Nucleic Acids Research, 2025, **53**, gkaf253
https://doi.org/10.1093/nar/gkaf253
Data Resources and Analyses

High-throughput screen of 100 000 small molecules in C9ORF72 ALS neurons identifies spliceosome modulators that mobilize G4C2 repeat RNA into nuclear export and repeat associated non-canonical translation

Maartje J. Luteijn[1,*], Varun Bhaskar[2,†], Dominic Trojer[1,†], Melanie Schürz[3,7], Hicham Mahboubi[2], Cornelia Handl[1], Nicolas Pizzato[1], Martin Pfeifer[1], Ruxandra Dafinca[4], Hans Voshol[1], Elisa Giorgetti[1], Carole Manneville[1], Isabelle P.M. Garnier[1], Matthias Müller[1], Fanning Zeng[1], Kathrin Buntin[1], Roger Markwalder[1], Harald Schröder[1], Jan Weiler[1], Dora Khar[1], Tim Schuhmann[1], Paul J. Groot-Kormelink[1], Caroline Gubser Keller[1], Pierre Farmer[1], Angela MacKay[1], Martin Beibel[1], Guglielmo Roma[1], Giovanni D'Ario[1], Claudia Merkl[1], Michael Schebesta[5], Marc Hild[5], Fiona Elwood[5], Björn F. Vahsen[6], Nina Ripin[1,6], Antoine Clery[6], Frederic Allain[6], Mark Labow[5], Daniela Gabriel[1], Jeffrey A. Chao[2], Kevin Talbot[4], Mark Nash[1], Jürg Hunziker[1,*], Nicole C. Meisner-Kober[1,3,7,*]

[1]Novartis Institutes for Biomedical Research, Department Global Discovery Chemistry, Basel, 4056, Switzerland
[2]Friedrich Miescher Institute for Biomedical Research, Department Genomic Regulation, Basel, 4056, Switzerland
[3]Paris-Lodron University of Salzburg, Department of Biosciences and Medical Biology, , Salzburg, 5020, Austria
[4] University of Oxford, John Radcliffe Hospital, Nuffield Department of Clinical Neurosciences, Oxford, OX3 9DU, United Kingdom
[5]Novartis Institutes for Biomedical Research, Department Discovery Sciences, Cambridge, MA 02139, United States
[6]ETH Zürich, Department of Biology, Institute f. Molekularbiol.u.Biophysik, Zürich, 8093, Switzerland
[7]Ludwig Boltzmann Institute for Nanovesicular Precision Medicine at the Paris Lodron University Salzburg, 5020, Austria

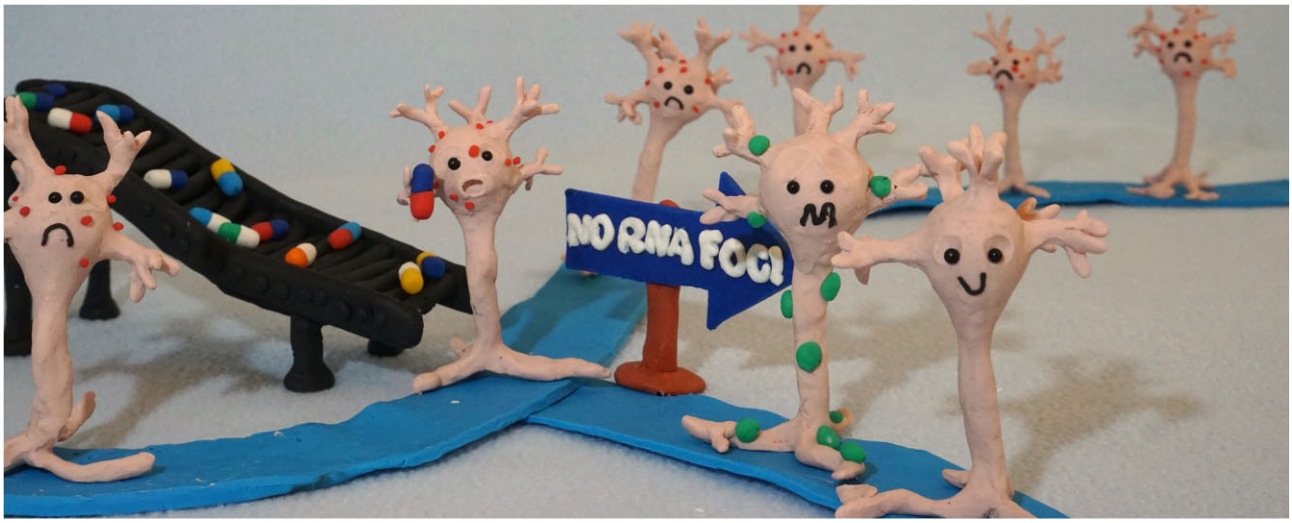
[*]To whom correspondence should be addressed: Email: nicole.meisner-kober@plus.ac.at
Correspondence may also be addressed to Maartje J. Luteijn. Email: m.luteijn@genetwister.nl
Correspondence may also be addressed to Jürg Hunziker. Email: juerg.hunziker@novartis.com
[†]Varun Bhaskar and Dominic Trojer should be regarded as Joint Second Authors.

Abstract

An intronic G4C2 repeat expansion in the C9ORF72 gene is the major known cause for Amyotrophic Lateral Sclerosis (ALS), with current evidence for both, loss of function and pathological gain of function disease mechanisms. We screened 96 200 small molecules in C9ORF72 patient iPS neurons for modulation of nuclear G4C2 RNA foci and identified 82 validated hits, including the Brd4 inhibitor JQ1 as well as novel analogs of Spliceostatin-A, a known modulator of SF3B1, the branch point binding protein of the U2-snRNP. Spliceosome modulation by these SF3B1 targeted compounds recruits SRSF1 to nuclear G4C2 RNA, mobilizing it from RNA foci into nucleocytoplasmic export. This leads to increased repeat-associated non-canonical (RAN) translation and ultimately, enhanced cell toxicity. Our data (i) provide a new pharmacological entry point with novel as well as known, publicly available tool compounds for dissection of C9ORF72 pathobiology in C9ORF72 ALS models, (ii) allowing to differentially modulate RNA foci versus RAN translation, and (iii) suggest that therapeutic RNA foci elimination strategies warrant caution due to a potential storage function, counteracting translation into toxic dipeptide repeat polyproteins. Instead, our data support modulation of nuclear export via SRSF1 or SR protein kinases as possible targets for future pharmacological drug discovery.

Received: December 20, 2024. **Revised:** March 3, 2025. **Editorial Decision:** March 5, 2025. **Accepted:** April 7, 2025

Graphical abstract



Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal disease that is caused by irreversible loss of motor neurons. Patients suffer from progressive paralysis and typically have a life expectancy of two to five years after being diagnosed [1]. Developing an effective therapy has been unsuccessful so far, in part due to the heterogeneity of genetic causes with ~130 genes linked to the disease [2], as well as the still largely unresolved molecular pathology. The most frequent genetic cause of ALS known to date is a G4C2 hexanucleotide repeat expansion in the first intron of the *C9ORF72* gene [3, 4]. In Northern European and Northern American populations, this mutation has been identified in approximately 40% of patients with familial ALS (fALS) and 8–10% of sporadic ALS cases (sALS) [5], and has also been associated with frontotemporal dementia (FTD), whereas it is less frequent in Asia and other parts of the global population [3, 4]. While up to 30x G4C2-repeats have been described in healthy individuals, C9ORF72 patients carry typically between 500 and 1200 (in some cases >4000) repeats, due to either inheritance of a mutant allele or spontaneous expansion through R-loop formation or other, unknown mechanisms [6, 7].

The molecular disease mechanism of the C9ORF72 mutation remains under debate. C9ORF72 haplo-insufficiency and/or a toxic gain of function of the repeat expansion are the current main hypotheses [8]. The G4C2 repeat RNA is generated from the *C9ORF72* mutant locus [9], folds into a stable G-quadruplex (G4) structure [10] and accumulates into nuclear RNA foci [11]. Many groups have tried to resolve what exact RNA sequences are found within the RNA foci: abortive or ineffectively spliced C9ORF72 transcripts, retained spliced-out repeat containing introns, non-canonical repeat-initiated transcripts, and/or mere G4C2 RNA remaining from inefficient processing of any of these. A study using single molecule FISH demonstrated that the RNA foci mainly consist of G4C2 containing lariat intermediates from splicing, which are stabilized due to the G4C2 repeat, exported to the cytoplasm and fed into cap-independent translation [12]. The repeat expansion is also transcribed in the antisense direction, giving rise to additional formation of C4G2-RNA/antisense repeat foci. RNA sense and antisense foci are also present in the nuclei of astrocytes, microglia, oligodendrocytes and fibroblasts [9] as well as, less frequently, in the cytoplasm of some cells [13].

By analogy with the CUG repeats in myotonic dystrophy type 1 [14], these RNA foci were proposed to sequester RNA binding proteins (RBPs), thereby resulting in potentially neurotoxic defects in RNA processing/metabolism [15, 16]. Extensive efforts have been undertaken to identify "the muscleblind protein of C9ORF72 ALS" in analogy to Myotonic Dystrophy Type I (DM1), where loss of function of the splice factors MBNL-1 (muscleblind-like 1) and CUGBP (CUG binding protein 1) are well established as major downstream effectors of disease pathophysiology. However, no individual predominant splice defects have been identified as single pathological cause for C9ORF72 so far. Rather, the opposite was shown; single cell (nuclear) transcriptome and epigenome sequencing of autopsied motor and frontal cortices revealed distinct molecular pathologies across different cell types and brain regions [17].

G4C2 repeat RNA can be transported to the cytoplasm via the splice-factor SRSF1 and nuclear export factor NXF1 [18] and be targeted for RNA decay via the RNA exosome [19]. Additionally, the G4C2 as well as C4G2 repeat RNAs get translated by a non-canonical, repeat-associated, non-AUG mechanism [repeat-associated non-canonical (RAN) translation]. RAN translation occurs in all three reading frames, thereby creating GA-, GR-, GP-, PR-, and PA dipeptide repeat polyproteins (DPRs) [10, 20–26]. These DPRs are prone to aggregation and, when overexpressed, have been shown to be toxic in various models [27–32]. DPR overexpression has been associated with disruption of nucleocytoplasmic trafficking [33, 34], ribosomal RNA biogenesis, mitochondrial function [35], and pre-mRNA splicing [36–39]. In a mouse model expressing an expanded hexanucleotide repeat sequence (C9-BAC mouse), DPR toxicity was reduced by high-affinity human antibodies targeting GA or GP DPRs [40]. However, a study demonstrated that DPRs accumulated for decades in an FTD patient before symptoms appeared [41], and mice with widespread poly-GP expression did not develop any overt motor phenotype [42], questioning the role of DPR toxicity

in the development of ALS/FTD. Finally, the repeat also epigenetically and transcriptionally alters expression of the *C9ORF72* gene, leading to reduced levels of C9ORF72 protein [3, 24, 43]. C9ORF72 haploinsufficiency has been implicated in impaired vesicular trafficking, axonal transport, autophagy deficits and synaptic dysfunction [44–48] and is therefore being investigated as an additional potential disease mechanism. A few articles support a disease mechanism in ALS/FTD driven by the combined effects of C9ORF72 loss of function and the toxic gain of function from the G4C2 repeat expansion [47, 48].

Despite all of this progress, it is still not entirely clear how and to which extent each of these three alterations contribute to the development of ALS [16]. The major difficulty is to study RNA toxicity independent of DPR toxicity in patient cells, since their biogenesis is linked and removing the repeat RNA inherently leads to loss of DPRs. By over-expression of non-repetitive constructs that only generate DPRs and no RNA foci, DPRs were shown to be toxic independent of G4C2 repeat RNA [49]. Also "RNA only" constructs were made, to address if RNA can be toxic independent of DPRs. These constructs lack a start codon and have a stop codon every 15 repeats, to prevent DPR formation. Initially these constructs were found not to be toxic in *Drosophila* [29, 50] but in later studies showed some toxicity in zebrafish models [51]. A study in a small cohort of 63 C9ORF72 patients however showed no correlation between the number of RNA foci and the burden of ALS pathology at autopsy [52].

Recent research on therapeutic strategies for C9ORF72-related ALS and FTD has focused on antisense oligonucleotides (ASO) targeting the G4C2 RNA either within or outside the repeat. ASOs targeted to the repeat have been shown to clear RNA foci effectively, reduce DPR levels, and improve behavioral deficits in patient cell lines and mouse models [(53–55), and reviewed in Ly & Miller, 2018 (56)]. Follow up studies with repeated intrathecal delivery of a backbone optimized, repeat targeted ASO in a single C9ORF72 ALS patient was well tolerated, leading to significant reductions in levels of cerebrospinal fluid poly(GP) [57]. More advanced ASOs, like WVE-004 by Wave Life Sciences were developed to selectively reduce repeat-containing transcripts and DPR levels despite targeting a sequence outside the repeat, thereby also reducing RAN DPR levels and protecting motor neurons from toxicity [55]. However, clinical development of both variant selective ASO candidates tested in Phase 1 C9ORF72 ALS clinical trials so far was discontinued due to a lack of efficacy and/or safety; BIIB078, a repeat targeted ASO developed by Ionis and Biogen, not only reduced the levels of toxic RAN DPR proteins in the cerebrospinal fluid but also increased neurofilament light chain, a marker of neuronal damage, and its development was discontinued already after Phase 1 due to potential safety concerns [58]. WVE-004 [59] successfully reduced poly-GP levels in the cerebrospinal fluid and was well tolerated but yet did not demonstrate a therapeutic benefit to the patients. These outcomes raise the question whether it is sufficient to target the sense G4C2 RNA for treating C9 ALS/FTD . Researchers are also exploring alternative ASO approaches and other modalities to treat C9ORF72-related ALS. For example, ASOs targeting the antisense strand (rG2C4) have shown potential in reducing TDP-43 pathology in patient-derived neurons [60]. Immunotherapies targeting polyGA DPRs have also demonstrated efficacy in mouse models, reducing molecular and behavioral deficits without altering G4C2 repeat RNA levels [61]. Gene editing tools, such as CRISPR/Cas9, have shown promising results for excising the G4C2 expansion from the C9ORF72 gene, potentially restoring normal function of the locus. While proof-of-concept studies demonstrated successful removal of the repeat expansion *in vitro* and in patient-derived motor neurons [61–63], substantial further preclinical research will be needed to resolve in particular hurdles of drug delivery for efficient therapeutic gene editing in relevant tissues and cells.

Efforts to identify small molecules to eliminate RNA foci have initially focused on small sets of known scaffolds targeting either CGG repeat or G-quadruplex RNA and DNA directly [64, 65]. These compounds were further matured into chimeric scaffolds to recruit endonucleases to the repeat RNA, inducing targeted RNA degradation [66], whereas the safety and thus clinical viability of these aromatic RNA stacking compounds remains to be investigated. A recent screen of ca. 3300 annotated bioactive molecules using an *in vitro* RAN translation assay has identified five validated hits, including the relatively unspecific intercalator propidium iodide and two other compounds binding the repeat RNA directly [67]. Furthermore, new strategies to alleviate C9ORF72 haploinsufficiency include small molecule inhibitors of PKIFYVE kinase [68, 69]. While interesting starting points, the mode of action, specificity and activity of these compounds in cell-based models of ALS still remain to be investigated. Thus, despite exciting progress, there is still a need for potent pharmacological tools that specifically target distinct regulatory steps in G4C2-RNA or DPR metabolism to accelerate our progress in investigating and drugging C9ORF72 ALS.

Here, we report an unbiased screen of 96 200 compounds for RNA foci clearance in ALS patient derived neurons, which identified a set of pharmacological tool compounds to interfere with either expression from the C9ORF72 locus or G4C2 repeat RNA metabolism. Characterization of the mode of action of two prioritized hits establishes a new mechanistic link between SF3B1 / the U2 snRNP subunit of the spliceosome and G4C2 repeat RNA foci metabolism and allows to differentially modulate nuclear RNA foci versus RAN DPRs. By further characterizing these tool compounds, our data support the hypothesis that nuclear RNA foci formation is a measure to protect neurons against ALS and FTD development by sequestering the G4C2 RNA, thereby counteracting RAN translation.

Material and methods

Biological resources

Cell lines

The fibroblast lines ND40063, ND41002 as well as the lymphoblastoid cell lines ND08980 were obtained from the NINDS repository at the Coriell Institute for Medical Research. Control fibroblast lines from healthy donors were ordered from Invitrogen. An additional C9ORF72 ALS fibroblast line was kindly provided by the McGill University. All primary cell lines are additionally specified in Supplementary Table S1. All lines were characterized by Southern Blotting. HeLa (ATCC CCL2), HEK293T, NCI-H23 [H23] lung adenocarcinoma cells were obtained from ATCC® (CRL-5800™). iPSC lines that were differentiated into motor neurons were derived from skin biopsy fibroblasts, collected under ethical approval granted by the South Wales Research Ethics

Committee (WA/12/0186) in the James Martin Stem Cell Facility, University of Oxford, under standardized protocol. For details see Dafinca *et al.* (2020) [70]. Cell line Ed-C9 is the C902-02 iPS line in which 1000 G4C2 repeats were replaced by 2 G4C2 repeats using CRISRP-Cas9, for detailed information see Ababneh *et al.* (2020) [71]

Constructs

Human Ngn2 cDNA was gene synthesized according to sequences available from the Ensembl database (Ensembl Gene ID ENSG00000178403 or accession number NM_024019.3) and cloned under the control of TRE tight (Tetracycline Response Element) promoter in a PiggyBac/Tet-ON all-in-one vector. This vector contains a CAG rtTA16 cassette allowing constitutive expression of Tet-ON system and an Hsv-tkNeo cassette for generation of stable iPS clones.

For the cell lines expressing 152x G4C2 repeats, a short 19xGGGGCC-repeat construct was synthesized by GenewizTM with flanking restriction enzyme sites to facilitate unlimited exponential GGGGCC-repeat cloning steps to extend the length of the repeat. The construct [AvrII-CCC:GGG(=SmaI)GCC-17xG4C2-GGG-GCC:GGC(=NaeI)GCCACCATG-BstBI] was cloned into the pTRE3G plasmid (ClontechTM) using AvrII-BstBI cloning sites. Subsequent serial G4C2 repeat extending cloning steps were performed by digesting out the G4C2 repeat insert with SmaI-BstBI whilst also digesting the same longest repeat construct backbone plasmid with NaeI-BstBI. The insert was then ligated into the repeat containing backbone plasmid with the blunt SmaI and NaeI one sites combining into a perfect GGGGCC sequence and the BstBI sites guaranteeing directional cloning (sticky end ligation). This guaranteed destroying the original SmaI-NaeI sites so the same cloning step can be repeated indefinitely. pTRE3G-152x repeat construct was recloned in a PiggyBac backbone for generation of stable iNeuron lines, under selection of puromycin. For the RAN-expressing HEK293 cells, the construct was further adjusted by adding Myc, Flag and HA tags in different reading frames at the end of the 152x repeats.

Oligonucleotides

- ASOs were synthesized by standard solid phase synthesis using 2′MOE-PS-DNA chemistry, and purified by ion exchange chromatography. A3, C5, and the GAPDH targeted ASO were Gapmers with five bases of 2′MOE-PS modified DNA bases in the 5′ and 3′ flanking regions, and unmodified DNA at the central 10 nucleotides to allow for RNAseH cleavage. A2 was fully 2′MOE-PS modified (steric blocker). The ASOs were added as a final concentration of 5 μM to the cell growth media. ASOs were incubated for 24 or 72 h before the cells were fixed with 4% PFA and check for RNA foci by FISH. ASOs were same as in Donnelly *et al.* (2013) [43]. Sequences:

A2: 5′-CCGGCCCCGGCCCCGGCCCC-3′
A3: 5′-CCGGCCCCGGCCCCGGCCCC-3′
C5 (5′-GCCTTACTCTAGGACCAAGA-3′)

- G4C2 oligo's / bio-RNA IP. RNA oligo's were synthesized by Microsynth.

(G4C2)4: 5′GGGGCCGGGGCCGGGGCCGGGGCC'3-biotin
(C4G2)4: 5′CCCCGGCCCCGGCCCCGGCCCCGG'3-biotin
(A4U2)4: 5′AAAAUUAAAAUUAAAAUUAAAAUU'3-biotin
(A4U2)4: 5′AAAAUUAAAAUUAAAAUUAAAAUU3'
(G4C2)4: 5′GGGGCCGGGGCCGGGGCCGGGGCC'3

- Oligos for FCS: unmodified or 5′TMR-C6-modified RNA oligo's were obtained from Microsynth.

SRSF1 canonical motif: 5′TMR-AAAUCAGAGGAAAA;
(G4C2)4:5′-GGGGCCGGGGCCGGGGCCGGGGCC-3′ (A4U2)4: 5′-AAAAUUAAAAUUAAAAUUAAAAUU-3′;
(G4C2)2: 5′TMR-GGGGCCGGGGCC-3′

- RT-qPCR primers / TaqMans were ordered from Invitrogen

Exon 2–3 junction of C9ORF72 (Hs00376619_m1)- FAM
Intron after exon 1b (custom design ID: AICSWFV)-FAM
18S-VIC (Catalog #: 4319413E)

- RT-PCR primers

eGFP_qPCR_Fw	AGTCCGCCCTGAGCAAAGA
eGFP_qPCR_Rv	TCCAGCAGGACCATGTGATC
β-actinFw_intron4	TCTGCCTGACATGAGGGTTA
β-actin_Rv_Exon5	AGCACTGTGTTGGCGTACAG
C9_Exon1a_Fw	TCAAACAGCGACAAGTTCCG
C9_intron1a_Rv	GGAGAGAGGGTGGGAAAAAC
18S_Fw	TCGAGGCCCTGTAATTGGAA
18S_Rv	CTTTAATATACGCTATTGGAGCTGGAA
MCL-1_Fw	AGACCTTACGACGGGTTGG
MCL-1_Rv	ACCAGCTCCTACTCCAGCAA
C9_Ex_Intr1b_F1	CGAGTGGGTGAGTGAGGAG
C9_intron1b_R1	AGTCGCTAGAGGCGAAAGC

- siRNAs
- SMARTpool: ON-TARGETplus NXF1 siRNA (Dharmacon, # L-013680–01-0005)
- SMARTpool: ON-TARGETplus SRSF1 siRNA (Dharmacon, # L-018672–01-0005)
- ON-TARGETplus Non-targeting Pool (Dharmacon #D-001810–10-05)
- SF3B1 stealth siRNAs (Invitrogen, #HSS146413 and #HSS146415)
- Stealth RNAi™ siRNA Negative Control, Med GC (Invitrogen #:12 935 300)

Reagents:

Compounds

Compounds used are: SML-1/ Spliceostatin J(1), SML-2/ Spliceostatin C (2), SML-3/FR901464 (3), SML-4 (4), Pladienolide B/PlaB (5), Herboxidiene (6), SRPIN340 (7), Amiloride (8), TG003 (9), and Tautomycin (10)

- Methods for making SML-1/Spliceostatin J (1): In a typical approach, strain FERM BP-3421 was fermented at 100 L scale in a mannitol-based medium (consisting of 20 g/L mannitol, 2.5 g/L yeast extract, 5 g/L ammonium sulfate, 4 g/L potassium chloride, 0.2 g/L dipotassiumhydrogenphosphate, adjusted to pH 7.4 prior to

sterilization). After cultivation (7 days at 24°C), the resulting broth was extracted with 120 L ethyl acetate and the dried organic phase (315 g) distributed between methanol/water 9:1 and cyclohexane. The methanolic phase was evaporated and the thus obtained defatted extract subjected to reversed-phase chromatography (YMC RP18 ODSA SP, gradient: methanol/water 40:60 → 75:25). Compound 1 enriched fractions were further purified using RP-18 chromatography (Merck Lichrosphere RP-18, gradient: acetonitrile/water (+0.03% formic acid) 35:65 → 45:55) to yield 3.1 g of pure SML-1/Spliceostatin J (Supplementary Data 1).

- Screening collection: 96 000 compounds of the Novartis collection were selected for the screen and preplated at either 2 mM or 5 mM in 90% DMSO / 10% H2O. These comprised an unbiased set of ca 67k compounds selected based on chemical attractiveness and maximising chemical diversity, and a biased set of ca 15k Natural products, ca 4.7k compounds with previous relation to RNA metabolism, 3k peptidomimetics, 2k compounds not hitting in any other screen ("dark matter"), 960 compounds with previous relation to known ALS related genes, and ca 3k compounds with known mode of action.
- Commercial compounds were obtained as follows: Pladienolide B (Santa Cruz, cat # sc-391691), Herboxidiene (Cfm Oskar Tropitzsch), Amiloride (Sigma Aldrich, cat # A7410-1G), tautomycin (Biovision, cat # B1237-10).

Antibodies

SRSF1 (Invitrogen #32–4500, 1:250 for ICC and 1:2000 for WB); SF3B1 (Abcam, ab170854, 1:100 for ICC and 1:1000 for WB); SF3B1 IP (LSBio, (clone 16) LS-C179473, 10 ug per IP), hnRNPH (Abcam 154 894, 1:5000 WB); β-Actin (8H10D10) (Cell Signaling, #3700 WB 1:1000); Laminin B1 (Abcam 133 741 (1:5000)

Generation of iPS cells

Reprogramming of all lines were done by using Sendai viruses expressing Oct3/4, Sox2, Klf4, c-Myc using CytoTune-iPS Reprogramming Kit (Invitrogen) according to kit instructions. After 3–4 weeks, colonies with stem cell morphology were selected and manually passaged more than five times. After establishment, iPSCs were transferred into feeder free conditions on matrigel (Corning) and maintained in mTeSR1 medium (Stem Cell Technologies) or in NutriStem hESC XF medium (Biological Industries cat-# 05–100-1A) with Pen/Strep supplement (Gibco). Cells were dissociated with TrypLE (Gibco) every 3–4 days, plated at the density of 10 000/cm^2 for Nutristem or 20 000/cm^2 for mTeSR in the presence of 10 μM ROCK inhibitor. The medium was replaced every day (mTeSR1) or every other day (Nutristem).

The absence of reprogramming vectors was confirmed using RT-PCR for Sendai virus expressed Oct3/4, Sox2, Klf4 and c-Myc as described in CytoTune-iPS Reprogramming Kit (Invitrogen). IPSC clones were analyzed for expression of pluripotency markers by scorecard assay (Invitrogen) and for surface markers SSEA-3, TRA1-81 and SSEA-1 (negative marker) using fluorochrome conjugated antibodies (BD Biosciences) and fluorescence activated cell sorting using standard procedures. Karyotype analyses was performed by full-genome SNP analyses by Life&Brain Gmbh (Bonn). All lines showed a normal karyotype except of ALS2 in which an isochromosome 20 could be detected.

Generation and differentiation of Ngn2 neurons

Approx 1 × 10^6 iPS cells were nucleofected by Amaxa nuclefector device using Human Stem Cell Nucleofector® Kit 1 (Lonza #VPH-5012) and Prg#B-016 with 4 ug of Ngn2 plasmid and 1 ug of the dual helper plasmid. Afterward cells were replated on matrigel plates with NutriStem medium containing 10μM of Rock inhibitor. Antibiotic selection (G418:0.1mg/ml) was applied after 48 h. Stable clones appear within 1 week.

For differentiation into iNeurons, 1 × 10^6 of iPS cells were plated on a 6 cm matrigel plate in proliferation medium (DMEM/F12 with Glutamax supplemented with 2% B27 (ThermoFisher, cat-# 17504–044) and 1% N2 (ThermoFisher, cat-# 17502–048), 10 ng/ml hEGF (ThermoFisher, cat-# PHG0315), 10 ng/ml hFGF (ThermoFisher cat-# CTP0263), 1% Pen/Strep (ThermoFisher cat-# 15070–063) containing Rock inhibitor (10μM) for 1d and Doxycycline (1μg/ml) for 3 days. Three days later, induced neurons were frozen or given to differentiation medium (Neurobasal supplemented with 2% B27, 1% N2, Pen/Strep, 1 mM Sodium Pyruvate (ThermoFisher cat-# #11360–039) and the following growth factors at 10 ng/ml BDNF (Cat.#450–02), GDNF (Cat #450–10), hNT3 (Cat.#450–03) (all from PeproTech).

RNA FISH

G4C2-RNA FISH was done accordingly: cells were grown on μClear 384-well plates (Greiner #781 091) or ibiTreat, u-Plate 96-Wel (Ibidi #89 626). Cells were fixed for 60 min with 4% PFA and after washing permeabilized with 0.2% Triton-X (Sigma #T8787) in phosphate-buffered saline (PBS) for 15 min at RT. Pre-hybridization was done for 30 s at 67°C Perfect Hyb buffer (H7033-125ML). FISH probe (Exiqon/Qiagen, design ID 212097, seq: DigN5/GCCCCGGCCCCG/3DigN Tm(RNA):87°C; 50μM) as added to hybridization buffer (miRCURY LNA microRNA ISH Buffer Set (FFPE), 2x, #9000, Qiagen) at a final concentration of 40 nM. Plate was sealed with PCR tape (3M) and incubated for 3 h at 67°C. After incubation, plate was washed for 5 min at 67°C with 2x SSC (Gibco #15 557 044) + 0.1% Triton-X and 3 × 5 min at 67°C with 0.1 x SSC. Cells were blocked for 1 h with 1% BSA in PBS at RT and primary anti-DIG antibody (Roche, Sheep-anti-DIG, #11 333 089 001) was added 1:400 to AB-Delutent (DAKO, #S2022). Plate was incubated O/N at 4°C and the next day washed 3x with PBS + 0.1% Triton X-100. Secondary antibody was added 1:700 for 2 h at RT and cells were washed 1 × 5 min with PBS + 0.1% Triton X-100, incubated with Hoechst (1:1000) for 10 min and washed again, before imaging.

On average 25 single plane images per 96-well (average 3000 cells per well) with a 20x N/A lens were taken by the Operetta (Perkin/Elmer). RNA foci quantification was done with the Harmony analysis software (version 4.1). Nuclear area was determined by DAPI staining, and the RNA foci were quantified within the nuclear area. RNA foci numbers per cell differ from experiment to experiment, depending on the cell type (iNeurons contain more RNA foci than fibroblasts), passage, the background levels and foci detection sensitivity determined by signal to noise ratio. Signal-to-noise ratio differs from FISH to FISH.

Electrophysiology

All the experiments were performed at room temperature. Cells were superfused with HEPES buffer saline solution containing (in mM): 140 NaCl, 5 KCl, 10 HEPES, 10 Glucose, 1 $MgCl_2$, 2 $CaCl_2$, pH 7.4. Patch electrodes had resistances of 2–4 MΩ, and were filled with solution containing (in mM): 140 K Gluconate, 5 HEPES, 2 $MgCl_2$, 1 EGTA, 2 Na_2ATP, 3 Na_3GTP, pH7.7, 288 mOsm. Electrical signals were recorded by a MultiClamp 700B amplifier, a Digidata 1440A converter and pCLAMP 10 software (Axon). In current clamp configuration, all firing curves were obtained with incremental current steps of 10 pA from -30 pA to + 100 pA. Current pulses were 1 s in duration with a 10 s interpulse interval.

Primary screen

iNeurons cells from ALS80/5a Plates were expanded at the iPS stage to generate one large enough batch for both the primary screen and validation, and differentiation was induced by addition of Dox for 3 days in proliferation medium prior to cryopreservation. Plates were coated with PDL–laminin prior to seeding of iNeurons. The plates were kept in the fridge for 6–7 h and then at 37°C O/N prior to cell seeding. Around 2200–2700 cells/well were seeded into 1536 well plates at day 3 after Ngn2 induction. After 5 days in differentiation, cells received compounds to a final concentration of 8 μM and 0.5% DMSO and were further incubated for 48 h at 37°C, 5% CO2. Cells were fixed with PFA (4% final conc), and FISH was performed according to protocol (see FISH section). Five fields per 1536-well were taken with a 20x N/A lens were taken by the Opera (Perkin/Elmer # HH10940200). Image analysis was done with AcapellaTM image analysis software. Nuclear area was determined by Draq5 staining, RNA foci were quantified after setting thresholds for size and intensity within the nuclear area. Plates were normalized using a plate-based normalization module (to the mean of the whole plate) both for cell and total spot count. Spots (sum) per Nuclei (sum) was calculated as ratio of total spot count/cell count in order to normalize for cell numbers. In primary screening, pattern correction was applied using RLOGREG module to account for plate pattern effects. Validation was performed essentially under the same conditions, without plate pattern correction and except for plate coating which was done using 4°C cold Matrigel (1:75) in Neurobasal medium. Each compound was tested in eight concentrations from 0.001 to 8 uM and four replicates each.

Counterassay for transcription/translation inhibitors

Two related reporter assays were used with hPEST destabilised Luciferase either transiently transfected into H23 human lung adenocarcinoma cells (male origin) or stably integrated into DMS-273 human small cell carcinoma cells (female origin). Both systems contain PEST-sequence destabilized luciferases under a CMV promoter and have a protein half-life of 60 min thus being able to detect acute suppression or inhibition of gene expression and to differentiate from general toxicity. Luc2P luciferase (Promega) was cloned into the pDONR221 vector and then subcloned into the M19 destination vector and transfected into NCI-H23 cells (lung adenocarcinoma cells from a 51 year old healthy man). NLucP is a pF5AgNanoLucP construct from Promega which was transfected into DMS-273 cells (small cell carcinoma cells from a 50 year old healthy woman) using Fugene 6 for 48 h and then selected on 1 mg/mL G418 for 7–10 days. Luciferase was detected with NanoGlo and SteadyGlo (Promega) in NLucP cells and Luc2P cells, respectively, and measured on the ViewLux Multilabel Plate imager (Perkin Elmer). Cells were seeded, incubated overnight, and then treated with the compounds at different concentrations for 6 h (0.5% DMSO) prior to luciferase detection. To rule out biochemical inhibition of Luciferase activity a short-term compound incubation of 15–20 min was performed in addition.

Classification of validated hits MoA annotation for the Novartis compounds was done according to M. Schirle and J. L. Jenkins (2015) [72].

RNA isolation and RT-PCR

RNA from fibroblasts or iNeurons was isolated according to the Trizol (Invitrogen, #15 596 026) RNA isolation protocol. RNA was treated with DNAseI from the Turbo DNA free kit (Invitrogen, cat # AM1907) for 15′ at 37°C and inactivated by heating for 10 min at 75°C. DNA. DNAse was removed, either by another Trizol RNA purification, or addition of the inactivating slurry (cat # AM1907). RNA concentrations were determined with Nanodrop (2000/2000c. Thermofisher). Per well/qPCR reaction, 5 pg/mL of RNA per sample was added and RT-qPCR reaction was done according to instructions of the QuantiTect Multiplex RT-PCR Kit (Qiagen, # 204 643). In this kit, cDNA is generated from RNA using random primers, following a qPCR reaction. qPCR protocol used:

STEP	TIME	TEMP.
Reverse transcription	20 min	50°C
PCR initial activation step	15 min	95°C
Two-step cycling		
Denaturation	30 s	94°C
Annealing/extension	30 s	60°C

Hybridization probes (Taqmans) spanning either the exon 2–3 junction of C9ORF72 (Hs00376619_m1, Thermofisher Scientific) or the intron after exon 1b (custom design ID: AICSWFV, Thermofisher Scientific) were used, carrying a FAM fluorescent label. 18S ribosomal RNA hybridization probes with a VIC label were used as an internal normalization control (Hs99999901_s1, Thermofisher Scientific). Within each well of the qPCR reaction plate, C9ORF72 and 18S RNA Ct values were determined. ΔCt values were calculated by subtracting Ct of 18S from Ct values of C9ORF72 ($Ct_{C9ORF72}$-Ct_{18S}). Then ΔΔCt was calculated by subtracting $\Delta Ct_{\text{treatment}}$-$\Delta Ct_{\text{control/DMSO}}$ and the fold up- or downregulation was calculated with $2^{-\Delta\Delta Ct}$.

RNA-protein binding experiments

Recombinant SRSF1 was prepared essentially as in Samatanga *et al.* (2017) [73]. Briefly, SRSF1 RRM1 + 2 ORF corresponding to amino acids 1 to 196 was previously cloned in the pET24 expression vector [74]. A GB1 tag was fused at the N-terminal extremity of the protein to increase its solubility and stability [74]. The protein was overexpressed for 3 h at 37°C in E. coli BL21 (DE3) codon plus cells in LB medium. Protein was purified by two successive nickel affinity chromatography (QIAGEN) steps using an *N*-terminal 6xHis tag cloned between GB1 and the N-terminal extremity of the protein, dialysed against 20mM NaHPO4 pH7, 50 mM L-Arg, 50 mM

L-Glu, 0.05% β-mercaptoethanol buffer and concentrated to 0.1 mM with a 10-kDa molecular mass cutoff Centricon device (Vivascience). Recombinant hnRNPH was obtained from Mybiosource (lo t# Y311415).

Binding of SRSF1 or hnRNP-H to 5′TMR labeled RNA was measured by determination of the fluorescence anisotropy with 2D-FIDA. Briefly, the 5′TMR labeled RNA was thermally denatured for 2 min at 80 °C in assay buffer (20mM NaHPO4 pH7, 100 mM L-Glutamine, 0.05% beta-mercaptoethanol), refolded by cooling to room temperature, and diluted to 0.5 nM; this ensured an average of <1 fluorescent particles in the confocal volume in our setup. Fluorescently labeled RNA was incubated with different concentrations of recombinant SRSF1 or hnRNPH for at least 15 min at room temperature. Measurements were performed in 96-well glass-bottom microtiter plates (Whatman) at ambient temperature (constant at 23.5 °C) on a Clarina2 reader (Perkin Elmer), which is an Olympus IX70 based instrument equipped with two fluorescence detectors, a polarization beam splitter in the fluorescence emission path, and an additional linear polarization filter in the excitation path. A HeNe laser ($\lambda = 543$ nm, laser power = 495 μW) was used for fluorescence excitation. The excitation laser light was blocked from the optical detection path by an interference barrier filter with optical density (OD) = 5. TMR in assay buffer (at 0.5 nM) was used for the adjustment of the confocal pinhole (70 μm) and for the determination of the G-factor of the instrument. The molecular brightness q was extracted from the 2D-FIDA raw data for each polarization channel by using the FIDA algorithm. The G-factor (calculated by using $P_{(\text{true})}$ TMR = 0.034) was determined after every 11 measurements. Competition titrations were performed in presence of constant concentrations of unlabeled (G4C2)4 or (A4U2)4 RNA that was prefolded in presence or absence of 200 uM KCl. Error bars indicated Stdev from 10 consecutive FIDA measurements of the same sample, and data are representative of two independent experiments.

CD spectroscopy

Oligonucleotides were diluted to a concentration of 5 μM in different buffers as indicated, heated to 95°C and refolded at room temperature. Samples were measured on a Jasco J-815, with N2 purging (5 L/min) using glass cuvettes of 0.1 mm path length. Sample measurement was done in 20 mM NaP pH7, 100mM NaCl and or 100 mM KCl, 100 mM L-Arg, 5mM DTT. An average of three CD scans, over the wavelength range of 320–180 nm, was acquired at a scan rate of 20 nm min^{-1} with an 8s response time. CD spectra were corrected for buffer contributions. Data were collected at 25°C.

RNA-Seq

RNA sequencing libraries were prepared using the Illumina TruSeq RNA Library Prep Kit v2 following the manufacturer's instructions. Each library was sequenced in paired-end mode, 2 × 76 bp, using the HiSeq2500 platform. Number of reads per sample are between 67 million (ALS-P009-SML-4) and 28 million (FTD_PlaB). Reads were mapped to the hg38 genome and the transcripts from Ensembl version 76 by using an in-house gene and exon quantification pipeline1. On average, between 96–76% of the total reads (depending on the library) were mapped to the genome or the transcripts, and 88–68% of the aligned reads mapped to expressed sequences. The genome and the transcript alignments were used to derive gene counts based on the human Ensembl gene IDs (v38). Gene counts, which represent the total number of reads aligned to each gene, were then transformed into counts per million (CPM; normalization by total number of mapped reads per sample) and fragments per kilobase of exon per million of fragments mapped (FPKM; further normalization by effective gene length). After assessment by principal component analysis and multidimensional scaling, all samples were retained for further analysis. Only genes with counts above 1 CPM in at least three samples were inculded. Differential expression analysis was performed on the CPMs using a limma/voom workflow with R [75]. The following test was performed: Treatment versus control using the donor as a blocking factor. Results are reported in terms of log2 fold changes and negative log10 adjusted P values (Benjamini Hochberg false discovery rate).

Tet-off assay

HeLa cells were seeded on a 96-well Ibidi plate (25.000 per well) and the next day transfected with 0.5 μl/well Lipofectamin 2000. A total of 200 ng plasmid DNA was added per well; 160 ng/well of the 152 repeat containing plasmid (TRE3G-152x-emGFP) and 40 ng/well of the TetOff protein vector (EF1a-TetOFF3G). 15 h after Lipofectamin transfection, the compounds were added (Doxycycline 1 μg/ml, PlaB (50 nM) or both together) for different time points. After 6 h, cells were fixed with 4% PFA and the RNA foci were checked by FISH.

RNA bio-IP

The RNA oligo IP was done according Haeusler *et al.* (2014) [76]. Briefly, 50 μg of total nuclear protein was used. After N/C fractionation, cells were lysed in 10 mM HEPES, pH 7.0, with 200 mM KCl, 1% Triton X-100, and 10 mM MgCl2 and 1 mg/ml heparin. Protein concentrations were measured Pierce BCA protein assay kit (catalog #: 23 225). RNA oligo's were folded in different structures, before binding to the beads. For the G4 structure, RNA in annealing buffer (10 mM Tris, pH 7.4, 6 and 100 mM KCl) was heating to 95°C for 10 min and cooling slowly to RT. For a refolded structure, denaturing was done in the absence of KCl. 30 μl of RNA oligo's (100 uM) were bound to 300 μl Dynabeads™ M-280 Streptavidin (catalog #11205D) for 15min at RT in annealing buffer. Beads were washed, added to the protein lysate and incubate for 3–4 h under at 4°C constant rotation. Beads were washed in lysis buffer with increasing concentrations of KCl (0.4M, 0.8 M and 1.6 M) for 10 min each at 4°C. After washing, protein loading dye with DTT was added to the beads and stored at −20°C. The next day, beads with loading dye were boiled for 10 min at 70°C before loading it to a 4–12% BisTris gradient gel.

Mass spectrometry

Affinity-purified samples were separated by SDS-PAGE as described above. Subsequently, gel lanes were excised in 16 equally sized gel slices and processed for mass spectrometry by in-gel digestion with a Trypsin/Lys-C mixture (Promega). Eluates from the gel slices were loaded on a 2 cm x 100 u trap column (Thermo Easycolumn) and separated by liquid chromatography on a 15 cm x 75 Easy-Spray column (ThermoFisher), coupled to a Q-Exactive mass spectrometer (ThermoFisher). Peaklists from the mass spec runs were searched using Mascot 2.4 (Matrix Science) against a Uniprot human

database (version 2014). Search results were summarized and analyzed using Scaffold (v. 4.6.4, Proteome Software). Spectral count was used as a proxy for protein abundance.

SILAC

SILAC HeLa cells were grown in DMEM for SILAC (ThermoFisher # 88 364) with 10% dialyzed FBS (ThermoFisher #26 400 044), 1% P/S and supplemented with heavy ($^{13}C/^{15}N$) and light ($^{12}C/^{14}N$) labeled L-Arg and L-Lys ThermoFisher # A33972). Cells divided at least 10 times to allow for complete change to heavy labelled proteome. Light HeLa cells were treated for 16 h with 25 nM of PlaB). Heavy labeled HeLas were treated with same amount of DMSO for 16 h. After treatment, cell numbers were measured three times with a countess cell counter (Invitrogen) and $20*10^6$ cells of heavy and light cells were mixed. Nuclear/cytoplasmic fractionation was performed (see below methods) and protein concentration was measured using BCA method. 0.85 mg of nuclear extract for the SRSF1 (Invitrogen #32–4500) and SF3B1 (LSBio, LS-C179473) IP was used. 10 mg of cytoplamsic extract was used for the SRSF1 cytoplasmic IP and 10 ug antibody was coupled to 100 ul proteinG Dynabeads (Invitrogen #10009D). Lysate/beads were incubated for 40′ at RT and washed 3 × 10′ at RT with IP buffer, and 2 × 10′ with LWB (co-IP kit, Invitrogen #14321D). SILAC samples were analyzed using MaxQuant v 1.5.8.3 [77] using the same database.

Nuclear/cytoplasmic fractionation was done according to the protocol of Wang *et al.* (2006) [78]. Lysis process was checked by microscopy and nuclei were sonicated (3 × 10'' on ice) in lysis buffer (10 mM HEPES, pH 7.0, with 200 mM KCl, 1% Triton X-100, and 10 mM MgCl2 and 1 mg/ml heparin) and incubated for 10′ before spinning debris down. Supernatant was used for IPs. For cytoplasmic fractions, only the first two RSB40 fractions were pooled, and the last wash was discarded. For RNA isolation, intact nuclei were dissociated in Trizol and RNA was isolated accordingly.

Knockdown experiments

ALS patient fibroblasts were seeded on 96-well Ibidi plates (3000–5000 per well) and the next day, siRNAs were incubated with RNAiMax lipofectamin (0.3 ul/well) and added to the cells. siRNAs targeting SF3B1 (10 nM final concentration) were added and 48 h later replaced by fresh SF3B1 siRNAs in fresh media. After total incubation of 4 days, cells were fixed in 4% PFA. siRNAs targeting SRSF1 and NXF1 (50 nM final) were added once and incubated for 72 h.

Immunocytochemistry for RAN-DPRs

HEK293T cells were grown in 96-well Greiner microclear plates and fixed with 4% PFA after the indicated treatments. To visualize DPRs, cells were permeabilized with 0.5% Triton in PBS for 30min and blocked with a 5% normal goat serum in PBS solution. Primary antibodies were incubated for 3 h or overnight at 4°C in blocking solution: α-Myc tag (Thermo, MA1-980, 1:250); α-HA tag (CellSignaling, #3724, 1:1600); α-Flag tag (Sigma, F3165, 1:400); α-polyGA (Millipore, MABN889, 1:500); α-polyGP (Proteintech, 24494–1-AP, 1:250). The following day, cells were washed with PBS, incubated with fluorescently-conjugated secondary antibodies (Thermo) for 2h at RT. Cells were then washed, stained with DAPI to visualize nuclei, and plates were sealed until imaging. Images in suppemental figures were acquired in high content settings with an InCell6500 (GE Healthcare), using a 40x objective or with an Olympus IX83 equipped with spectral filters for DAPI, GFP, Cy3, Cy5 and Cy7 and a 40x objective (UPLXAPO40X NA 0.95). Image analysis of images generated with the Olympus IX83 was done using the EV-Analyzer pipeline for cell detection [79]. Images in the main figure were aquired with Zeis confocal microscope and DPR quantification was done with the Harmony analysis software (version 4.1).

Cell proliferation and viability measurements

For live-cell tracking of viability, cells were plated at the same density (27′500/well) across all wells in Corning microclear 96-well plates or Ibidi plate. The plates were then placed in an IncuCyte automated live-cell imager (Essen Bioscience), installed inside an incubator. Images were acquired every 3 h prior to, and after compound treatments. Two fields per well were imaged. The IncuCyte S3 analysis software (Essen Bioscience) was used to quantify the proliferation of live cells across treatments and throughout the experiment. Data were exported to Excel to generate graphs (Supplementary Fig. 6c). After 24 h of experiment, the Ibidi plate was fixed in 4% PFA for immunostaining as described above to determine the subcellular presence of DPRs.

Plasmid construction and stable cell line construction for stable smFISH HEK cell lines

24x cassettes each of PP7 and MS2 stem loops were cloned into the piggybac plasmid containing the 152x-G4C2 repeats placed under a TRE3G promoter using ligation free cloning using NEBuilder (NEB). 2 μg of the cloned construct along with 2 μg plasmid coding for transposase were transfected into HEK-293 Tet-off cells (Takara) using Lipofectamine 2000 (Invitrogen). Transfected cells were selected for ten days with 1.5 μg/mL puromycin (Invivogen). Individual colonies were isolate, grown, and were tested for integration using colony PCR and restriction digestion.

Single-molecule FISH and image analysis

HEK-293 Tet-off cells were induced upon splitting the cells in Doxycycline free DMEM supplemented with 10% FBS for 72 h. 4 × 104 cells were seeded on Laminin-Poly-lysine coated glass coverslips placed in a 12-well tissue culture plate two days before smFISH was performed. Fresh DMEM + 10% FBS containing Doxycycline (1 μg/mL) was added for 45 min to shut off transcription. For the Pladienolide B treated cells, Pladienolide B (25 nM) was added to the cells for 90 min and were also induced for 45 min by Dox. In addition to the above-mentioned coverslip, a coverslip containing the non-induced cells was included as control. Subsequently, cells were fixed using 4% paraformaldehyde (Electron Microscopy Sciences) in 1x PBS for 10 min. Following fixation cells were washed twice with 1x PBS before permeabilizing using 70% ethanol at 4°C. Subsequently, the cells on coverslips were prehybridized in wash buffer (2x SSC (Invitrogen), 10% (v/v) formamide (Ambion)) twice for 5 min. FISH probes against the PP7 stem loops in the 3′ UTR (Quasar 570) were then hybridized with a solution that contained 125 nM of each FISH probes, 2x SSC, 10% (v/v) formamide and 10% (w/v) dextran sulfate (Sigma) at 37°C overnight. Cells were washed twice with above-mentioned wash buffer for 30 min each. Thereafter, the cells were rinsed twice with PBS before mounting on slides using ProLong Gold Antifade Mountant with DAPI for counterstaining DNA (Molecular Probes).

Images were acquired using a wide-field microscope (Zeiss) with a Plan-APOCHROMAT 100 × 1.4 NA oil objective equipped with an AxioCam 506 mono camera and an X-Cite 120 EXFO metal halide light source. Optically sectioning of 240 nm z-step was employed, spanning a z-depth of 5 μm. Exposure times of 1600 ms was used to acquire images of each plane in the channel Quasar 570, and 40 ms in the DAPI channel. For analysis, Z-stack images were Z-projected using maximum intensity and the background in the Quasar 570 channel was subtracted using rolling-ball background subtraction in Fiji software (2). Subsequently, the region of interest containing individual cells were selected and the total number of RNAs in each cell was estimated using spot detection algorithm of Trackmate plugin in Fiji (1). The nucleus of each cell that were previously selected for analysis was segmented using the DAPI channel and was used as the region of interest to estimate the number of RNAs in each nucleus. The number of cytoplasmic RNAs in each cell were calculated by subtracting the total number of nuclear RNAs from the total number of cellular RNAs in that cells. The cytoplasmic fraction of RNAs was calculated by dividing the total cytoplasmic RNAs in each cell by the corresponding total cellular RNAs. The final results were visualized in a boxplot using R (3).

Differentiation of iPS cells to motor neurons

The iPS cell lines used in this study were previously described in Dafinca *et al.* (2020) [70]. The iPSCs were differentiated to motor neurons according to previously published methods [70, 80].

Immunostaining on motor neurons

Motor neurons were fixed with 4% paraformaldehyde-PBS for 15 min and incubated with 10% donkey serum in Phosphate-buffered Saline (PBS) with 0.2% Triton-X for 1 h at room temperature. The cells were incubated overnight at 4oC with rabbit anti-cleaved caspase-3 (Cell Signalling, 1/500), mouse anti-?- III Tubulin (Tuj1) (Covance, 1:1000), or rabbit anti-PABP (Abcam, 1:1000. The cells were washed with 0.1% TritonX/PBS three times for 10 min and incubated with Alexa Fluor 488, Alexa Fluor 568 or Alexa Fluor 637 conjugated donkey anti-rabbit and anti-mouse secondary antibodies (Life Technologies) for 1 h at room temperature. Nuclei were stained with DAPI. Fluorescence was visualized using a Spinning Disc Confocal Microscope Zeiss.

RNA FISH on iPSC-derived MNs

RNA-FISH was performed on iPSC-derived MNs using an adapted, previously published protocol [9, 81]. Cells were fixed with 4% Paraformaldehyde for 20 min and then permeabilised using 0.1% Triton-X in PBS for 20 min. After three 5 min-washing steps with PBS, the coverslips were dehydrated by the serial application of 70%, 80% and 100% ethanol, for 1 min each, and then airdried. The cells were then rehydrated using PBS and briefly washed using 2 x saline-sodium citrate (SSC). The coverslips were then treated with pre-hybridisation solution, consisting of 50% Formamide and 2 x SSC, and incubated at 80oC in a hybridisation oven for 30 min. The pre-hybridisation solution was then replaced with hybridisation solution containing 2xSSC, 0.16% BSA, 0.8 mg/mL salmon sperm, 0.8 mg/mL tRNA, 50% Formamide, 8% dextran sulphate, 1.6 mM vanadyl ribonucleoside, 5 mM EDTA and 0.2 ng/μL probe 5Cy3-GGCCCC GGCCCC GGCCCC GGCCCC), and hybridization was performed in a hybridisation oven at 80oC for 2 h. The coverslips were then washed 3x with 50% formamide/0.5xSSC wash solution at 80oC, for 30 min each. Hereafter, the cells were washed 3x with 0.5 x SSC at RT, for 10 min each. After another washing step with PBS, the cells were blocked with 10% donkey serum in PBS/0.1%Triton-X for 30 min at RT, and then incubated with a mouse anti-TUJ1 antibody in PBS/0.1%Triton-X (1:500, BioLegend) at 4oC ON. Immunostaining and imaging were continued as described above.

Statistical information

All error bars represent SDs, as noted in the figure legends. Significance of most figures was determined by paired One-Way Anova, and three stars (****) indicates $P < 0.0001$. Significance of Fig. 3E was determined by t-test, and three stars (***) indicates $P < 0.0001$. Significance was calculated with GraphPad Prism (v 8.12). For RNA foci and RT-qPCR with a biological replicate number of $n > 3$ in rare occasions, data points which were clear outliers due to technical artefacts were excluded from the statistical analysis, provided that the data were reproduced by at least three independent biological experiments. For most RNA foci and DPR analysis, each datapoint is the average of ∼25 images per well, containing on average 3000 cells per well, or otherwise noted in figure legends.

Code availability

No custom code availability has been used in this study.

Ethics approval and consent to participate

All primary cells from patient and healthy donors were obtained under full Ethical/Institutional Review Board approval and patient consent as specified in the "Material and methods" section.

Results

Large scale (96 200) small molecule screen for G4C2-RNA foci disassembly in C9ORF72 iPS neurons identifies spliceosome targeting compounds

To screen for small molecules interfering with the metabolism of G4C2-RNA foci in a relevant cellular context, we generated iPS derived neurons from ALS and FTD patient cells carrying the C9ORF72 repeat expansion (750–1100 repeats; Supplementary Table S1) by insertion of a Doxycycline (Dox) inducible Neurogenin-2 (Ngn2) expression construct (iNeurons) [82] (Fig. 1A). We confirmed neuronal morphology by expression of the neuronal marker Tuj1 and the ability to fire action potentials (Fig. 1B and Supplementary Fig. S1A). Southern blotting confirmed that the repeat length was maintained during reprogramming (Supplementary Fig. S1C). Consistently, FISH (fluorescent *in situ* hybridization) demonstrated that G4C2-RNA foci were maintained after reprogramming and differentiation (Fig. 1C and Supplementary Fig. S1B), whereas a high variability of RNA foci levels was observed between donors and even between iPS clones and subclones from the same donor, with no apparent correlation with repeat lengths (Fig. 1D and Supplementary Fig. S1B). Consistent with previous reports [3, 24], ALS lines showed slightly lower expression of C9ORF72 mRNA but increased RNA levels of the repeat containing intron compared to HD

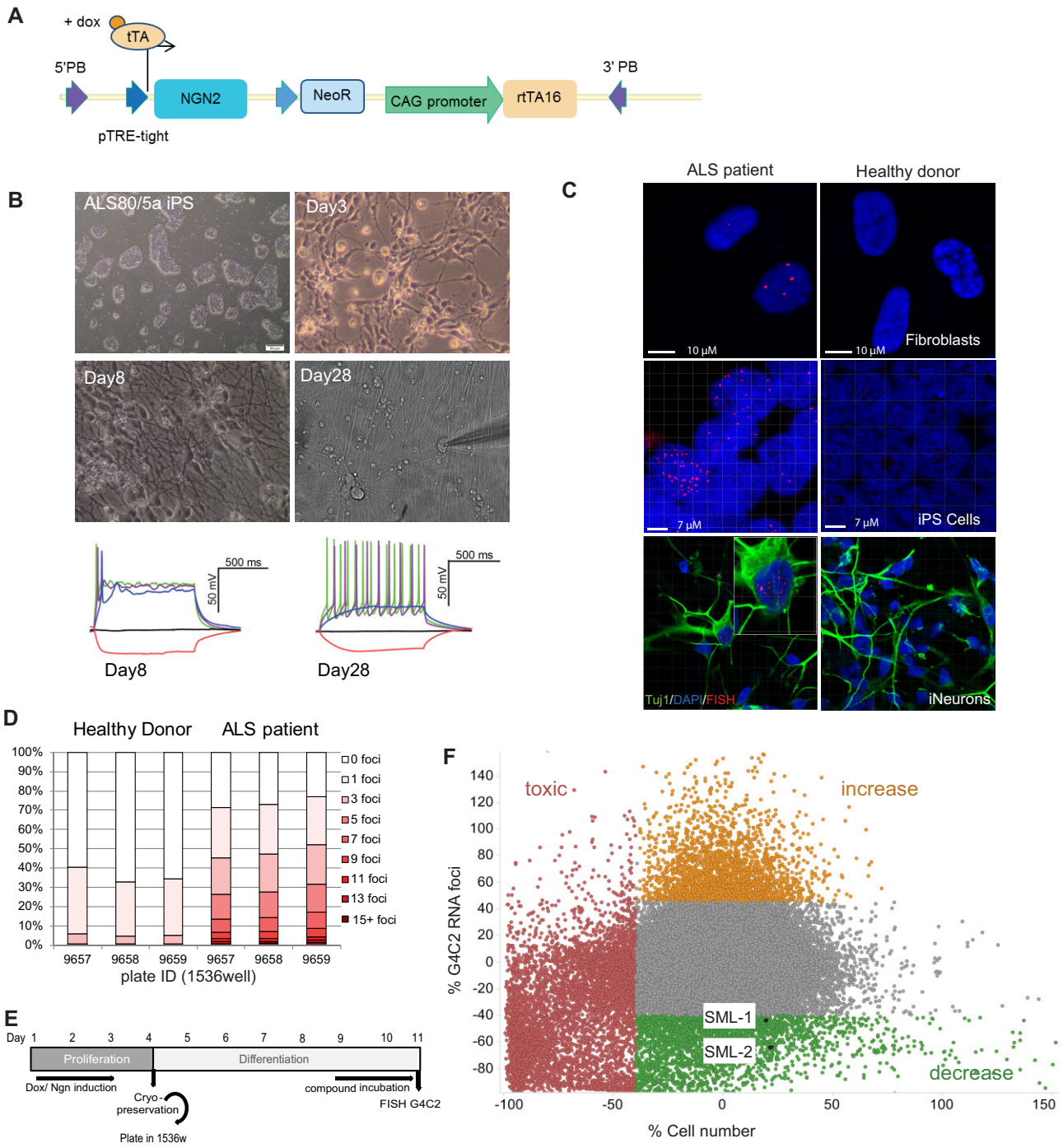


Figure 1. High-throughput screen of 96 000 small molecules for RNA foci clearance in C9orf72 iNeurons. (**A**) Construct used for stable integration of inducible Ngn2 expression in HD and ALS iPS cells, adapted from Zhang *et al.* (2013) [72]. 5′PB and 3′PB: flanking arms for PiggyBac transposase genome integration. Ngn2: open reading frame of human Neurogenin2. Dox: Doxycyclin. NeoR: Neomycin resistance gene. The reverse tetracycline-controlled transactivator (rTA) is encoded on the same construct (rTA16: rTetR fused to VP16) and expressed from the CAG (CMV early enhancer/chicken β actin) promoter. (**B**) Dox induced expression of stably integrated, transgenic Ngn2 in C9orf72 ALS iPS cells results in neuronal morphology and electrical activity after 8 days and repetitive action potential firing after 28 days. Representative data for one ALS line are shown, data for HD line are shown in Supplementary Fig. S1A. (**C**) G4C2-RNA foci are retained in C9orf72 ALS patient iNeurons as detected by FISH. Nucleus stain: DAPI. Left panels, ALS patient cells. Right panels, healthy donor lines (Supplementary Table S1) (**D**) Quantification of RNA foci in 1536-well format. Single cell histograms of RNA foci per cell are shown for healthy donor (hDF90) and ALS (ALS80/5a) iNeurons. Each bar represents pooled data for nine wells on the same plate. (**E**) Screening paradigm. iPS cells were expanded and Ngn2 expression was induced for 3 days with Dox prior to cryopreservation. For the HTS, iNeurons were plated into 1536-well plates for additional 5 days and incubated with compounds for 48 h. RNA foci were visualized by FISH and quantified by automated imaging. (**F**) Primary screening data. Each dot represents data for one compound at 8 μM and shows % change in RNA foci and cell number as compared to DMSO controls. SML (splice modulator like) compounds are highlighted within the primary screening data.

(healthy donor) cells (Intron 1a, Supplementary Fig. S1D). C9ORF72 mRNA expression was comparable between differentiations, whereas some variability was observed for intronic RNA (Supplementary Fig. S1D).

iPS cells were differentiated to neurons by induction of transgenic Ngn2 expression with Doxycycline. After 3 days, cells were seeded into 1536-well plates in differentiation medium. Compounds were added on Day9 after Ngn2 induction and incubated with the cells for 48 h prior to fixation and RNA foci detection by FISH (Fig. 1E). RNA foci were quantified by plate-based imaging of >1000 cells per well and automated image analysis. For assay development, we used treatment with repeat targeted ASOs as positive controls [23] (data not shown), however due to the required longer incubation time (5–7 days) these were not suitable as internal references for the screen. Instead, we used healthy donor iPS neurons (DMSO treated) to determine the background foci detection in the automated image analysis (low value). Additionally, ALS patient neurons were treated with DMSO only (high value). These reference samples were included in replicates on every plate and used to determine the dynamic window of RNA foci detection and % DMSO inhibition thresholds (Supplementary Fig. S1E and F). A total of 96 200 small molecules from the Novartis compound collection, stratified for maximal chemical diversity and including approximately 3000 compounds with a known mode of action, were screened for RNA foci clearance at a single replicate and a compound concentration of 8 uM. Each data point in Fig. 1F shows relative changes in RNA foci levels and cell numbers for individual compounds. Thresholds for primary hit selection were set at ≥40 % RNA foci increase / decrease and ≤ 40 % reduction in cell number. These thresholds were set based on the variability between all reference wells across the screen (Supplementary Fig. S1E). In total, 2952 compounds were selected for validation in a dose response assay in iNeurons from the same line as in the primary screen, performed in triplicates and concentrations from 0.1 to 10 uM (Supplementary Fig. S2A). During validation, we determined IC50 values for RNA foci elimination as well as cell toxicity (based on nuclear numbers, Supplementary Fig. S2A and B). As typically observed in high-throughput screens, a large fraction of primary hits dropped out at this stage due to either low potencies (IC50 values > 10 uM) and/or cell toxicity (Supplementary Fig. S2A and B) or lack of replication from the primary screen. Hits which did not show a dose response were excluded as well (data not shown). Around 261 of the primary hits were finally confirmed to modulate RNA foci levels in a dose dependent manner and with an IC50 of < 10 uM, comprising 223 compounds decreasing RNA foci (Fig. 2A, right panel) and 38 compounds increasing RNA foci (Fig. 2A, middle panel).

To filter for compounds affecting gene expression in general we used two related reporter assays with hPEST destabilized NanoLuc Luciferase in stably engineered HeLa cells (data not shown). This assay filtered out 122 compounds that unspecifically interfere with gene expression such as by blocking transcription, RNA processing or translation machineries, as well as DNA binders or intercalators, independent of the C9ORF72 repeat. 139 compounds were found to be G4C2-specific and were further tested in C9ORF72 lines from three independent donors, resulting in 82 final validated hits (Fig. 2A). Among these, we identified the Brd4 inhibitor JQ1 as a strongly RNA foci increasing compound (Fig. 2B and C), which is consistent with a previous report [83] and intrinsically provides a validation for the screen. Testing of stereopure isomers of JQ1 for RNA foci clearance (Fig. 2B, Supplementary Fig. S2C and D) confirmed that only the active (+) but not the inactive (-) stereoisomer of JQ1 are active in RNA foci clearance. Consistent with the known epigenetic mechanism, JQ1(+) also upregulated C9ORF72 mRNA, which was again not observed for JQ1(-) (Supplementary Fig. S2E). A number of additional, known epigenetic modulators were found among the validated hits (Fig. 2C). Inhibitors of Brd4, CREBBP, KDM4B, H3K27 HMT and BPRF1 consistently led to an increase in RNA foci, while HDAC inhibitors decreased the number of RNA foci. Chemical substructure clustering (Fig. 2D) revealed that the majority of validated hits were singletons from different scaffolds, in line with the fact that the screening deck was stratified based on maximizing chemical diversity rather than including multiple analogs of the same pharmacophore. Additionally, 39 scaffolds were identified with at least two representatives. Two of these compounds caught our attention in particular; although they were previously undescribed molecules, their structures had high chemical similarity with Spliceostatin A (SSA), a known modulator of the core spliceosome that targets the branch point binding protein SF3B1 (Sap155) of the U2 snRNP [84, 85]. We therefore term these compounds *splice modulator like-1*/**SML-1** (Spliceostatin J) and **SML-2** (Spliceostatin C). The structures are shown in Fig. 2D, the chemical characterization is summarized in Supplementary Data 1.

RNA foci elimination by SML compounds is due to SF3B1 modulation

Since the G4C2 repeat is in an intron, we set out to further investigate the mode of action of these compounds. We selected eight additional close analogs of SML-1 and SML-2 from our compound collection to perform a structure activity relationship analysis (Fig. 3A). Unfortunately, we were not able to obtain SSA itself for this study. The primary screening hits SML-1 and SML-2 were confirmed to clear G4C2-RNA foci within only 6 h in ALS iNeurons (Supplementary Fig. S3A), whereas only slight cell toxicity was observed up to 16 h treatment (Supplementary Fig. S3B). The compounds showed similar RNA foci clearance activity in additional systems, including C9ORF72 primary patient fibroblasts (Supplementary Fig. S3C). Due to practical considerations, we, therefore, decided to use patient fibroblasts rather than neurons and with compound treatment for only 8 h for IC50 determination of all SML analogs. These data revealed a dynamic structure-activity-relationship (SAR) at the epoxide containing heterocycle (Fig. 3A and Supplementary Fig. S3C), which is similar to published SAR data for SSA (Spliceostatin A) in splicing inhibition [86]. From this series, we selected SML-3 (also known in the literature as FR901464 [87] as the most potent derivative with an IC50 of 2 nM. Remarkably, SML-4, a previously undescribed compound which only differs from SML-3 in the opening of the epoxide, was entirely inactive (Fig. 3A and Supplementary Fig. S3C) and was selected as negative control in subsequent studies. Consistent with the fibroblast data, SML-3 was also confirmed to clear RNA foci with high potency (IC50 2 nM) in ALS iNeurons whereas SML-4 was again inactive (Fig. 3B and C).

We next aimed to confirm that the SML series of compounds indeed also targets the spliceosome. For this purpose,

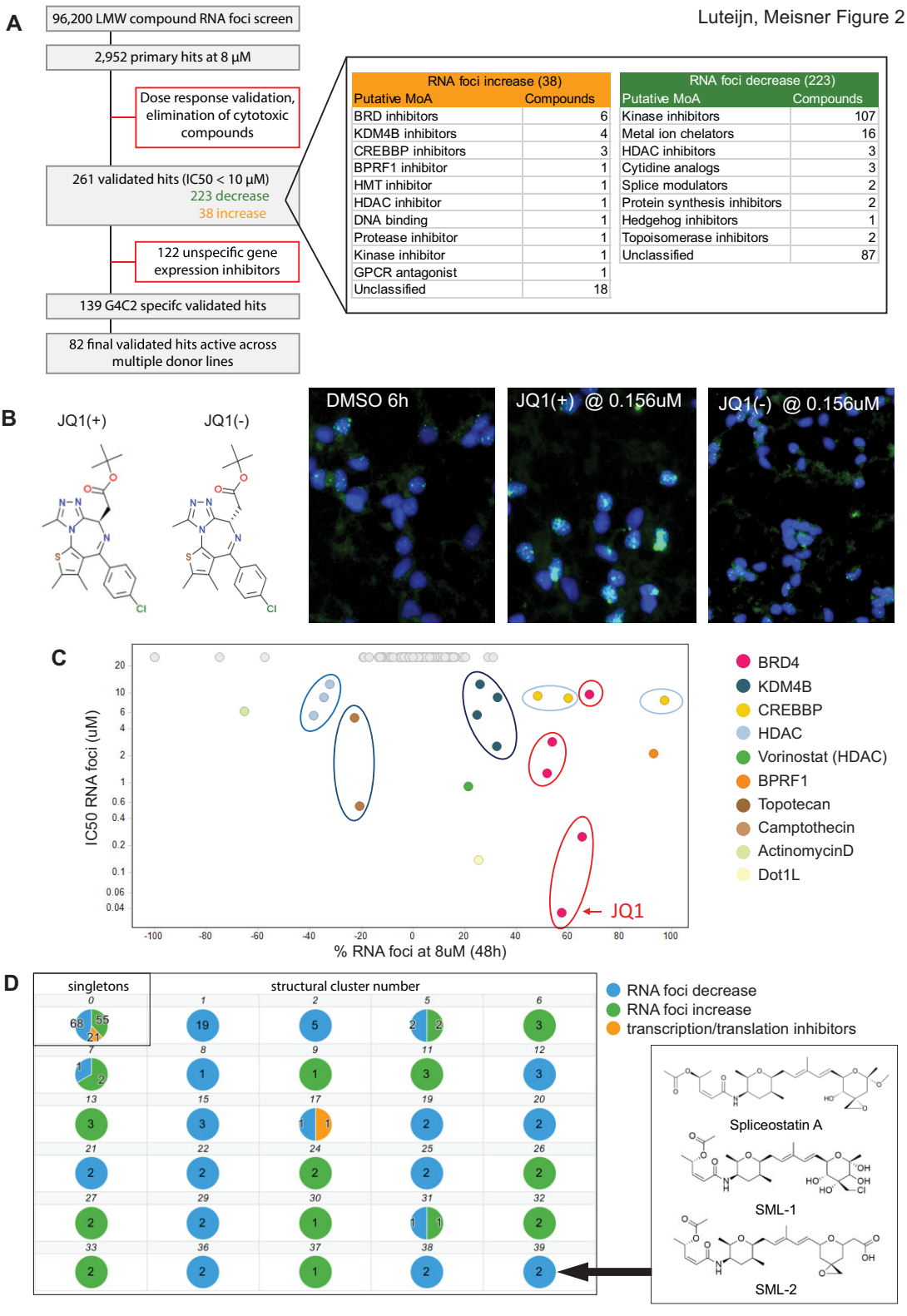


Figure 2. Validation and stratification of primary hits reveals pharmacological RNA foci modulation by epigenetic modulators and novel Spliceostatin A analogs. (**A**) Screening flowchart (left panel) with classification of the 261 validated hits based on known targets and historical data (right panel). (**B**) JQ1(+) results in massive upregulation of both RNA foci as well as C9orf72 mRNA within 6 h, whereas the Brd4 inactive enantiomer JQ1(-) was inactive also in RNA foci clearance, for quantification see Supplementary Fig. S2C–E. (**C**) IC50 and % change of RNA foci for small molecules targeting epigenetic regulatory proteins, as annotated for each of these compounds based on in house data from independent, previous screening and drug discovery campaigns. Among the validated RNA foci increasing hits, the screen led to *de novo* identification of the Brd4 inhibitor JQ1, which already had been previously reported to regulate the C9orf72 locus [73]. (**D**) Clustering of the 261 validated hits by chemical similarity. Each circle represents one scaffold. The number of analogs belonging to the same chemical cluster is shown within the center of each circle, numbers on top indicate arbitrary cluster identities. The cluster number 39, comprising two analogs of the spliceosome modulator Spliceostatin A, SML-1 and SML-2, is highlighted and the structures of the compounds are shown.

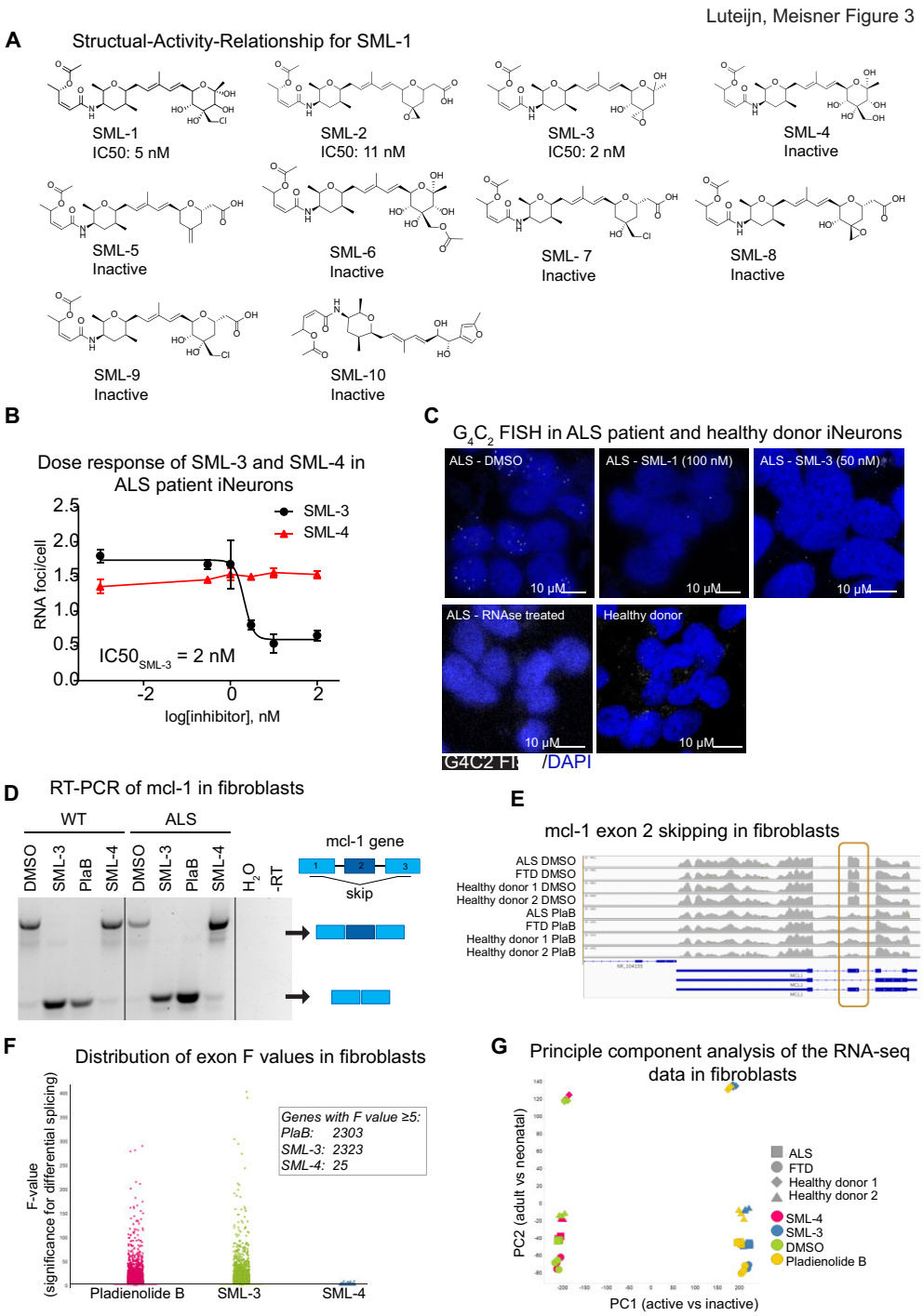


Figure 3. SML compounds phenocopy Pladienolide B in alternative splicing and eliminate RNA foci via SF3B1 modulation. (**A**) SAR (structure activity relationship) for compounds similar to Spliceostatin A (SSA) with IC50 values of RNA foci clearance in C9orf72 fibroblasts (dose response curves in Supplementary Fig. S3c). (**B**) Dose response of SML-3 and SML-4 in RNA foci clearance after 8 h of treatment of ALS iNeurons. Each data point represents an average of three wells, error bars are SDs. (**C**) G4C2 FISH (white) of ALS patient iNeurons treated with 50nM of SML-1 or 100 nM of SML-3. ALS iNeurons treated with RNAse and Healthy donor iNeurons are shown as controls for the FISH. Nuclei stain: DAPI. (**D**) RT-PCR of *mcl-1* on cDNA of fibroblasts treated for 4 h with 25 nM of SML-3, SML-4 and PlaB or equivalent DMSO concentrations, showing alternative splicing of *mcl-1* exon 2. (**E**-**G** RNASeq of ALS and healthy donor (HD) fibroblasts with and without treatment with PlaB or DMSO for 4 h. (**E**) alignment of reads to the *mcl-1* gene, showing exon2 skipping and intron retention upon PlaB treatment, confirming the RT-PCR data in (**D**). (**F**) Transcriptome wide distribution of alternatively spliced RNA transcripts in PlaB or SML-3 treated fibroblasts as determined by RNASeq. *F*-value represents the difference in reads along the transcripts for each exon compared to DMSO treated samples. The higher the *F*-value, the stronger this difference, most likely attributed to alternative splicing. Libraries were generated from triplicates per compound condition. ALS patient samples (ALS G9 fibroblasts) and Healthy Donors (Hdf 90/1a) were used for the limma/voom analysis. (**G**) Principal component analysis of alternatively spliced transcripts, grouping affected intronic sequences by compound and cell line. Every intron was treated as "gene", resulting in 43 572 "genes" after filtering (counts per million (CPM) ≥ 5). The first component separates active from inactive compounds, the second component separates hdf63 from other donors (neonatal versus adult). In both components, SML-3 and PlaB affected introns grouped closely together while SML-4 clustered with DMSO. Additional data of the RNASeq study are shown in Supplementary Fig. S3.

we used SML-3 as the most potent derivative for side by side comparison with known analogs. Since SSA was not available at the time of the study, we used Pladienolide B (PlaB) as a reference compound, which is another well documented modulator of SF3B1 and with the same mode of action as SSA [84, 88–91]. Binding of either SSA or PlaB to SF3B1 does not block but instead modulates the protein's function; upon compound binding, SF3B1 undergoes conformational changes which allow the U2 snRNP to bind to non-canonical splice sites, consequently resulting in intron retention and alternative splicing [89, 92, 93]. Fibroblasts from C9ORF72 ALS patients and healthy donors were treated with SML-3, SML-4, PlaB or DMSO followed by total RNA extraction to measure effects on alternative splicing by RT-PCR and RNA sequencing. An incubation time of only 4 h was used to avoid secondary effects of modulated downstream genes. Treatment of ALS fibroblasts with SML-3 induced skipping of the *mcl-1* exon 2, a well-established alternative splicing event induced by SF3B1 inhibitors [94], while SML-4 showed no effect in both, RT-PCR (Fig. 3D) as well as RNA sequencing (Fig. 3e). At the transcriptome wide level more than 2000 transcripts were alternatively spliced in SML-3 treated cells, as quantified by the *F*-value of the limma/voom differential splice test (Fig. 3F). In addition, largely the same transcripts were affected in PlaB and SML-3 treated cells, shown by the strong correlation of adjusted *P* values in PlaB and SML-3 treated cells (Supplementary Fig. S3D), the large overlap of top-ranking affected transcripts (Supplementary Fig. S3E) and the clustering in a principal component analysis (Fig. 3G). Furthermore, SML-4 was confirmed as a true inactive analog and was nearly identical to DMSO treatment at the transcriptome-wide level (Fig. 3F and G).

As the newly identified compounds SML-1, SML-2 and SML-3 thus not only show high chemical similarity to known SF3B1 modulators but also phenocopied their effects in alternative splicing (demonstrated for SML-3), we next aimed to confirm whether SF3B1 is indeed the efficacy target also for RNA foci modulation by these compounds. We therefore tested whether vice versa, unrelated SF3B1 modulators would also phenocopy SML compounds in RNA foci clearance. Indeed, both PlaB as well as Herboxidiene which are both SF3B1 modulators [95] chemically unrelated to the Spliceostatin and SML scaffolds also cleared RNA foci in fibroblasts (Fig. 4A), iNeurons (Fig. 4B and C) as well as in motor neurons (Fig. 4D and Supplementary Fig. S4A), thereby strongly indicating that SF3B1 modulation is indeed responsible for RNA foci clearance by these compounds.

To further corroborate this, we knocked down SF3B1 by siRNA transfection of ALS fibroblasts (Fig. 4E, left panel). Consistent with the previously described gain of function of SF3B1 in alternative splicing rather than inhibition upon binding to SSA, PlaB or Herboxidiene, reduced SF3B1 expression led to an increase rather a decrease in RNA foci (Fig. 4E, lower left panel for quantification and right panel for images). This indicates that, even in the absence of SML compounds, SF3B1 is involved in intrinsic G4C2 RNA metabolism. Additionally, we treated SF3B1 depleted cells with PlaB, SML-3, SML-4 and DMSO and observed a significant reduction of PlaB activity in ALS patient fibroblasts and a trend for SML-3, supporting the conclusion that these compounds require SF3B1 to be present in order to mediate RNA foci clearance (Supplementary Fig. S4B). Interestingly, and in line with these data, combining FISH and immunocytochemistry revealed that G4C2 RNA foci cluster in the periphery of SF3B1 positive nuclear speckles in ALS neurons (Fig. 4F).

SML compounds induce RNA foci disassembly by a posttranscriptional mechanism, independent of C9ORF72 gene context and splicing

Interestingly, the RNA sequencing data revealed that C9ORF72 itself is one of the targets downregulated upon both, SML-3 and PlaB treatment, which was not observed for SML-4 (Supplementary Fig. S5A). This was further confirmed by quantification of C9ORF72 processed mRNA (exon 2–3 junction) levels by RT-qPCR in ALS neurons (Supplementary Fig. S5B). The C9ORF72 mRNA downregulation was observed in both ALS as well as wild type cells (Supplementary Fig. S5A), demonstrating that SF3B1 modulating compounds downregulate C9ORF72 expression independent of the repeat expansion. Consistently, no general effects of SML-3 and PlaB were observed on the expression levels of endogenous transcripts containing three or moreG4C2 repeats in the RNA-Seq data (Supplementary Fig. S5C).

The most obvious explanation for how SF3B1 modulation might mediate RNA foci clearance would therefore be due to modulation of C9ORF72 pre-mRNA processing and altered splicing of the repeat containing intron. However, we observed no evidence for alternative splicing of the C9ORF72 pre-mRNA induced by SML-3 or PlaB in ALS or healthy donor lines, as assessed by alignment of sequencing reads to the C9ORF72 gene (Supplementary Fig. S5D). We therefore next generated reporter lines to test whether the elimination of RNA foci by these small molecules is dependent on the C9ORF72 gene context. We engineered stable iNeurons and HeLa cell lines with 152x G4C2 repeats upstream of an eGFP coding sequence (Fig. 5A). RT-PCR revealed the expression of RNA transcripts with lengths corresponding to ca. 100x repeats, suggesting loss of some repeats during clonal selection (data not shown). The reporter cells showed nuclear G4C2-RNA foci with similar levels as our ALS lines (Fig. 5B). In contrast to what we expected, however, we observed a clearance of the RNA foci after incubation with PlaB and SML-3 in both, HeLa and iNeuron G4C2 reporter cells (Fig. 5C and Supplementary Fig. S5E). This demonstrates that the SF3B1 compounds can promote G4C2 nuclear RNA foci clearance independent of the C9ORF72 gene context. The constructs used for our 152x G4C2 reporter lines did further not contain any intron/exon context, suggesting that RNA foci clearance is not dependent on splicing of the repeat containing transcripts *in cis* either. We observed only a minor change of eGFP RNA levels upon SML-3 treatment for 4 h, while we did observe a strong increase of unspliced β-actin pre-mRNA as a control (Supplementary Fig. S5F), supporting that the reporter transcript is not non-canonically spliced or downregulated through another mechanism upon SF3B1 compound treatment. Since the reporter constructs are further not under the control of the endogenous C9ORF72 promoter, this also indicates that the compounds do not act at the level of transcription. To unambiguously test whether SF3B1 modulation affects G4C2-transcription versus the metabolism of RNA foci, we generated constructs with 152x G4C2 repeats under control of a tetracycline responsive element (TRE) driven promoter to shut off transcription upon Doxycycline addition (Tet-off assay, Fig. 5d). This reporter allows measurement

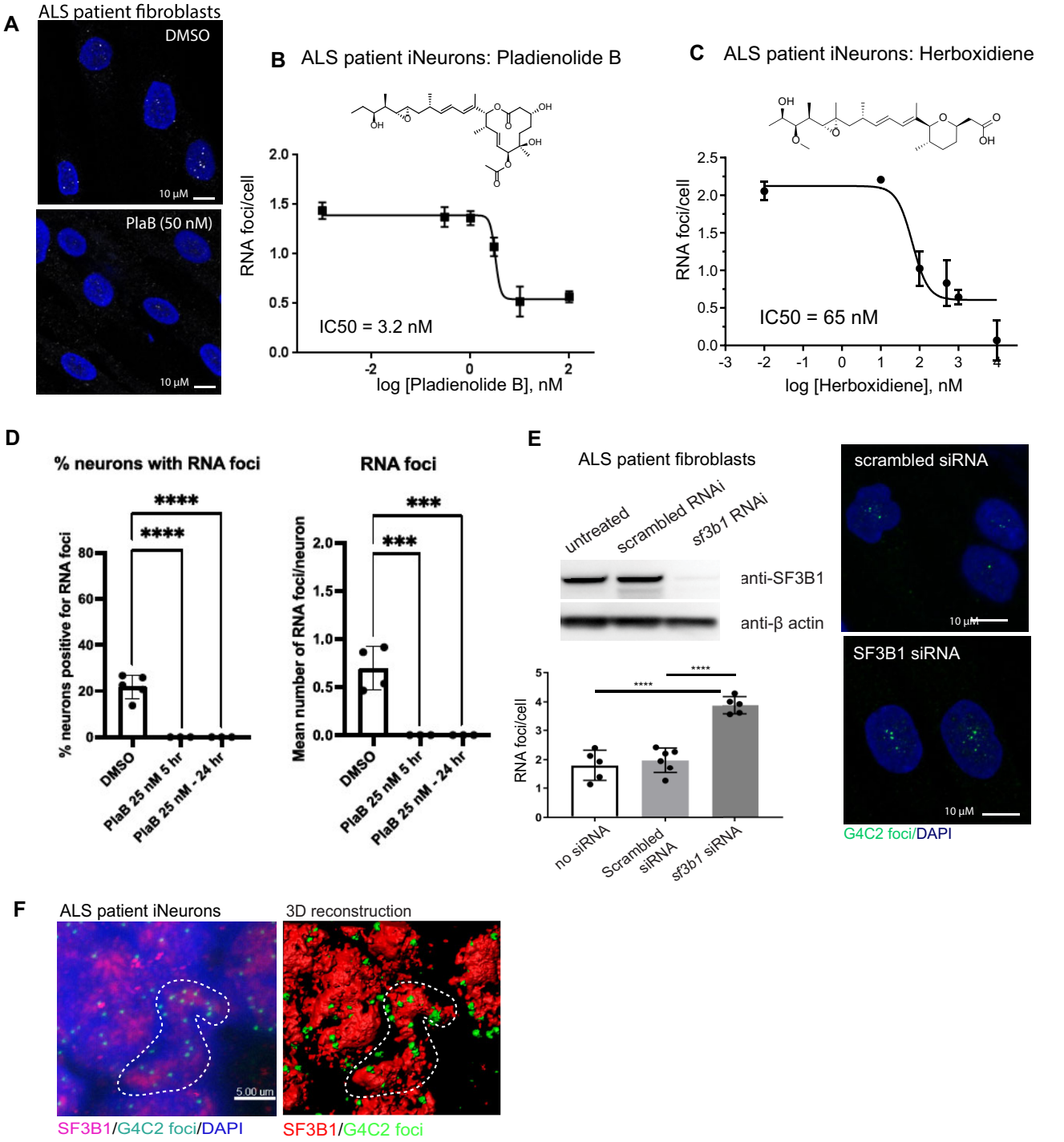


Figure 4. Pharmacological as well as gene expression modulation of SF3B1 alters G4C2 RNA foci levels. (**A**-**C**) RNA foci clearance by orthogonal chemotypes targeting SF3B1. (**A**) Confocal images of RNA foci in ALS fibroblasts treated with DMSO or 50 nM of Pladienolide B (PlaB). Note that fibroblasts contain less foci per cell than iNeurons. Herboxidiene (**B**) and PlaB (**C**) were added in different concentrations for 8 h to ALS iNeurons. Each condition is an average of three wells, error bars are SD. (**D**) Quantification of RNA foci in ALS patient iPS derived motor neurons after treatment with PlaB (Between 3 and 4 fields of view were analyzed (each dot represent a field of view). Number of neurons counted for each condition: DMSO: 349 neurons, PlaB (25 nM, 5 h): 102, PlaB (25 nM, 24 h): 156, Statistical analysis: One-Way ANOVA with Dunnet's multiple comparisons test, error bars: SD). (**E**) Left: Western blot of SF3B1 in ALS patient fibroblasts treated twice over a total of 4 days with siRNA targeting *sf3b1* or scrambled siRNAs. Loading control: b-actin. Middle panel: Representative images of G4C2 RNA foci (green dots) in siRNA treated ALS fibroblasts, showing increased number of RNA-foci in SF3B1 knockdown cells. Right panel: average number of RNA-foci per cell upon SF3B1 knockdown ($n = 3$ wells, 2500 cells per well, error bars: SD, significance: two-tailed One-Way Anova, **** indicates $P < 0.0001$). (**F**) Detection of RNA foci and SF3B1 in ALS patient iNeurons by confocal microscopy. Nuclei stain: DAPI. Right panel: 3D reconstruction (Imaris). RNA foci accumulate in the periphery of SF3B1 speckles.

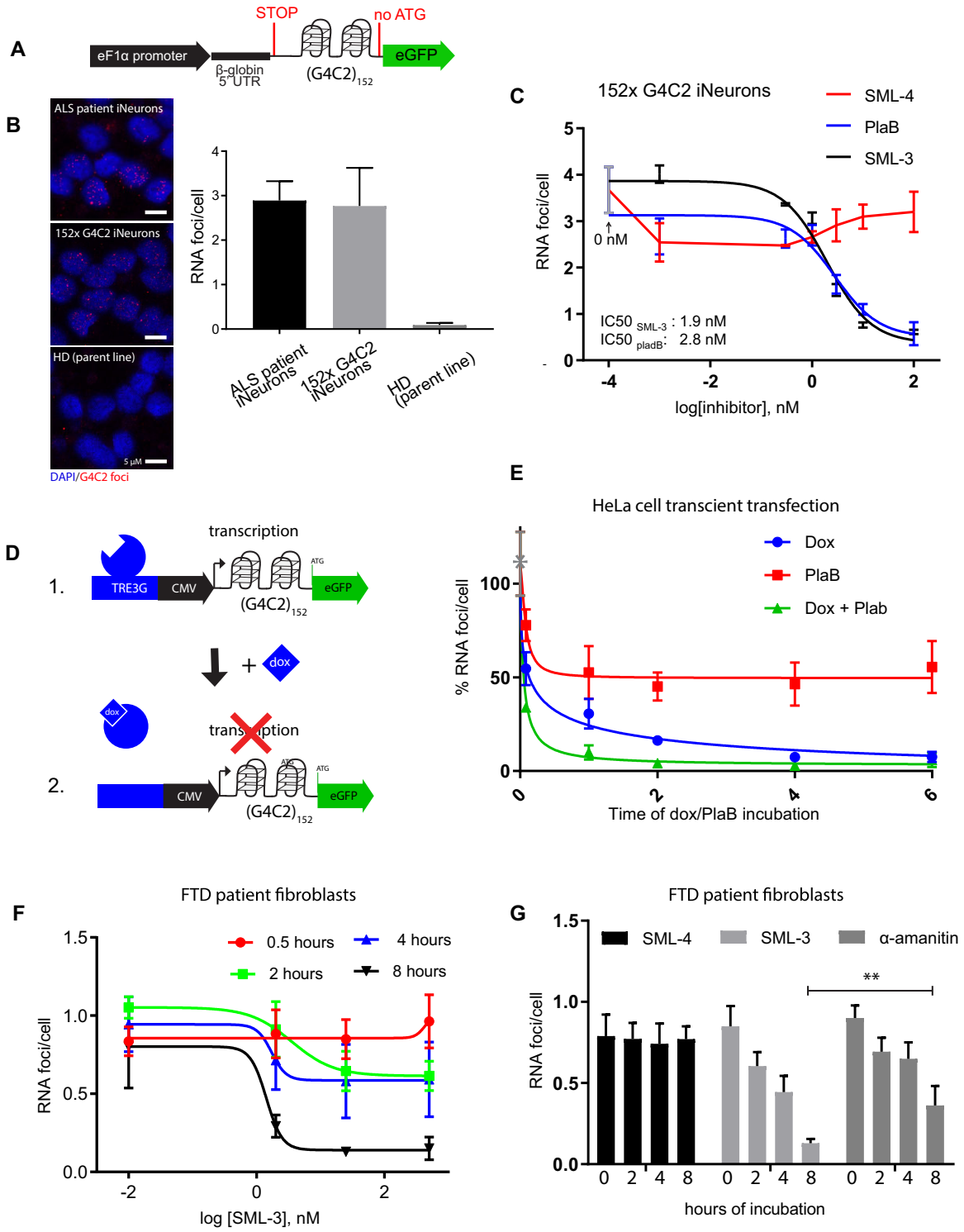


Figure 5. G4C2-RNA foci are eliminated posttranscriptionally. (**A**) A construct carrying 152x G4C2 repeats under expression of the *ef1a* promoter and a stop codon upstream of the repeats. The construct was randomly integrated in HD iNeurons with PiggyBac transposase to generate stable 152x G4C2 iNeuron reporter lines. (**B**) Confocal images confirm the presence of G4C2-RNA foci in the nucleus (DAPI) of reporter lines (left panel). The reporter line, ALS patient iNeurons (ALS G9) and the healthy donor parent line were analyzed side by side for comparison. A quantification is shown in the right panel. (**C**) RNA foci in 152x G4C2 reporter iNeurons, treated with PlaB, SML-3, and SML-4 at different concentrations for 8 h. Each bar represents an average of three wells (~5000 cells per well) and error bars are SDs. (**D**) A schematic overview of the TetOFF construct with 152x G4C2 repeats as used in Fig. 5E. Upon Doxycycline addition, transcription is turned off. (**E**) RNA foci in HeLa cells transfected with 152x G4C2 repeats under a TetOFF promoter, were treated for 0.5–6 h with 50 nM PlaB and/or Dox ($n = 3$ wells, 15 000 cells/well, error bars represent SDs). Grey dot represents $t = 0$ (before Dox addition). (**F**) RNA foci clearance by SML-3 as a function of time and dose after treatment with SML-3 (FTD patient fibroblasts treated with SML-3 for 0.5, 2, 4, and 8 h, $n = 3$ wells, 300 cells/well, error bars are SDs). Data for SML-4 are shown in Supplementary Fig. S5i. (**G**) RNA foci in FTD patient fibroblasts treated with SML-3 (500 nM), SML-4 (500 nM) or the RNA polymerase II/III inhibitor α-amanitin (20 ug/ml) for different time points. RNA foci clearance induced by SML-3 is faster than by blocking transcription by α-amanitin. The difference between the 8 h time-points for SML-3 and α-amanitin is statistically significant ($P > 0.05$ t-test, $n = 3$ wells 300 cells/well, error bars are SDs).

of RNA foci half-lives upon transcription inhibition. The presence of RNA foci after transient expression of this reporter in HeLa cells was confirmed by confocal imaging (Supplementary Fig. S5G). Upon transcription inhibition by Dox, nuclear RNA foci disappeared with a half live shorter than 1 h, revealing a high intrinsic turnover rate (Fig. 5E and Supplementary Fig. S5G). Addition of PlaB together with Dox resulted in an increased rate of RNA foci elimination above the background of transcriptional inhibition. This shows that SF3B1 modulating compounds act posttranscriptionally by accelerating the turnover of nuclear RNA foci. Consistently, the kinetics of endogenous RNA foci clearance upon SML-3 treatment was fast in C9ORF72 patient fibroblasts as well, with a $T_{1/2}$ of ca 2–4 h (Fig. 5F) and thus faster than the response to global inhibition of RNA Polymerase-II transcription by α-amanitin (Fig. 5G). Treatment with α-amanitin led to a minor dose- and time-dependent loss of cells, whereas the cell numbers remained stable over increased incubation time and dose for SML-3, suggesting no significant toxicity of this compound under the conditions of the experiment (Supplementary Fig. S5h). SML-4 again showed no activity (Supplementary Fig. S5i and Fig. 5g). These data demonstrate that SF3B1 modulating compounds lead to an accelerated RNA foci elimination rate, thereby further corroborating a posttranscriptional mechanism.

SF3B1 modulation increases RAN DPR levels and enhances reduction of cell growth by DPRs

Since nuclear G4C2 RNA foci were eliminated by SML compounds while overall reporter RNA levels were unchanged, we next tested the effect of these compounds on dipeptide repeat proteins (DPRs). Detection of endogenous G4C2 derived DPRs in C9ORF72 neurons is particularly challenging due to low expression levels and limited affinities of commercially available antibodies to the linear, unstructured epitopes. Using antibodies commercially available at the time of the study, we were unable to detect RAN DPRs in our patient iPS derived Ngn2 and motor neurons by western blotting or immunocytochemistry (data not shown). To nevertheless investigate the effects of SF3B1 modulation on RAN DPRs, we therefore engineered a stable RAN reporter cell line in HEK293 cells, expressing 152xG4C2 repeats fused to frameshifted 3xmyc, 3xHA and 3xFLAG tags, in frame with either poly-GA, poly-GP and poly-GR dipeptide repeat (DPRs) polyproteins, respectively (Fig. 6A). To prevent canonical translation, the constructs were engineered without a start codon and by additionally inserting three upstream stop codons. Upon addition of Doxycycline, both, poly-GP and poly-GA DPRs were expressed, whereas the signals for poly-GR were substantially lower (Fig. 6B). RAN control constructs with the tags only were not expressed and also not induced by Doxycycline (Fig. 6A), confirming that RAN translation is dependent on the repeat RNA. In line with the endogenous localization of DPRs shown previously [26], the subcellular localization of the DPRs visualized by the tags was mostly nuclear for GA and GP, with some inclusions in the cytoplasm, whereas poly-GR was mostly detected in cytoplasmic inclusions (Fig. 6b, white arrows). Furthermore, the subcellular localization of the tags fused to the DPRs differed from the localization of the tags only expressed from control reporter constructs by ATG-dependent, canonical translation (Supplementary Fig. S6A), thereby indicating that the localization is determined by the DPR. Using antibodies targeting the DPR directly, we again were not able to obtain a robust signal for either of the polyGA, polyGP or polyGA DPRs (Supplementary Fig. S6B). To investigate the effects of SF3B1 compounds on RAN DPRs, we therefore detected the DPRs indirectly via antibodies to the myc, HA or FLAG tags. After 24 h of treatment with SML-3 or PlaB, we observed a massive increase in both GA and GP DPRs (Fig. 6C for quantification, and Fig. 6D for exemplary images and additional quantification in Supplementary Fig. S6E and F), which is contrary to our initial hypothesis that RNA foci clearance would also result in a reduction of RAN translation products. We did not detect an increase in GR-DPRs (as monitored by the Flag tags, Supplementary Fig.Fig. S6G); however this might be due to the generally low levels detected in our transgenic DPR-expressing lines. DPR levels increased with time of exposure to SML-3 (Supplementary Fig. S6C), and in a dose dependent way (Supplementary Fig. S6D). Since several independent studies have shown that DPR proteins can be toxic to cells (e.g. [96]), we tested whether the increased amounts of DPR induced by SML-3 and PlaB might lead to an increase in DPR cell toxicity as assessed based on cell numbers and growth. RAN reporter cells were treated with PlaB, SML-3 or SML-4 each with or without Dox-induction over a course of 96 h and cell confluency was monitored every 3 h. Combining SML-3 or PlaB with DPR induction indeed resulted in significantly decreased cell growth (Fig. 6E, light red and light blue line) compared to addition of SML-3 or PlaB alone (Fig. 6E, red and blue line). No increase in cell toxicity was observed by SML-4 (Fig. 6E, dark and light green lines). We would like to note that although Flag, myc and HA tags have been overexpressed numerous times in HEK cells in other studies, we cannot entirely exclude any potential contribution of the tags to the toxicity. Given the continuous development of improved tools to detect RAN DPRs within the field, further experiments assessing the effect of SF3B1 modulators on native RAN translation in patient cells would be ideal in follow up studies.

SML compounds mobilize G4C2 RNA from nuclear RNA foci into nuclear export by recruiting SRSF1 to repeat RNA G-quadruplex structures

The fact that the levels of DPRs are increased rather than decreased upon SF3B1 modulator treatment suggests that the G4C2 repeat transcripts are fed into translation, potentially by mobilization of RNA from the ‘depot’ of nuclear foci. This might therefore involve increased export to the cytoplasm, where the G4C2-repeat transcripts are translated. In fact, SF3B1 modulating compounds were shown previously to induce nuclear export of unspliced RNA [95], which we also confirmed to be true for the new SF3B1 compounds by RT-PCR of unspliced β-actin transcripts within nuclear and cytoplasmic fractions of HeLa cells, pretreated with DMSO, SML-4, SML-3 and PlaB (Supplementary Fig. S7A and B). To test whether SML-3 and PlaB also trigger nuclear export of G4C2 RNA, we introduced 24xPP7 and MS2 stem-loop elements into our 152xG4C2 reporter (Fig. 7A) and generated stable transgenic HEK293 cell lines to allow for single-molecule FISH (smFISH) detection of the repeat RNA. As previously shown by us and others, the smFISH probes targeted to the PP7 stem-loops allow for quantitative imaging at single RNA molecule resolution [97]. After addition of 50 nM PlaB for 90 min, we observed a significant shift in repeat transcripts

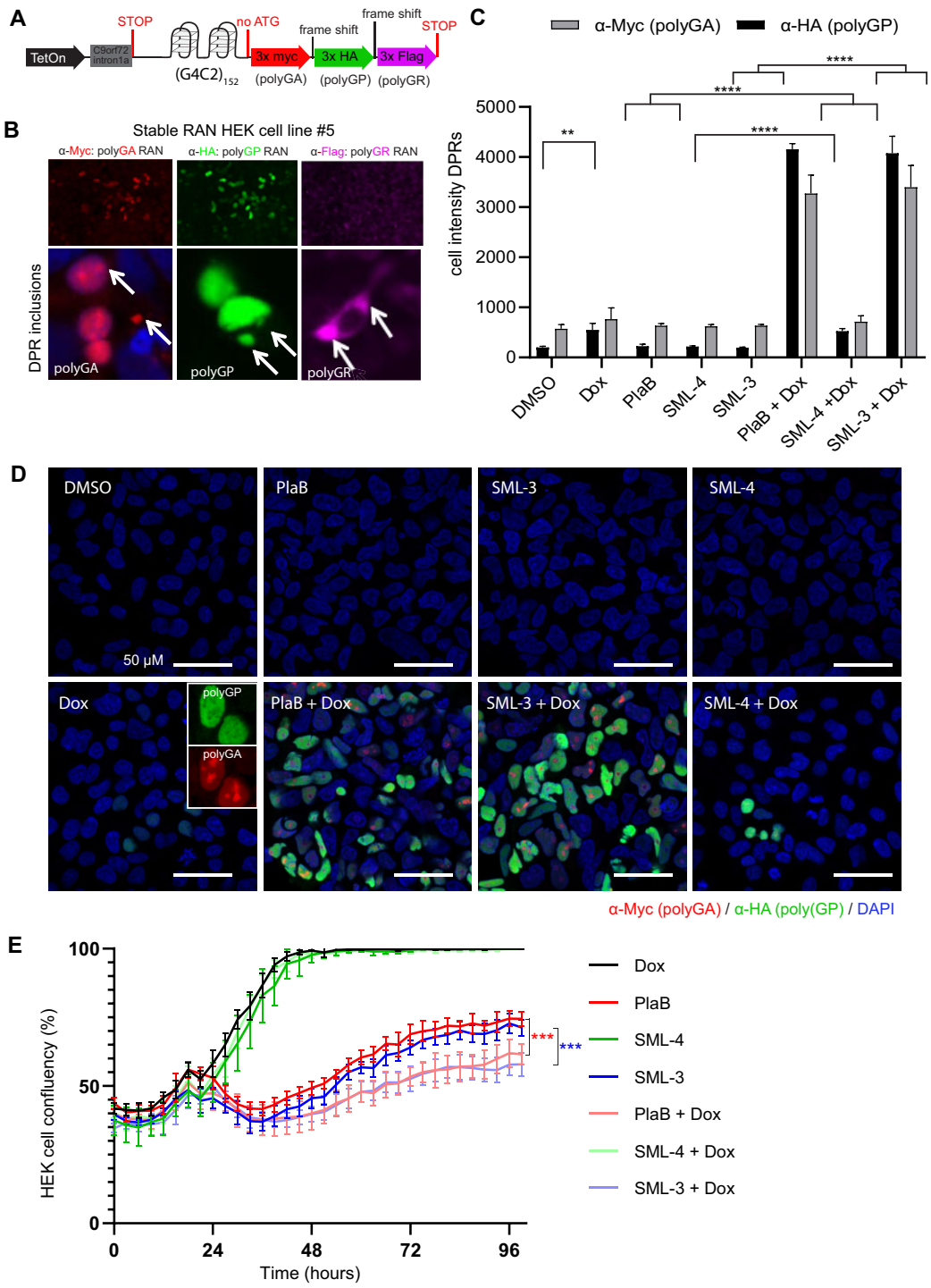
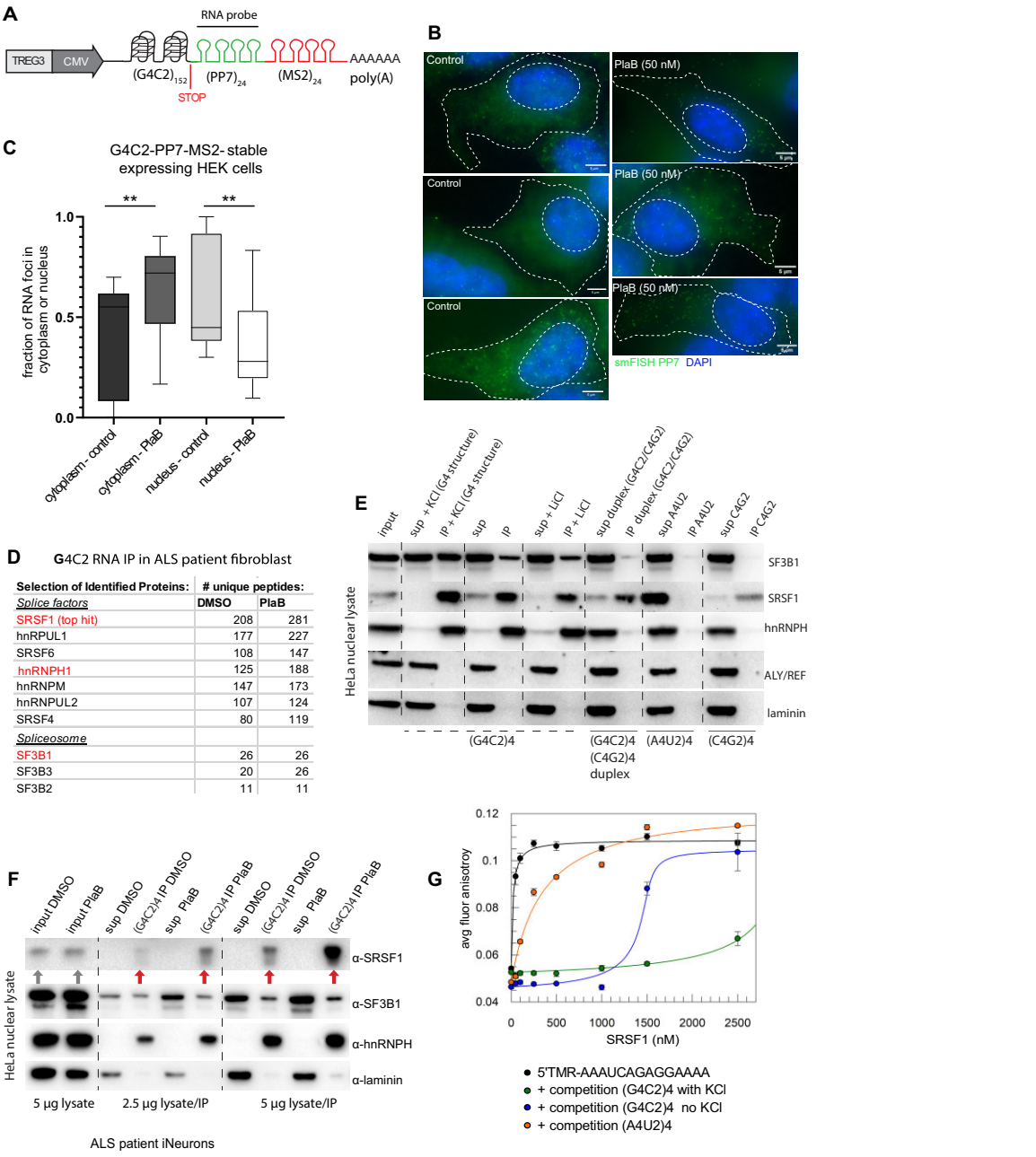


Figure 6. SML compounds increase DPR levels and enhance DPR toxicity, measured by cell proliferation. (**A**) Schematic representation of the Dox-inducible RAN reporter construct used for stable insertion into HEK293 cells, including an upstream Stop codon, no ATG and fusing three reporter tags in different frames downstream of the repeat. Depending on the reading frame, polyGA DPRs carrying a Myc-tag, polyGP DPRs carrying an HA-tag, or polyGR DPRs carrying a FLAG-tag are expressed upon Dox induction of reporter transcription. (**B**) Widefield fluorescent microscopy images showing expression of the different DPRs (Myc-tag: polyGA; HA-tag: polyGP; Flag-tag: polyGR) in RAN reporter lines upon Dox induction, showing high expression of polyGA and polyGP DPRs, and very low expression of polyGR. Zoomed-in confocal figures (lower panel) show subcellular localization of DPR inclusions in the HEK293 152xG4C2 reporter lines, with mostly nuclear localization with some cytoplasmic aggregates for polyGA and polyGP. PolyGR showed mostly cytoplasmic localization (**C**) Quantification of the images represented in Fig. 5E. Each datapoint is an average of 6 wells, with approx. 5'000 cells per well. Statistical differences are shown (t test; $****P < 0.0001$ $**P = 0.002$). (**D**) Confocal images of polyGA and polyGP in HEK293 RAN reporter cells treated with 100 nM of SML-3 and PlaB for 24h, showing an increased production of DPRs induced by SML-3 and PlaB compared to Dox only or Dox + SML-4. Zoomed-in frame (Dox panel) of PlaB + Dox treated HEK293 RAN reporter cells suggests nucleolar polyGA inclusions, consistent with previous reports [104]. (**E**) Cell proliferation measured by continuous bright field imaging (Incucyte) for RAN reporter cells with and without Dox-induced RAN expression and 100 nM SML-3, SML-4 or PlaB treatment for up to 100 h ($n = 8$ wells and error bars are SD). SML-3 and PlaB treatment significantly enhanced the inhibitory DPR effect on cell proliferation, showing that increased levels of DPRs are inducing reduced cell proliferation (t test, $P < 0.0001$).



Selection of Identified Proteins:	# unique peptides:	
Splice factors	DMSO	PlaB
SRSF1 (top hit)	208	281
hnRPUL1	177	227
SRSF6	108	147
hnRNPH1	125	188
hnRNPM	147	173
hnRNPUL2	107	124
SRSF4	80	119
Spliceosome		
SF3B1	26	26
SF3B3	20	26
SF3B2	11	11

Figure 7. Spliceosome modulating compounds induce G4C2 repeat RNA export. (**A**) Schematic representation of a reporter construct used for generating stable HEK293 reporter cells with PP7 and MS2 stem loops downstream of 152xG4C2 repeats that allow visualization of single RNA molecules by FISH. The G4C2 reporter RNA was under control of a TREG3 promoter to allow Tetracyclin induced inhibition of transcription, thereby preventing transcript accumulation at transcriptional start sites. (**B**) Confocal images of single molecule FISH against PP7 stem-loop RNA (green) in HEK293 G4C2 single molecule FISH reporter cells. Nuclei stain: DAPI. A redistribution of G4C2-PP7 RNA molecules in cells treated with PlaB to the cytoplasm can be observed. (**C**) Quantification of single molecule FISH against PP7 stem-loop RNA in HEK293 G4C2 single molecule FISH reporter cells upon compound treatment. Cells were treated for 1.5 h with 50 nM Pladienolide B and 45 min of Dox. Controls: Dox only. A significant difference between the cytoplasmic and nuclear fractions of G4C2 RNA molecules in control and PlaB treated cells was observed (Welch's t test, $P = 0.0028$). (**D**) MassSpec analysis of proteins co-immunoprecipitated with biotinylated (G4C2)4-RNA in nuclear lysates from PlaB or DMSO pretreated HeLa cells. Selection of top hits and spliceosomal proteins are shown (full data set in Supplementary Data S4). (**E**) Western blot for SF3B1, SRSF1, export factor ALY/REF, hnRNPH (positive control), and laminin (loading control) after co-immunoprecipitation with different biotinylated RNA oligonucleotides in nuclear HeLa extracts. (G4C2)4 RNA oligos were pre-folded in different structures prior to the IP. KCl or LiCl were used to facilitate or disrupt G-quadruplex formation, respectively. Hybridization to the complementary strand was used to fold the (G4C2)4 oligos into dsRNA. Confirmation of RNA secondary structures by CD spectroscopy is shown in Supplementary Fig. S7c. Biotinylated (A4U2)4 or (C4G2)4 RNA oligos were used for comparison. The blot shows one representative data set out of 15 independent experiments. (**F**) Western blot for SRSF1, SF3B1, hnRNPH (positive control) and laminin (loading control) after co-immunoprecipitations with biotinylated (G4C2)4 RNA using nuclear extracts of HeLa cells pre-treated for 4 h with PlaB (50nM) or DMSO prior to lysis. An increased binding of SRSF1 to the G4C2 RNA oligo in lysates from PlaB pretreated cells was observed. (**G**) Binding of recombinant SRSF1 to 5'TMR labeled RNA (canonical motif: AAAUCAGAGGAAAA, 0.5 nM) measured by fluorescence correlation spectroscopy (2D-FIDA). RNA-protein complex formation is monitored based on an increase of the fluorescence anisotropy upon protein binding to the TMR labelled RNA. Binding of SRSF1 to its canonical motif in presence of unlabelled (G4C2)4 RNA (5 uM) prefolded in absence or presence of 200 µM KCl was fitted to a model of 1:1 competition. Competition with (A4U2)4 RNA is shown as control for unspecific RNA binding.

to the cytoplasm as compared to the no compound control (Fig. 7b and c), confirming that RNA foci disassembly triggered by spliceosome modulating compounds is accompanied by enhanced export of the transcripts to the cytoplasm.

To gain an insight in the molecular factors involved in RNA foci clearance and nuclear G4C2 RNA export by SF3B1 modulation, we performed RNA-protein immunoprecipitation experiments. Biotinylated $(G4C2)_4$ RNA oligonucleotides were spiked into nuclear lysates from HeLa cells with and without compound pre-treatment. The co-precipitated proteins were then analyzed by LC/MS-proteomics for unbiased identification of proteins differentially binding the repeat RNA in response to SF3B1 compounds. Among the top hits of proteins that were increased in binding to G4C2 RNA in response to PlaB treatment we detected the splice factor and nuclear export adaptor SRSF-1 (also known as ASF/SF2) (Fig. 7D and Supplementary Data S4). This was particularly interesting as previous studies have shown that SRSF1 is involved in the nuclear export of intron-1-retained C9ORF72 repeat transcripts, required for cytoplasmic RAN translation of DPRs [98]. Since we had also observed that PlaB increased RAN translation and G4C2-repeat export in our HeLa/HEK293 cells expressing 152x G4C2 repeats, we set out to further elucidate the involvement of SRSF1 in the relation between spliceosome modulation and RNA foci clearance.

G4C2 repeat RNA can fold into stable G quadruplex (G4 structures) *in vitro* [10], and it has been shown that this conformation is also predominant *in situ* within the RNA foci [99, 100]. We therefore set out to investigate the interaction of SRSF1 and synthetic biotinylated $(G4C2)_4$ RNA as a function of RNA folding, as well as in comparison to other control oligonucleotides by RNA-protein immunoprecipitation and western blotting. The splicing factor hnRNPH was used as a control, since it co-localizes with RNA foci [36] and was shown to bind $(G4C2)_4$ RNA directly [101]. ALY/REF was included as an additional export factor interacting with the splice machinery [102]. hnRNPH, SF3B1 and SRSF1 were all pulled down with $(G4C2)_4$-RNA but not with control RNA sequences (A4U2 and C4G2) (Fig. 7E). No pull down of ALY/REF was detected with any of the biotinylated oligos. A preformed duplex of G4C2/C4G2-RNA did not pull down hnRNPH, consistent with its known single stranded RNA binding preference. A strongly diminished binding of SF3B1 and, to a lesser extent, SRSF1 was observed to the G4C2/C4G2 duplex. G4 formation is generally facilitated by potassium and disrupted in presence of lithium ions, respectively [10], which we confirmed for our biotinylated $(G4C2)_4$ oligos by Circular Dichroism (CD) spectroscopy (Supplementary Fig. S7C). Binding of both, SRSF1 and SF3B1 increased when the RNA was pre-folded into G4 structure using potassium chloride, whereas hnRNPH binding remained unchanged. Consistently, disruption of the G4 RNA structure by lithium chloride resulted in a small decrease in both, SF3B1 and SRSF1 binding as compared to binding with addition of KCl (Fig. 7E). Addition of unbound $(G4C2)_4$ or $(A4U2)_4$-RNA to the $(G4C2)_4$ RNA-IP did compete in binding for SRSF1, and, to a lesser extent, also SF3B1 (Supplementary Fig. S7D). These data were highly consistent with the known RNA binding properties of these splice factors [73] and supported the validity of our RNA-protein interaction assay in nuclear extracts.

To further corroborate the data from the LC-MS proteomics, we additionally tested the effect of PlaB on the binding capability of SF3B1, SRSF1 and hnRNPH to the biotinylated (G4C2)4 RNA in nuclear extracts of HeLa cells pretreated with the compounds (Fig. 7F). PlaB resulted in slightly increased binding of SF3B1 to (G4C2)4-RNA whereas again a massively increased binding of SRSF1 to $(G4C2)_4$-RNA upon PlaB Hela cell pretreatment was observed (Fig. 7F, red arrows), consistent with the proteomics data. Total nuclear SF3B1 and SRSF1 levels were unchanged (Fig. 7F grey arrows).

The binding of SRSF1 to G-quadruplex folded $(G4C2)_4$-RNA induced by PlaB treatment and detected by RNP-immune precipitation from nuclear lysates could be either due to a direct binding to the RNA or mediated via other RBPs. To unambiguously address whether SRSF1 can bind the G4C2 repeat RNA sequence directly and via its canonical RNA binding site, we analyzed the interaction of recombinant SRSF1 (Supplementary Fig. S7E) with G4C2 repeat RNA using fluorescence correlation spectroscopy (Fig. 7G, Supplementary Fig. S7F and G). As a positive control for recombinant SRSF1 protein activity, we used its canonical binding motif (AUCAGAGGAA), fluorescently labeled with TMR, which was bound by SRSF1 with nanomolar affinity ($Kd = 21 \pm 3$ nM; Fig. 7G), similar to published Kd values [73]. Unlabeled $(G4C2)_4$ RNA directly competed with binding of the canonical RNA motif to SRSF1 revealing an affinity of $Kd < 21$ nM. $(A4U2)_4$ RNA showed only weak competition ($Kd \sim 200$ nM). This showed that SRSF1 can bind to repeat RNA directly in the low nanomolar range, similar to hnRNPH (Kd 12 nM, Supplementary Fig. S7F). The direct competition with the AUCAGAGGAA RNA further suggests that $(G4C2)_4$ RNA is bound to the canonical RNA binding pocket of SRSF1. The affinity for $(G4C2)_4$ RNA slightly increased in presence of K^+ (Fig. 7E) and only weak binding to (G4C2)2 was observed (Supplementary Fig. S7G). Together these data demonstrate that SRSF1 is able to bind directly to a pre-folded G4-structure via its canonical RNA binding motif, and that this interaction is substantially enhanced in response to small molecule modulation of SF3B1.

Protein-protein interaction between SRSF1 and SF3B1-containing spliceosomal subcomplexes was previously shown by immune precipitation and LC/MS proteomics [103]. To confirm the interaction between SF3B1 and SRSF1 in a cellular model where SML compounds clear G4C2 RNA foci, we investigated whether there is a change in SRSF1 and SF3B1 protein complex composition upon PlaB treatment. We separately cultured HeLa cells with either heavy isotope labeled amino acids or normal amino acids and treated these cells with either DMSO (heavy) or 25 nM PlaB (light) for 16 h. Heavy and light cells were mixed, lysed and their nuclear lysate was used in either SRSF1 or SF3B1 IPs. The proteins pulled down from these lysates were analyzed by mass spectrometry (Supplementary Dataset S1 and S2). SF3B1 as well as SRSF1 were among the proteins with highest peptide count in the respective immune precipitations, showing that the IP was successful. Interestingly, a large overlap (approximately 50 %) in co-precipitating proteins that were changed upon PlaB treatment was observed between SRSF1 and SF3B1 interactomes (Supplementary Fig. S7H). For example, both complexes had reduced binding of the pre-mRNA splice factor RBM22, the helicases DDX41 and DDX56 and increased binding of pre-mRNA splice factor PR38B and the importin subunit IMA1 upon PlaB treatment (Supplementary Fig. S7I, Supplementary Dataset S3). Furthermore, log(Heavy/Light) ratios of SRSF1 and SF3B1 IPs for proteins that are present in

both samples largely overlapped, as visualized by the highly significant correlation depicted in Supplementary Fig. S7i ($R^2 = 0.835$). These data demonstrate that the interactome of SRSF1 and SF3B1 not only largely overlap but also respond to PlaB treatment to the same extent.

Together, our data show that SF3B1 modulating small molecules license the G4C2-repeat RNA from nuclear RNA foci into RAN-translation in the cytoplasm resulting in elevated DPR toxicity. Elimination of nuclear RNA foci by spliceosome modulating compounds is neither a consequence of halted transcription, nor altered splicing of repeat containing pre-mRNA or induction of G4C2 RNA decay. Instead, these compounds alter the canonical spliceosome assembly in such a way to allow redistribution of nuclear SF3B1 and SRSF1 to non-canonical G4C2 target RNA, thereby promoting its nuclear export and translation into DPR polyproteins (Fig. 8). These data support the previously proposed hypothesis that nuclear RNA foci might act as RNA storage granules, protecting the cell against DPR production and ultimately cell toxicity [52, 104].

Discussion

We have performed a large G4C2-RNA foci screen of approximately 100 000 small molecules in C9ORF72 iNeurons miniaturized to the 1536 well format, showing a high confirmation rate across donors and iPS lines. Our screening campaign proves the feasibility of high-throughput small molecule screening in primary patient derived neurons and establishes FISH based RNA foci detection in general as an HTS readout. Filtering for unspecific gene expression inhibitors turned out to be essential, since ca 50% of the primary hits were interfering with the canonical gene expression machineries. Our high hit rate of transcription inhibitors is consistent with a very recent report of a small screen with a focused library of 1430 approved drugs for RAN DPR modulation in iPSC neurons which identified nucleoside analogs resulting in both, reduction of RAN DPRs as well as G4C2 RNA [100]. Furthermore, we observed that G4C2-repeat containing foci are dynamic structures with a high turnover and $T_{1/2}$ of ca. 4 h that are pharmacologically tractable. We identified several RNA foci clearing compounds involved in chromatin modification and C9ORF72 expression regulation, suggesting heavy epigenetic and transcriptional control of the C9ORF72 locus in ALS patient cells. In addition, we identified compounds targeting proteins that had previously also scored in a genetic CRISPR-Cas9 screen by Kramer and colleagues [105]. Besides JQ1, we identified a significant number of epigenetic modulators including BET bromodomain inhibitors in our screen. This is in line with conclusions from a previously reported focused small molecule screen of 4125 annotated bioactive compounds for alleviation of DPR toxicity in a PR20 reporter assay in U2OS cells [106]. All of these compounds and target proteins might provide additional starting points for follow up investigations.

Our RNA foci screen further identified SF3B1 modulators which we selected for mechanistic investigation, aiming to gain deeper insight in possible pharmacological entry points of the repeat RNA metabolism. SF3B1 modulating compounds such as Pladienolide B, Spliceostatin or the newly identified SML compounds lock SF3B1 in an open conformation, thereby stalling the spliceosome in the A-complex and reducing stringency and affinity of RNA binding [95, 107–109]. This results in alternative splicing or splice inhibition of over 2000 pre-mRNAs, thereby downregulating a large number of canonically transcribed mRNAs. While the C9ORF72 mRNA is among the genes affected by SF3B1 modulators, its downregulation is independent of the repeat expansion since it also occurs in cells from healthy donors, and does not involve altered splicing of the repeat containing C9ORF72 intron as revealed by RNASeq. Additionally, G4C2 repeat RNA is affected by these SF3B1 targeted compounds independent of the C9ORF72 gene context.

In consequence of SF3B1 modulation, the interactome of G4C2 RNA is altered, resulting in increased binding of the splicing and nuclear export factor SRSF1 and facilitating the movement of G4C2 RNA from nuclear foci to the cytoplasm. As a result, there is a significant increase in translation into dipeptide repeats, ultimately leading to cell toxicity (Fig. 8). Note that while our data show a clear correlation between DPR increase and cell toxicity, causality of the DPRs themselves for the toxicity was not established within our study. Previous studies have used different approaches to test causality of DPR for neuronal phenotypic consequences. For example, a 152x reporter interrupted by multiple stop codons was used to prevent long DPR production in a Drosophila eye degeneration model [110], allowing the investigation of neuronal toxicity driven solely by the repetitive RNA. Using such a reporter in our system would imply that the DPRs can no longer be monitored by the tags and instead require antibodies recognizing endogenous DPRs, which were ineffective even for the non-interrupted reporter (see Supplementary Fig. S6B). To further establish that toxicity observed upon RNA foci elimination by SF3B1 modulation not only correlates with increase RAN DPRs but is also mediated by them, further experiments will be required ideally in patient cells, such as by rescue experiments with DPR eliminating tools as they become available.

SF3B1 modulation by SSA or meayamycin was reported to induce alternative splicing changes similar to those observed with SF3B1 knockdown [90, 111], which indicates that the compounds phenocopy SF3B1 loss of function in transcriptome wide splice consequences. Notably, our data reveal opposing effects of genetic versus pharmacological modulation of SF3B1 on RNA foci: SF3B1 knockdown increased RNA foci levels, whereas small molecule inhibition reduced them. On one hand these data show that SF3B1 is generally involved in constitutive RNA foci turnover, thus explaining the RNA foci increase upon SF3B1 knockdown. On the other hand, the differential effect of the SF3B1 lmw modulators might indicated that they enhance rather than inhibit this role of SF3B1 in RNA foci turnover, thus explaining the RNA foci elimination by the SF3B1 targeted compounds. Whereas the exact mechanism remains to be further investigated, one possibility might be that they induce a non-canonical gain of function of SF3B1 irrelevant to transcriptome wide splicing, but relevant for the turnover of repeat expanded RNA. Such non-canonical gain of function may involve de novo recognition of cryptic and potentially non-productive splice sites, which has been shown to occur *in vitro* [84, 85].

Our data are consistent with the described role of SRSF1 in nuclear export of G4C2 repeat RNA [18] and now link this non-canonical export of a pathological RNA to the canonical splice machinery. The exact molecular mechanisms of how SF3B1 modulation recruits SRSF1 to the RNA foci will need further investigation. Interestingly, in our RNP immune pre-

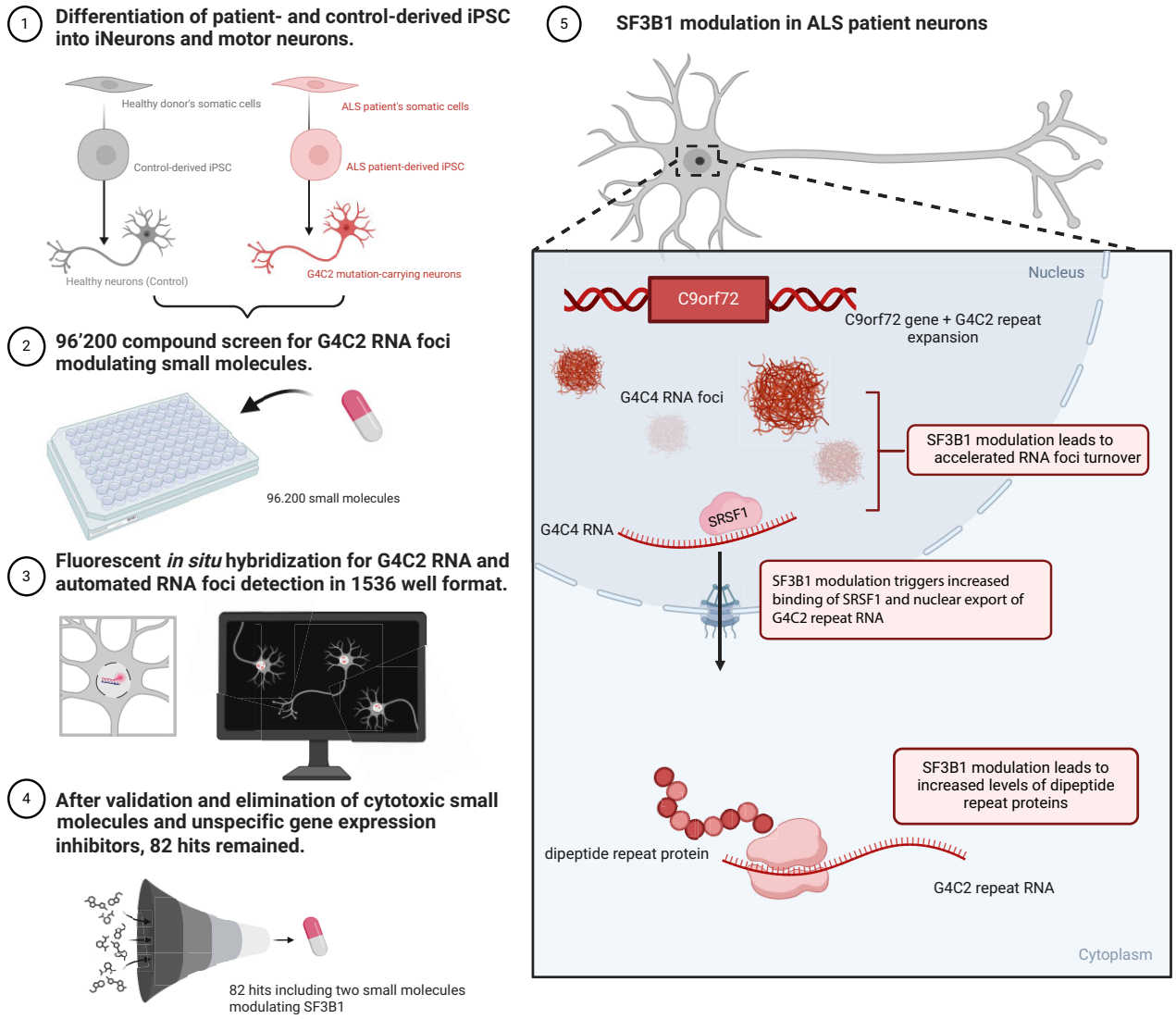


Figure 8. Model for compound screen design and turnover of nuclear RNA foci in response to SF3B1 modulating compounds explained in the discussion. Created with BioRender.com.

cipitation experiments, we observed an increased binding of not only SRSF1 but also SRSF4 and SRSF6, while ALY/REF and hnRNPH were unaffected. This indicates a relatively selective redirection of splice factors to G4C2 RNA and further supports the central role of SR proteins in nuclear export of G4C2 RNA reported by other groups. Also, our data reveal that not only SRSF1 but also SF3B1 binds to the repeat RNA even in the absence of PlaB/SML induced non-canonical splicing. RNA foci cluster around SF3B1 positive nuclear speckles and SF3B1 depletion by RNAi results in an increase in RNA foci. This suggests that not only SRSF1 but a larger spliceosomal subunit including SF3B1 contributes to constitutive repeat RNA turnover and export. In fact, a study using single molecule imaging to investigate the spatio-temporal dynamics of G4C2 repeat RNA has demonstrated that repeat containing introns are exported as lariats [12]. This is well in line with our conclusion that non-canonical spliceosome assembly on the repeat containing RNA promotes G4C2 nuclear export, and that this is enhanced when reducing spliceosome stringency by SF3B1 targeted small molecules.

The cytoplasmic metabolism of the repeat RNA upon nuclear export elicited by PlaB/SML compounds, and in particular how the G4C2 repeat RNA is targeted for RAN translation, represent interesting questions for further investigation. The increased DPR levels could be a simple consequence of increased repeat RNA levels available in the cytoplasm as revealed by our single molecule FISH experiments. However, the relative levels of increased RAN translation products were higher than the levels of nucleo/cytoplasmic redistribution of the RNA, suggesting an additional modulation of the cytoplasmic RNA fate. SRSF1 was shown to also promote initiation of translation of the exported mRNA *in cis* [112]. Hence, SRSF1 bound to the repeat RNA could additionally increase the efficiency of RAN translation.

Previous studies have suggested that RNA foci do not contribute to ALS disease pathology while DPRs do [16, 113]. Yet, the toxicity of DPRs remains debated, as their presence does not always result in the development of ALS/FTD [42, 114]. On the other hand, the presence of DPRs in the brains of asymptomatic individuals may still be consistent with the

hypothesis of RAN DPRs driving the disease rather than RNA foci since the thresholds required for consequences on motor neuron function and survival may vary individually. An important implication of our data is that RNA foci could in fact even have a protective role, preventing translation into toxic DPRs, since removal of nuclear RNA foci by SF3B1 modulation worsens rather than alleviates DPR toxicity in cell growth assays. The typically late disease onset of ALS could accordingly be explained by increased leakage of repeat RNA from nuclear ´storage´ foci into the cytosol over time, due to an increasingly defective nuclear export pathway caused by aging [115] and aggravated by the DPR proteins themselves [33]. This might create a toxic positive feedback loop of repeat RNA leakage to the cytosol, increasing DPRs and consequently further damaging the nuclear export pathway, ultimately driving the motor neurons over the "catastrophic cliff" that has been proposed as trigger for the onset of ALS. Neurons could be particularly sensitive to this toxic loop due to lack of dilution of the DPR protein levels during cell division.

The SF3B1 modulating compounds identified in this study now provide both, novel as well as known and commercially accessible tools to study the relative contributions of RNA foci and DPRs to the C9ORF72 ALS pathology in a differential way, as opposed to genetic modulation where repeat RNA expression and thus RNA foci reduction typically leads to decreased RAN translation. DPRs are difficult to detect in patient cells due to their low levels, and phenotypes in ALS patient cells *in vitro* are typically subtle. Enhancing DPR production through RNA foci clearance, without affecting C9ORF72 transcription, may accelerate the onset of subtle, progressive and disease relevant phenotypes in iPS derived neurons, which are increasingly attractive models for ALS drug discovery [116]. We would like to point out that SF3B1 ligands such as PlaB, Spliceostatin or SML-3 not only clear G4C2 RNA foci but have pleiotropic activity in alternative splicing. This complicates the interpretation of any phenotypic outcomes and demands careful controls such as isogenically matched lines as well as additional pharmacological tools targeting different nodes in the same pathway. Our data identify a new pharmacological entry point into the nuclear export and RNA metabolism pathway of the G4C2 repeat RNA which will facilitate further tool compound and drug discovery.

Finally, our data warrant caution to be exercised when using nuclear RNA foci as a primary readout in the discovery of C9ORF72 therapeutic candidates. Unless acting through true RNA degradation, small molecules clearing RNA foci could in fact worsen ALS disease pathology instead of alleviating it. The negative results from a Phase 1 clinical trial with an ASO targeting C9ORF72 (BIIB078) where the drug worsened the disease in ALS patients, underline the concern that the disease mechanism is more complex than thought. In light of our data, using reduction of RNA foci as primary readout for preclinical drug development may indeed overlook the possibility that the RNA is fed into potentially more harmful pathways. Going forward, this shifts the bottleneck increasingly towards the development of truly predictive preclinical models to enable the investigation of multiple pharmacological strategies interfering with C9ORF72 pathobiology. This will ultimately help to identify the most promising therapeutic entry points prior to initiating studies in patients, whereas our data support the modulation of nuclear G4C2 RNA export by targeting SR proteins or SR protein kinases as a promising area for exploration.

Conclusions

In summary, we report the first pharmacological high-throughput screen in ALS patient iPS cell derived neurons and describe how spliceosome modulating compounds can be used to differentially modulate G4C2 RNA foci and repeat associated DPRs in C9ORF72 ALS pathology. SF3B1 modulators eliminate G4C2 RNA foci from the nucleus in at least eight different cellular models of ALS. Surprisingly, this is neither due to altered splicing of the repeat containing intron nor accelerated RNA decay; instead, these compounds alter the interactome of G4C2 RNA, enhance binding of nuclear export factors and mobilise the G4C2 RNA from nuclear foci into the cytoplasm. In consequence, this leads to massively increased translation into dipeptide repeats and ultimately, enhanced DPR mediated cell toxicity. These data not only reveal a new mechanistic link between the canonical splice machinery and non-canonical G4C2 repeat RNA metabolism, but further provide direct support for the recently proposed hypothesis that RNA foci are not toxic per se and rather act as a protective storage intermediate, counteracting the formation of toxic DPRs.

We believe that the new insights into C9ORF72 repeat RNA biology and the identified RNA foci modulating compounds will be of great interest to the ALS field that is in strong need for better tools and models to enable therapeutic discovery. In our work we provide tool compounds for controlled interference with the G4C2 repeat RNA lifecycle, including analogs that are commercially available. These tools and mechanistic insights will greatly facilitate the development of preclinical models, further dissection of C9ORF72 pathobiology as well as research on related RNA repeat expansion disorders. Importantly, our data warrant caution on therapeutic strategies targeting the RNA foci due to their potential storage function, counteracting translation into toxic DPRs. Instead, our data support modulation of nuclear export via SRSF1 or SR protein kinases as possible targets for future pharmacological drug discovery – whereas our screen provides precedence that pharmaceutical high-throughput screens in patient iPS neurons are feasible.

Acknowledgements

We would like to thank Dr Guy Rouleau, Dr Patrick Dion and McGill University for kindly providing access to ALS patient material. We also express our gratitude to the patients for donating samples to enable this research. In addition, we like to thank Juan Valcárcel, Maria Carmo-Fonseca, Yuanjing Liu, and Chandra Vargese for insightful discussion and advice.

Author contributions: M.J.L., J.H., M.L., D.G., M.N., and N.M.K. devised the study. M.J.L. and N.M.K. have written the manuscript. C.Ma., I.P.M.G., and M.M. generated the patient derived iNeurons. C.H., C.Ma., C.Me., N.P., and D.T. characterized lines used in this study, F.Z. performed the electrophysiology. P.G.K., N.M.K., C.H., and N.P. designed repeat constructs and cloning strategies, C.H. and N.P. generated stable repeat expressing lines. The HTS was a team effort. D.T. developed the FISH assay, M.P. ran the high-throughput screen, using iNeurons provided at scale by C.H., and D.G. supervised the screen. The screening collection was mainly collated by P.F., D.G., C.G.K., J.H., A.M.K., M.L., M.N., and N.M.K. Compound stratification based on HTS results was mainly done by D.G., A.M.K., J.H., M.L., M.N., and N.M.K.

Strategy for compound progression was devised mainly by D.G., P.G.K., C.G.K., A.M.K., M.L., J.H., M.N., and N.M.K., experiments for compound progression were performed mainly by M.P., D.T., C.H. S.M.L. analogs were provided by K.B., R.M., H.S., and J.H.. ASO.s were designed by J.W. and synthesized by D.K. RNASeq experiments were performed by M.J.L., E.G., and G.R. RNASeq data were analyzed by M.B., G.d.A., and C.G.K. Recombinant SRSF1 protein was generated by A.C. and F.A., FCS experiments were done by N.R. (hnRNPH) and N.M.K. (SRSF1). MS proteomics was performed and analyzed by M.J.L. and H.V. H.M. generated the RAN stable cells and optimized the DPR detection assays under the guidance of M.H. and F.E. RAN experiments collectively were performed by M.J.L., H.M., R.D., and M.S. Experiments with iPS motor neurons were performed by R.D. under supervision of K.T. Single molecule FISH experiments were designed and performed by V.B.R. and J.C. All other experiments were performed and designed by M.J.L. All authors read and approved the final manuscript.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

The following authors are or were employees of the Novartis Institutes of Biomedical Research at the time this work was performed: Maartje Luteijn, Dominic Trojer, Hicham Mahboubi, Nicolas Pizzato, Martin Pfeifer, Hans Voshol, Elisa Giorgetti, Carole Manneville, Isabelle P.M. Garnier, Matthias Müller, Fanning Zeng, Kathrin Buntin, Roger Markwalder, Harald Schröder, Jan Weiler, Dora Khar, Tim Schuhmann, Paul J. Groot-Kormelink, Caroline Gubser Keller, Pierre Farmer, Angela MacKay, Martin Beibel, Guglielmo Roma, Giovanni D'Ario, Claudia Merkl, Michael Schebesta, Marc Hild, Fiona Elwood, Mark Labow, Daniela Gabriel, Mark Nash1, Jürg Hunziker and Nicole C. Meisner-Kober.

Funding

European Funds for Regional Development (Grant EVTT, Award Number: 'EFRE/IWB 20102-F1900731-KZP'); Salzburger Landesregierung (Grants EVTT-BPro and CONSONANT, Award Numbers: 'WISS2025 F2200397-KZP', 'WISS2025 P1812596'). Funding to pay the Open Access publication charges for this article was provided by private (Novartis).

Data availability

The raw data of the RNAseq experiments (cell lines FTD, HD1 and HD2) has been uploaded to SRA under the project code PRJNA843992 (Temporary Submission ID: SUB11547954 Release date: 2023-06-01). The ALSG9 cell line is subject to restrictions in publication of full transcriptomic data including SNPs as restricted by the patient consent form. The gene level count data were generated as described in https://pubmed.ncbi.nlm.nih.gov/27302131/ (reference genome 38, Ensembl version 78) and are provided in Supplementary data S5.

The full data of the MS-proteomics can be accessed via PRIDE:

PRIDE submission #587817: Content: SF3B1 SILAC pulldown (16/16 files); Project Accession PXD034093; (Supplementary Data Set S1)

PRIDE submission #587811: Content: SRSF1 SILAC pulldowns from 2 subcellular fractions (32/32 files); Project Accession PXD034092; (Supplementary Data Set S2)

PRIDE submission #586786: Content: G4C2 pulldown (30/32 files); Project Accession PXD034016;

PRIDE submission #587874: Content: 2 missing files from G4C2 pulldown (2/32 files); Project Accession PXD034097;

The raw data of the compound structure determination by NMR are included in Supplementary file 1. The structures of all compounds used in the RNA foci screen (Fig. 1 and Supplementary Fig. S1) are not made publicly available due to them containing information on substances that are not pursued in the manuscript and proprietary and confidential to Novartis. All other data sets used/processed during this study are available upon reasonable request from the corresponding authors [MJL, JH, NMK]. All DNA constructs, and SML compounds can be made available upon request based on material transfer agreements. ALS/FTD patient and HD cell lines should be requested through the original providers as specified.

References

1. Westeneng HJ, Debray TPA, Visser AE *et al.* Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol* 2018. **17**:423–33. https://doi.org/10.1016/S1474-4422(18)30089-9
2. Wales S, Kiernan DSc AM C, Cheah MBiostat BC *et al.* Seminar amyotrophic lateral sclerosis. *Lancet* 2011;**377**:423–33.
3. DeJesus-Hernandez M, Mackenzie IR, Boeve BF *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 2011. **72**(2):245–56 https://doi.org/10.1016/j.neuron.2011.09.011
4. Renton AE, Majounie E, Waite A *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 2011;**72**:257–68. https://doi.org/10.1016/j.neuron.2011.09.010
5. Majounie E, Renton AE, Mok K *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* 2012. **11**:323–30. https://doi.org/10.1016/S1474-4422(12)70043-1
6. Beck J, Poulter M, Hensman D *et al.* Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am J Hum Genet* 2013;**92**:345–53. https://doi.org/10.1016/j.ajhg.2013.01.011
7. van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E *et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol* 2013;**12**:978–88. https://doi.org/10.1016/S1474-4422(13)70210-2
8. Balendra R, Isaacs AM. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nat Rev Neurol* 2018. **14**:544–58 https://doi.org/10.1038/s41582-018-0047-2
9. Mizielinska S, Lashley T, Norona FE *et al.* C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol* 2013;**126**:845–57. https://doi.org/10.1007/s00401-013-1200-z

10. Fratta P, Mizielinska S, Nicoll AJ *et al.* C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci Rep* 2012. **2**:1016. https://doi.org/10.1038/srep01016
11. Zhang K, Donnelly CJ, Haeusler AR *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* 2015;**525**:56–61. https://doi.org/10.1038/nature14973
12. Wang S, Latallo MJ, Zhang Z *et al.* Nuclear export and translation of circular repeat-containing intronic RNA in C9ORF72-ALS/FTD. *Nat Commun* 2021. **12**:4908. https://doi.org/10.1038/s41467-021-25082-9
13. Cooper-Knock J, Higginbottom A, Stopford MJ *et al.* Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol* 2015. **130**:63–75. https://doi.org/10.1007/s00401-015-1429-9
14. Pettersson OJ, Aagaard L, Jensen TG *et al.* Molecular mechanisms in DM1 - A focus on foci. *Nucleic Acids Res* 2015. **43**:2433–41. https://doi.org/10.1093/nar/gkv029
15. Kumar V, Hasan GM, Hassan MI. Unraveling the role of RNA mediated toxicity of C9orf72 repeats in C9-FTD/ALS. *Front Neurosci* 2017;**11**:2433–41. https://doi.org/10.3389/fnins.2017.00711
16. Swinnen B, Robberecht W, Van Den Bosch L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J* 2020. **39**:e101112. https://doi.org/10.15252/embj.2018101112
17. Li J, Jaiswal MK, Chien J-F *et al.* Divergent single cell transcriptome and epigenome alterations in ALS and FTD patients with C9orf72 mutation. *Nat Commun* 2023;**14**:5714. https://doi.org/10.1038/s41467-023-41033-y
18. Hautbergue GM, Castelli LM, Ferraiuolo L *et al.* SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nat Commun* 2017. **8**:16063. https://doi.org/10.1038/ncomms16063
19. Kawabe Y, Mori K, Yamashita T *et al.* The RNA exosome complex degrades expanded hexanucleotide repeat RNA in C9orf72 FTLD /ALS. *EMBO J* 2020. **39**:e102700. https://doi.org/10.15252/embj.2019102700
20. Ash PEA, Bieniek KF, Gendron TF *et al.* Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 2013;**77**:639–46. https://doi.org/10.1016/j.neuron.2013.02.004
21. Gendron TF, Bieniek KF, Zhang YJ *et al.* Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* 2013;**126**:829–44. https://doi.org/10.1007/s00401-013-1192-8
22. Jiang J, Zhu Q, Gendron TF *et al.* Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* 2016;**90**:535–50. https://doi.org/10.1016/j.neuron.2016.04.006
23. Lagier-Tourenne C, Baughn M, Rigo F *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci USA* 2013;**110**:E4530–9. https://doi.org/10.1073/pnas.1318835110
24. Mori K, Arzberger T, Grässer FA *et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol* 2013;**126**:881–93. https://doi.org/10.1007/s00401-013-1189-3
25. Mori K, Weng SM, Arzberger T *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* 2013. **339**:1335–8. https://doi.org/10.1126/science.1232927
26. Zu T, Liu Y, Bañez-Coronel M *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci USA* 2013. **110**:E4968–77. https://doi.org/10.1073/pnas.1315438110
27. He F, Flores BN, Krans A *et al.* The carboxyl termini of RAN translated GGGGCC nucleotide repeat expansions modulate toxicity in models of ALS/FTD. *Acta Neuropathol Commun* 2020. **8**:122. https://doi.org/10.1186/s40478-020-01002-8
28. Lee YB, Baskaran P, Gomez-Deza J *et al.* C9orf72 poly GA RAN-translated protein plays a key role in amyotrophic lateral sclerosis via aggregation and toxicity. *Hum Mol Genet* 2017. **30**:318–20. https://doi.org/10.1093/hmg/ddx350
29. Mizielinska S, Grönke S, Niccoli T *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* 2014. **345**:1192–4. https://doi.org/10.1126/science.1256800
30. Park JH, Chung CG, Seo J *et al.* C9orf72-associated arginine-rich dipeptide repeat proteins reduce the number of golgi outposts and dendritic branches in drosophila neurons. *Mol Cells* 2020. **43**:821–30. https://doi.org/10.14348/molcells.2020.0130
31. Wen X, Tan W, Westergard T *et al.* Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate invitro and invivo neuronal death. *Neuron* 2014;**84**:1213–25. https://doi.org/10.1016/j.neuron.2014.12.010
32. Zhang YJ, Gendron TF, Grima JC *et al.* C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci* 2016;**19**:668–77. https://doi.org/10.1038/nn.4272
33. Freibaum BD, Lu Y, Lopez-Gonzalez R *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* 2015;**525**:129–33. https://doi.org/10.1038/nature14974
34. Xi Z, van Blitterswijk M, Zhang M *et al.* Jump from pre-mutation to pathologic expansion in C9orf72. *Am Hum Genet* 2015;**96**:962–70. https://doi.org/10.1016/j.ajhg.2015.04.016
35. Lopez-Gonzalez R, Lu Y, Gendron TF *et al.* Poly(GR) in C9ORF72-related ALS/FTD compromises mitochondrial function and increases oxidative stress and DNA damage in iPSC-derived motor neurons. *Neuron* 2016;**92**:383–91. https://doi.org/10.1016/j.neuron.2016.09.015
36. Conlon EG, Lu L, Sharma A *et al.* The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *eLife* 2016. **5**:e17820. https://doi.org/10.7554/eLife.17820
37. Cooper-Knock J, Walsh MJ, Higginbottom A *et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* 2014;**137**:2040–51. https://doi.org/10.1093/brain/awu120
38. Kwon I, Xiang S, Kato M *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (1979)* 2014;**345**:1139–45. https://doi.org/10.1126/science.1254917
39. Prudencio M, Belzil VV, Batra R *et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci* 2015;**18**:1175–82. https://doi.org/10.1038/nn.4065
40. Nguyen L, Montrasio F, Pattamatta A *et al.* Antibody therapy targeting RAN proteins rescues C9 ALS/FTD phenotypes in C9orf72 mouse model. *Neuron* 2020. **105**:645–62. https://doi.org/10.1016/j.neuron.2019.11.007
41. Vatsavayai SC, Yoon SJ, Gardner RC *et al.* Timing and significance of pathological features in *C9orf72* expansion-associated frontotemporal dementia. *Brain* 2016;**139**:3202–16. https://doi.org/10.1093/brain/aww250
42. Peters OM, Cabrera GT, Tran H *et al.* Human C9ORF72 hexanucleotide expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. *Neuron* 2015;**88**:902–9. https://doi.org/10.1016/j.neuron.2015.11.018
43. Donnelly CJ, Zhang PW, Pham JT *et al.* RNA-induced toxicity from the C9ORF72 ALS/FTD repeat expansion is mitigated by antisense intervention. *Amyotroph Lateral Scler Frontotemporal*

Degener 2013. **80**:415–28. https://doi.org/10.1016/j.neuron.2013.10.015
44. Liang C, Shao Q, Zhang W *et al.* Smcr8 deficiency disrupts axonal transport-dependent lysosomal function and promotes axonal swellings and gain of toxicity in C9ALS/FTD mouse models. *Hum Mol Genet* 2019. **29**:1056. https://doi.org/10.1093/hmg/ddz230
45. Shi Y, Lin S, Staats KA *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med* 2018. **24**:313–25. https://doi.org/10.1038/nm.4490
46. Su MY, Fromm SA, Zoncu R *et al.* Structure of the C9orf72 ARF GAP complex that is haploinsufficient in ALS and FTD. *Nature* 2020. **585**:251–5. https://doi.org/10.1038/s41586-020-2633-x
47. Zhu Q, Jiang J, Gendron TF *et al.* Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. *Nat Neurosci* 2020;**23**:615–24. https://doi.org/10.1038/s41593-020-0619-5
48. Shi Y, Lin S, Staats KA *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med* 2018;**24**:313–25. https://doi.org/10.1038/nm.4490
49. Jovičič A, Mertens J, Boeynaems S *et al.* Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci* 2015;**18**:1226–9. https://doi.org/10.1038/nn.4085
50. Moens TG, Mizielinska S, Niccoli T *et al.* Sense and antisense RNA are not toxic in Drosophila models of C9orf72-associated ALS/FTD. *Acta Neuropathol* 2018. **135**:445-57. https://doi.org/10.1007/s00401-017-1798-3
51. Swinnen B, Bento-Abreu A, Gendron TF *et al.* A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol* 2018. **135**:427–43. https://doi.org/10.1007/s00401-017-1796-5
52. DeJesus-Hernandez M, Finch NCA, Wang X *et al.* In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers. *Acta Neuropathol* 2017. **134**:255-69. https://doi.org/10.1007/s00401-017-1725-7
53. Sareen D, O'Rourke JG, Meera P *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* 2013. **5**:208ra149. https://doi.org/10.1126/scitranslmed.3007529
54. Lagier-Tourenne C, Baughn M, Rigo F *et al.* Targeted degradation of sense and antisense *C9orf72* RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci USA* 2013. **110**:E4530–9. https://doi.org/10.1073/pnas.1318835110
55. Liu Y, Dodart J-C, Tran H *et al.* Variant-selective stereopure oligonucleotides protect against pathologies associated with C9orf72-repeat expansion in preclinical models. *Nat Commun* 2021;**12**:255–69. https://doi.org/10.1038/s41467-021-21112-8
56. Ly CV, Miller TM. Emerging antisense oligonucleotide and viral therapies for amyotrophic lateral sclerosis. *Curr Opin Neurol* 2018. **31**:648–54. https://doi.org/10.1097/WCO.0000000000000594
57. Tran H, Moazami MP, Yang H *et al.* Suppression of mutant C9orf72 expression by a potent mixed backbone antisense oligonucleotide. *Nat Med* 2022;**28**:117–24. https://doi.org/10.1038/s41591-021-01557-6
58. van den Berg LH, Rothstein JD, Shaw PJ *et al.* Safety, tolerability, and pharmacokinetics of antisense oligonucleotide BIIB078 in adults with C9orf72-associated amyotrophic lateral sclerosis: a phase 1, randomised, double blinded, placebo-controlled, multiple ascending dose study. *Lancet Neurol* 2024;**23**:901–12. https://doi.org/10.1016/S1474-4422(24)00216-3
59. Liu Y, Andreucci A, Iwamoto N *et al.* Preclinical evaluation of WVE-004, an investigational stereopure oligonucleotide for the treatment of C9orf72-associated ALS or FTD. *Mol Ther - Nucleic Acids* 2022;**28**:558–70. https://doi.org/10.1016/j.omtn.2022.04.007
60. Rothstein JD, Baskerville V, Rapuri S *et al.* G_2C_4 targeting antisense oligonucleotides potently mitigate TDP-43 dysfunction in human C9orf72 ALS/FTD induced pluripotent stem cell derived neurons. *Acta Neuropathol* 2023;**147**:1. https://doi.org/10.1007/s00401-023-02652-3
61. Ababneh NA, Scaber J, Flynn R *et al.* Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Hum Mol Genet* 2020;**29**:2200–17. https://doi.org/10.1093/hmg/ddaa106
62. Meijboom KE, Abdallah A, Fordham NP *et al.* CRISPR/Cas9-mediated excision of ALS/FTD-causing hexanucleotide repeat expansion in C9ORF72 rescues major disease mechanisms *in vivo* and *in vitro*. *Nat Commun* 2022;**13**:6286. https://doi.org/10.1038/s41467-022-33332-7
63. Piao X, Meng D, Zhang X *et al.* Dual-gRNA approach with limited off-target effect corrects C9ORF72 repeat expansion in vivo. *Sci Rep* 2022;**12**:5672. https://doi.org/10.1038/s41598-022-07746-8
64. Simone R, Balendra R, Moens TG *et al.* G-quadruplex-binding small molecules ameliorate *C9orf72* FTD/ALS pathology *in vitro* and *in vivo*. *EMBO Mol Med* 2018. **10**:22–31. https://doi.org/10.15252/emmm.201707850
65. Su Z, Zhang Y, Gendron TF *et al.* Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* 2014. **83**:1043–50. https://doi.org/10.1016/j.neuron.2014.07.041
66. Bush JA, Aikawa H, Fuerst R *et al.* Ribonuclease recruitment using a small molecule reduced c9ALS/FTD r(G4C2) repeat expansion *in vitro* and *in vivo* ALS models. *Sci Transl Med* 2021. **13**:eabd5991 https://doi.org/10.1126/SCITRANSLMED.ABD5991
67. Green KM, Sheth UJ, Flores BN *et al.* High-throughput screening yields several small-molecule inhibitors of repeat-associated non-AUG translation. *J Biol Chem* 2019. **294**:18624–38. https://doi.org/10.1074/jbc.RA119.009951
68. Staats KA, Seah C, Sahimi A *et al.* Small molecule inhibition of PIKFYVE kinase rescues gain- and loss-of-function C9ORF72 ALS/FTD disease processes *in vivo*. bioRxiv, https://doi.org/10.1101/685800, 28 June 2019, preprint: not peer reviewed.
69. Shi Y, Lin S, Staats KA *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med* 2018;**24**:313–25. https://doi.org/10.1038/nm.4490
70. Dafinca R, Barbagallo P, Farrimond L *et al.* Impairment of mitochondrial calcium buffering links mutations in C9ORF72 and TARDBP in iPS-derived motor neurons from patients with ALS/FTD. *Stem Cell Rep* 2020. **14**:892–908. https://doi.org/10.1016/j.stemcr.2020.03.023
71. Ababneh NA, Scaber J, Flynn R *et al.* Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Hum Mol Genet* 2020. **29**:2200–17. https://doi.org/10.1093/hmg/ddaa106
72. Schirle M, Jenkins JL. Identifying compound efficacy targets in phenotypic drug discovery. *Drug Discovery Today* 2016. **21**:82–9. https://doi.org/10.1016/j.drudis.2015.08.001
73. Samatanga B, Cléry A, Barraud P *et al.* Comparative analyses of the thermodynamic RNA binding signatures of different types of RNA recognition motifs. *Nucleic Acids Res* 2017. **45**:6037–50. https://doi.org/10.1093/nar/gkx136
74. Anczuków O, Akerman M, Cléry A *et al.* SRSF1-Regulated alternative splicing in breast cancer. *Mol Cell* 2015. **60**:105–17. https://doi.org/10.1016/j.molcel.2015.09.005
75. Ritchie ME, Phipson B, Wu D *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies.

Nucleic Acids Res 2015. **43**:e47. https://doi.org/10.1093/nar/gkv007
76. Haeusler AR, Donnelly CJ, Rothstein JD. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nat Rev Neurosci* 2016;**17**:e47.
77. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008. **26**:1367–72. https://doi.org/10.1038/nbt.1511
78. Wang Y, Zhu W, Levy DE. Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time RT-PCR. *Methods* 2006. **39**:356–62. https://doi.org/10.1016/j.ymeth.2006.06.010
79. Schürz M, Danmayr J, Jaritsch M *et al.* EVAnalyzer: high content imaging for rigorous characterisation of single extracellular vesicles using standard laboratory equipment and a new open-source ImageJ/Fiji plugin. *J Extracell Vesicles* 2022;**11**.e12282.
80. Maury Y, Côme J, Piskorowski RA *et al.* Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* 2015. **33**:89–96. https://doi.org/10.1038/nbt.3049
81. Dafinca R, Scaber J, Ababneh N *et al.* C9orf72 Hexanucleotide expansions are associated with altered endoplasmic reticulum calcium homeostasis and stress granule formation in induced pluripotent stem cell-derived neurons from patients with amyotrophic lateral sclerosis and frontotemporal demen. *Stem Cells* 2016. **34**:2063–78. https://doi.org/10.1002/stem.2388
82. Zhang Y, Pak CH, Han Y *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 2013. **78**:785–98. https://doi.org/10.1016/j.neuron.2013.05.029
83. Zeier Z, Esanov R, Belle KC *et al.* Bromodomain inhibitors regulate the C9ORF72 locus in ALS. *Exp Neurol* 2015;**271**:241–50. https://doi.org/10.1016/j.expneurol.2015.06.017
84. Corrionero A, Miñana B, Valcárcel J. Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes Dev* 2011;**25**:241–50.
85. Vigevani L, Gohr A, Webb T *et al.* Molecular basis of differential 3′ splice site sensitivity to anti-tumor drugs targeting U2 snRNP. *Nat Commun* 2017. **8**:2100. https://doi.org/10.1038/s41467-017-02007-z
86. Effenberger KA, Urabe VK, Jurica MS. Modulating splicing with small molecular inhibitors of the spliceosome. *WIREs RNA* 2017;**8**. 10.1002/wrna.1381. https://doi.org/10.1002/wrna.1381
87. Ghosh AK, Chen ZH. Enantioselective syntheses of FR901464 and spliceostatin A: potent inhibitors of spliceosome. *Org Lett* 2013. **15**:5088–91. https://doi.org/10.1021/ol4024634
88. Roybal GA, Jurica MS. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. *Nucleic Acids Res* 2010;**38**:6664–72. https://doi.org/10.1093/nar/gkq494
89. Yoshimoto R, Chhipi-Shrestha JK, Schneider-Poetsch T *et al.* Spliceostatin A interaction with SF3B limits U1 snRNP availability and causes premature cleavage and polyadenylation. *Cell Chem Biol* 2021;**28**:1356–1365.e4. https://doi.org/10.1016/j.chembiol.2021.03.002
90. Kaida D, Motoyoshi H, Tashiro E *et al.* Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol* 2007;**3**:576–83. https://doi.org/10.1038/nchembio.2007.18
91. Effenberger KA, Anderson DD, Bray WM *et al.* Coherence between cellular responses and in vitro splicing inhibition for the anti-tumor drug pladienolide b and its analogs. *J Biol Chem* 2014;**289**:1938–47. https://doi.org/10.1074/jbc.M113.515536
92. Wu G, Fan L, Edmonson MN *et al.* Inhibition of SF3B1 by molecules targeting the spliceosome results in massive aberrant exon skipping. *RNA* 2018;**24**:1056–66. https://doi.org/10.1261/rna.065383.117
93. Vigevani L, Gohr A, Webb T *et al.* Molecular basis of differential 3′ splice site sensitivity to anti-tumor drugs targeting U2 snRNP. *Nat Commun* 2017;**8**:2100. https://doi.org/10.1038/s41467-017-02007-z
94. Gao Y, Koide K. Chemical perturbation of Mcl-1 pre-mRNA splicing to induce apoptosis in cancer cells. *ACS Chem Biol* 2013. **8**:895–900. https://doi.org/10.1021/cb300602j
95. Carvalho T, Martins S, Rino J *et al.* Pharmacological inhibition of the spliceosome subunit SF3b triggers exon junction complex-independent nonsense-mediated decay. *J Cell Sci* 2017. **130**:1519–31. https://doi.org/10.1242/jcs.202200
96. Kwon I, Xiang S, Kato M *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (1979)* 2014;**345**:1519–31.
97. Hocine S, Raymond P, Zenklusen D *et al.* Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods* 2013;**10**:119–21.
98. Castelli LM, Cutillo L, Souza CDS *et al.* SRSF1-dependent inhibition of C9ORF72-repeat RNA nuclear export: genome-wide mechanisms for neuroprotection in amyotrophic lateral sclerosis. *Mol Neurodegener* 2021;**16**:53.
99. Geng Y, Liu C, Xu N *et al.* Crystal structure of a tetrameric RNA G-quadruplex formed by hexanucleotide repeat expansions of C9orf72 in ALS/FTD. *Nucleic Acids Res* 2024;**52**:7961–70.
100. Raguseo F, Wang Y, Li J *et al.* The ALS/FTD-related C9orf72 hexanucleotide repeat expansion forms RNA condensates through multimolecular G-quadruplexes. *Nat Commun* 2023;**14**:8272.
101. Haeusler AR, Donnelly CJ, Periz G *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 2014;**507**:195–200.
102. Stutz F, Bachi A, Doerks T *et al.* REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* 2000. **6**:638–50. https://doi.org/10.1017/S1355838200000078
103. Akerman M, Fregoso OI, Das S *et al.* Differential connectivity of splicing activators and repressors to the human spliceosome. *Genome Biol* 2015. **16**:119. https://doi.org/10.1186/s13059-015-0682-5
104. Moens TG, Mizielinska S, Niccoli T *et al.* Sense and antisense RNA are not toxic in Drosophila models of C9orf72-associated ALS/FTD. *Acta Neuropathol* 2018. **135**:445–57. https://doi.org/10.1007/s00401-017-1798-3
105. Kramer NJ, Haney MS, Morgens DW *et al.* CRISPR-Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity. *Nat Genet* 2018;**50**:445–57.
106. Corman A, Jung B, Häggblad M *et al.* A chemical screen identifies compounds limiting the toxicity of C9ORF72 dipeptide repeats. *Cell Chem Biol* 2019. **26**:235–43. https://doi.org/10.1016/j.chembiol.2018.10.020
107. Teng T, Tsai JH, Puyang X *et al.* Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A-SF3b complex. *Nat Commun* 2017;**8**:235–243.e5. https://doi.org/10.1038/ncomms15522
108. Daguenet E, Dujardin G, Valcarcel J. The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO Rep* 2015;**16**:1640–55. https://doi.org/10.15252/embr.201541116
109. Kotake Y, Sagane K, Owa T *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat Chem Biol* 2007. **3**:570–5. https://doi.org/10.1038/nchembio.2007.16
110. Mizielinska S, Grönke S, Niccoli T *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science (1979)* 2014;**1192**:570–5.
111. Papasaikas P, Tejedor JR, Vigevani L *et al.* Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* 2015;**57**:7–22. https://doi.org/10.1016/j.molcel.2014.10.030
112. Michlewski G, Sanford JR, Cáceres JF. The splicing factor SF2/ASF regulates translation initiation by enhancing

phosphorylation of 4E-BP1. *Mol Cell* 2008. **30**:179–89. https://doi.org/10.1016/j.molcel.2008.03.013
113. Tran H, Almeida S, Moore J *et al.* Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a drosophila model of C9ORF72 FTD/ALS. *Neuron* 2015;**87**:1207–14. https://doi.org/10.1016/j.neuron.2015.09.015
114. Geevasinga N, Menon P, Nicholson GA *et al.* Cortical function in asymptomatic carriers and patients with C9orf72 amyotrophic lateral sclerosis. *JAMA Neurol* 2015;**72**:1268–74. https://doi.org/10.1001/jamaneurol.2015.1872
115. Kim HJ, Taylor JP. Lost in transportation: nucleocytoplasmic transport defects in ALS and other neurodegenerative diseases. *Neuron* 2017;**96**:285–97. https://doi.org/10.1016/j.neuron.2017.07.029
116. Vasques JF, Mendez-Otero R, Gubert F. Modeling ALS using iPSCs: is it possible to reproduce the phenotypic variations observed in patients *in vitro*? *Regen Med* 2020. **15**:1919–33. https://doi.org/10.2217/rme-2020-0067

Received: December 20, 2024. **Revised:** March 3, 2025. **Editorial Decision:** March 5, 2025. **Accepted:** April 7, 2025