleotide comprising a phosphorothioate group, a nucleotide comprising phosphorodithioate (PS2), a nucleotide comprising a methylphosphonate group, a nucleotide comprising a 5'- phosphate, and a nucleotide comprising a 5'-phosphate mimic, for example a 5'-vinyl phosphate, a nucleotide comprising a 2'-deoxy-2'-fluro and a 2' methyl sugar base.
- 25 52. The agent according to any one of claims 49 to 51 wherein said double stranded nucleic acid molecule comprising sense and antisense nucleic acid molecules comprise modified sugar(s).
- 30 53. The agent according to claim 52 wherein said modified sugar is selected from the group: a modified version of the ribosyl moiety, such as -O- modified RNA such as 2'-O-alkyl or 2'-O-(substituted)alkyl e.g. 2'-0-methyl, T-0-(2- cyanoethyl), 2'-0-(2-methoxy)ethyl (2'-MOE), 2'-0-(2-thiomethyl)ethyl, 2'-O-butyryl, -O- propargyl, 2'-O-allyl, 2'-O-(2-amino)propyl, 2'-

O-(2-(dimethylamino)propyl), 2'-O-(2- aminoethyl, 2'-O-(2-(dimethylamino)ethyl); 2'-deoxy (DNA); 2'-O-(haloalkoxy)methyl, e.g. 2'-O-(2-chloroethoxy)methyl (MCEM), -O- (2,2-dichloroethoxy)methyl (DCEM); 2'-<3-alkoxycarbonyl e.g. T-0-[2- (methoxycarbonyl)ethyl] (MOCE), 2'-O-[2-(N-methylcarbamoyl)ethyl] (MCE), T-0-[2-(N,N- dimethylcarbamoyl)ethyl] (DCME); 2'-halo e.g. 2'-F, FANA (2'-F arabinosyl nucleic acid); carbasugar and azasugar modifications; 3 '-O-alkyl e.g. 3'-O-methyl, 3 '-O-butyryl, V-O- propargyl and their derivatives.

5 54. The agent according to any one of claims 49 to 53 wherein said antagonistic agent comprises a nucleotide sequence designed with reference to the target nucleic acid sequences selected from the group:

10 TGGCACTGGTGTGAGTTGTA (SEQ ID NO 110);
TGGTGTGAGTTGACGGAAA (SEQ ID NO 111);
TCGTGGAGTTGTACGGAAAGAAGA (SEQ ID NO 112);
AAGATATGACCTATGCAGTCGAAA (SEQ ID NO 113);
GAGAAACTGCCTACATCCGGGTTAA (SEQ ID NO 114);
15 CGGGTAAAGTTGATGGGCCAGAA (SEQ ID NO 115);
TGATGGGCCAGAACAGTCCAAGTTAT (SEQ ID NO 116);
CAGAAGTCCAAGTTATGGAAGATCT (SEQ ID NO 117);
GAGAAGCAGAGGATCACCAACGCTAT (SEQ ID NO 118); and
CGTCATAGCAGATCTCGCTCTCGTA (SEQ ID NO 119).

20 55. A pharmaceutical composition comprising an antagonist agent according to any one of claims 39 to 54 and including an excipient or carrier.

56. An antagonistic agent according to any one of claims 39 to 54 for use as a medicament.

25 57. An antagonistic agent according to any one of claims 39 to 54 for use in the treatment of a neurodegenerative disease.

58. The antagonistic agent according to claim 57 wherein said neurodegenerative disease is amyotrophic lateral sclerosis (ALS).

30 59. The antagonistic agent according to claim 58 wherein said neurodegenerative disease is sporadic and/or familial amyotrophic lateral sclerosis.

35 60. The antagonistic agent according to claim 58 or 59 wherein said neurodegenerative disease is ALS not caused by pathological C9ORF72-repeat expansion.

61. The antagonistic agent according to claim 57 wherein said neurodegenerative disease is sporadic frontotemporal dementia (FTD).
- 5 62. The antagonistic agent according to claim 57 wherein said neurodegenerative disease is Fragile X-associated tremor/ataxia syndrome (FXTAS).
63. A shRNA molecules comprising a nucleotide sequence, or variant thereof, selected from the group consisting of:
SRSF1-shRNA1 (SEQ ID NO 91):
10 GCUGAUGUUUACCGAGAUGGC UUCAAGAGA GCCAUCUCGGUAAACAUCAUCAGC;
SRSF1-shRNA2 (SEQ ID NO 92):
GGAGUUUGUACGGAAAGAAGA UUCAAGAGA UCUUCUUUCCGUACAAACUCC;
SRSF1-shRNA3 (SEQ ID NO 93):
GGAAAGAAGAU AUGACCUAUG UUCAAGAGA CAUAGGUCAUAUCUUCUUCC;
15 SRSF1-shRNA4 (SEQ ID NO 94):
GAAAGAAGAU AUGACCUAUGC UUCAAGAGA GCAUAGGUCAUAUCUUCUUUC;
SRSF1-shRNA5 (SEQ ID NO 95):
GCCUACAUCGGGUAAAGUU UUCAAGAGA AACUUUAACCCGGAUGUAGGC;
SRSF1-shRNA6 (SEQ ID NO 96):
20 GGGCCCAGAAGGUCCAAGUUUAU UUCAAGAGA AUAACUUGGACUUCUGGGCCC;
SRSF1-shRNA7 (SEQ ID NO 97):
GGCCCAGAAGGUCCAAGUUUAUG UUCAAGAGA CAUAACUUGGACUUCUGGGCC;
SRSF1-shRNA8 (SEQ ID NO 98):
GCCCAGAAGGUCCAAGUUUAGG UUCAAGAGA CCAUAACUUGGACUUCUGGGC;
25 SRSF1-shRNA9 (SEQ ID NO 99):
GGAAGAUCUCGAUCUCGAAGC UUCAAGAGA GCUUCGAGAUCGAGAUCUUC; and
SRSF1-shRNA10 (SEQ ID NO 100):
GCAGAGGAUCACCACGCUAUU UUCAAGAGA AAUAGCGUGGUGAUCCUCUGC.
- 30 64. The shRNA molecule according to claim 63 wherein said shRNA molecule comprises or consist of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 96.
65. The shRNA molecule according to claim 63 wherein said shRNA molecule comprises or consist of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 99.

66. The shRNA molecule according to claim 63 wherein said shRNA molecule comprises or consist of a nucleotide sequence, or variant thereof, set forth in SEQ ID NO 100.

5

10

15

20

25

30

35



Application No: GB2312083.5

Examiner: Dr Graham Feeney

Claims searched: In part 1-31

Date of search: 5 February 2024

Patents Act 1977: Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1-31	US 2019/0194660 A1 (HAUTBERGUE et al.) The whole document, esp. see the claims
A	-	WO 2014/007858 A1 (UNIV IOWA RES FOUND) claim 1
A	-	WO 2017/161273 A1 (THE CHILDREN'S HOSPITAL OF PHILADELPHIA) claim 1

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^X :

Worldwide search of patent documents classified in the following areas of the IPC

The following online and other databases have been used in the preparation of this search report

SEARCH-PATENT, SEARCH-NPL, Cas Online



International Classification:

Subclass	Subgroup	Valid From
C12N	0015/864	01/01/2006
A61K	0031/713	01/01/2006
A61K	0038/16	01/01/2006
A61K	0038/17	01/01/2006
A61P	0025/28	01/01/2006
C07K	0014/47	01/01/2006
C12N	0015/113	01/01/2010