



This is a repository copy of *C9orf72 Expansion Disrupts ATM-mediated Chromosomal Break Repair*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/118606/>

Version: Accepted Version

Article:

Walker, C., Herranz-Martin, S., Karyka, E. et al. (19 more authors) (2017) C9orf72 Expansion Disrupts ATM-mediated Chromosomal Break Repair. *Nature Neuroscience*, 20 (9). pp. 1225-1235. ISSN 1097-6256

<https://doi.org/10.1038/nn.4604>

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1038/nn.4604>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

1 ***C9orf72* Expansion Disrupts ATM-mediated Chromosomal Break Repair**

2

3 Callum Walker^{1,2,3¶}, Saul Herranz-Martin^{2¶}, Evangelia Karyka^{1,2,3}, Chunyan Liao³, Katherine
4 Lewis², Waheba Elsayed^{3,4}, Vera Lukashchuk², Shih-Chieh Chiang³, Swagat Ray³, Pdraig J.
5 Mulcahy², Mateusz Jurga³, Ioannis Tsagakis², Tommaso Iannitti², Jayanth Chandran², Ian
6 Coldicott², Kurt J. De Vos², Mohamed K. Hassan^{3,4}, Adrian Higginbottom², Pamela J. Shaw²,
7 Guillaume M. Hautbergue², Mimoun Azzouz^{1,2*#}, Sherif F. El-Khamisy^{1,3,4*#}

8

9 ¹SITraN and Krebs Institutes, Neurodegeneration and Genome Stability Group, University of
10 Sheffield, UK

11 ²Sheffield Institute for Translational Neuroscience, University of Sheffield, 385a Glossop
12 Road, S10 2HQ, Sheffield, UK

13 ³Krebs and Sheffield Institutes for Nucleic Acids, Department of Molecular Biology and
14 Biotechnology, Firth Court, University of Sheffield, S10 2TN, Sheffield, UK

15 ⁴Center for Genomics, Helmy Institute for Medical Sciences, Zewail City of Science and
16 Technology, Giza, Egypt.

17

18 [¶]Joint first authors

19 [#]Joint senior authors made equal contribution

20

21 *Correspondence should be addressed to Sherif El-Khamisy (s.el-khamisy@sheffield.ac.uk),
22 Tel. +44 (0) 114 2222 791 and Mimoun Azzouz (m.azzouz@sheffield.ac.uk), Tel. +44 (0) 114
23 2222 238.

24

25

26

27 **A hexanucleotide repeat expansion represents the most common genetic cause of**
28 **amyotrophic lateral sclerosis (ALS) and frontotemporal dementia, though the**
29 **mechanisms by which the expansion cause neurodegeneration are poorly understood. We**
30 **report elevated levels of DNA/RNA hybrids (R-loops) and double-strand breaks (DSBs) in**
31 **rodent neurons, human cells, and in *C9orf72*-ALS patient spinal cord tissues.**
32 **Accumulation of endogenous DNA damage is concomitant with defective ATM-mediated**
33 **DNA repair signalling and accumulation of protein-linked DNA breaks. We further**
34 **reveal that defective ATM-mediated DNA repair is a consequence of p62 accumulation,**
35 **which impairs H2A ubiquitylation and perturbs ATM signalling. Adeno-associated virus-**
36 **mediated expression of *C9orf72*-related RNA and dipeptide repeats in the murine central**
37 **nervous system causes elevated DSBs, ATM defects, and triggers neurodegeneration.**
38 **These findings identify R-Loops, DSBs, and defective ATM-mediated repair as**
39 **pathological consequences of *C9orf72* expansions, and suggest that *C9orf72*-linked**
40 **neurodegeneration is driven, at least in part, by genomic instability.**

41

42 Short tandem nucleotide repeats and microsatellites are common features of mammalian
43 genomes. Expansion of a hexanucleotide G4C2 repeat in the non-coding region of chromosome
44 9 open reading frame 72 (*C9orf72*) is the most common genetic cause for amyotrophic lateral
45 sclerosis (ALS) and frontotemporal dementia (FTD). Growing evidence suggests that *C9orf72*
46 repeat expansions also contribute to a wide spectrum of neurodegenerative diseases such as
47 Alzheimer's, Huntington's, multiple sclerosis, Parkinson's disease and cerebellar ataxias¹.
48 Approximately half of non-pathogenic *C9orf72* alleles possess two G4C2 repeats and the
49 remaining half ranges from 2 to 25 repeats². The pathogenic expanded repeat length, on the
50 other hand, varies from tens to thousands³. *C9orf72* expansions are bidirectionally transcribed
51 leading to the formation of intracellular sense and antisense RNA repeat expansion foci (RRE).

52 Moreover, the transcripts are prone to repeat-associated non-ATG (RAN) translation producing
53 dipeptide repeat proteins (DPRs). Although a molecular understanding of *C9orf72* pathological
54 phenotypes are beginning to emerge, the mechanisms by which the G4C2 repeat expansions
55 cause ALS/FTD are not clear.

56

57 During the transcription of repetitive sequences, the nascent RNA is prone to hybridisation with
58 the DNA template strand, displacing the complementary DNA strand and producing a three-
59 stranded nucleic acid structure called R-loops⁴. R-loops primarily occur at GC-rich
60 transcription sites, since guanine-rich RNA: cytosine-rich DNA hybrids are thermodynamically
61 more stable than the respective DNA: DNA duplex⁵. Once formed, R-loops can be very stable
62 structures, as they are bound together by Watson-Crick base pairing. These transcription by-
63 products are a major threat to genome stability, since they are prone to DNA breakage⁶. Given
64 the pure GC nature of the *C9orf72* repeat expansions and their propensity to form R-loops *in*
65 *vitro*⁷, we hypothesised that R-loop-mediated genome instability may play a role in
66 neurodegeneration linked to *C9orf72* repeats. To test this, we transfected MRC5 cells with 10
67 or 102 RREs and visualised R-loops using R-loop specific S9.6 antibodies. We concomitantly
68 visualised RNA foci using fluorescence in situ hybridization (FISH). Expression of 102 RREs
69 led to prominent RNA foci and triggered an approximate 7-fold increase in R-loop levels
70 compared to cells transfected with a shorter expansion containing 10 RREs, which also
71 displayed fewer RNA foci (Fig. 1a, $p=0.009$). Interestingly, R-loops and RNA foci co-localised
72 in cells expressing 102 RREs, suggesting a physical relationship. The R-loop signal was
73 specific since it disappeared following addition of RNase H1, an R-loop specific resolvase
74 (Fig. 1a).

75 We next evaluated the role of poly-GA DPRs in mediating R-loop formation. Transfection of
76 MRC5 cells with either 34 or 69 DPRs revealed a length-dependent predisposition to dipeptide

77 aggregates, in which DPRs were more abundant in cells transfected with 69 DPRs (Fig. 1b).
78 Expression of 34 DPRs led to ~4 R-loop foci per cell and expression of 69 DPRs led to ~8 R-
79 loop foci per cell, both of which were considered statistically higher than in control cells (Fig.
80 1b, $p=0.046$ and $p=0.003$, respectively), but were not statistically distinct from each other (Fig.
81 1b, $p=0.1$). These observations suggest that *C9orf72* RNA-repeat expansions and poly-GA
82 DPRs cause an increase in R-loop formation. To test if this is also true in *bona fide* post-mitotic
83 neurons, we transduced rat cortical neurons with adeno-associated serotype 9 (AAV9) viral
84 particles expressing 10, 102 RREs (Fig. 1c) or 34, 69 DPRs (Fig. 1d). Similar to human cells,
85 expression of 102 RREs led to increased R-loop foci compared to control cells and cells
86 transduced with 10 RREs (Fig. 1c, $p=0.018$). Likewise, 34 or 69 DPR expression increased the
87 number of R-loop foci (Fig. 1d; $p=0.001$; $p=0.0002$, respectively), though the difference
88 between 34 and 69 was not statistically significant (Fig. 1d, $p=0.07$). Together, these data show
89 that G4C2 repeat expansions and poly-GA DPRs cause R-loop formation in mammalian cells.

90

91 Persistent accumulation of R-loops causes DNA double-strand break (DSB) formation and
92 genome instability⁸. To test whether G4C2 repeat expansions cause DSBs we co-transfected
93 MRC5 cells with 10, 102 RREs alongside GFP and examined DNA DSBs by immunostaining
94 using antibodies for γ H2AX (Ser-139 phosphorylated histone H2AX), an established marker
95 for DSBs. As anticipated, expression of 102 RREs led to a significant increase in the number of
96 cells exhibiting more than 10 γ H2AX foci, when compared to control cells and cells expressing
97 10 RREs (Fig. 1e, $p=0.005$). Direct quantification of DSBs using the neutral comet assay
98 revealed a similar increase in 102 RRE transfected cells compared to controls (Fig. 1f,
99 $p=0.006$). Similarly, 34 or 69 DPR expression caused a significant increase in DSBs compared
100 to controls, as measured by γ H2AX immunostaining and the neutral comet assay (Fig. 1g,
101 $p=0.001$; Fig. 1h, $p=0.0003$, respectively).

102

103 We next speculated that R-Loops might drive the formation of DSBs in cells expressing
104 *C9orf72* repeat expansions. To directly test this hypothesis, we first assessed whether the
105 elevated levels of R-Loops observed in *C9orf72* cells could be reduced by overexpressing the
106 R-Loop resolution helicase, senataxin (SETX)⁹. Whilst control adenovirus expression of RFP
107 had no detectable impact on R-loops, expression of SETX from the same viral backbone led to
108 a marked reduction of R-Loop foci (Fig 1i,j, *left*). Overexpression of SETX, but not RFP, also
109 reduced the number of γ H2AX-positive cells (Fig. 1i, $p=0.01$; Fig 1j, $p=0.022$), indicating that
110 R-Loops are a major source of *C9orf72* expansion-driven DSBs. Furthermore, SETX
111 expression was able to reduce RRE- and DPR-driven cellular toxicity, as shown by a reduction
112 in the % of cleaved-PARP1-positive cells (Fig. 1k,m; $p=0.009$ and $p=0.038$, respectively) and
113 by a similar reduction in the % of Trypan Blue-positive cells (Fig. 1l,n; $p=0.037$ and $p=0.025$,
114 respectively). The effect of SETX was specific to RRE-102- and 69-V5-positive cells since it
115 did not impact cell viability in control cells expressing RRE-10 or 0-V5 constructs (Fig. 1k-n;
116 $p>0.05$). Taken together, our data demonstrate that *C9orf72* repeat expansions promote R-
117 Loop-driven DSB formation, which contributes to cellular toxicity.

118

119 The growing evidence that links *C9orf72* repeat expansions to cerebellar ataxias^{10,11} prompted
120 us to test whether the activation of the master DNA repair kinase, mutated in ataxia
121 telangiectasia (ATM), is dysregulated in cells expressing *C9orf72* expansions. Upon DNA
122 breakage, ATM becomes activated in a process that involves its autophosphorylation at
123 multiple sites, including serine 1981. This then sets in motion a cascade of downstream events
124 resulting in 53BP1 recruitment and DNA repair¹²⁻¹⁴. While control cells displayed an average
125 of 2 phosphorylated ATM (pATM) foci per nucleus, cells expressing DPRs (Fig. 2a, $p=0.002$)
126 or 102 RREs (Fig. 2b, $p=0.008$) consistently possessed no, or in a few cases only one, focus.

127 The reduced number of pATM foci is unlikely to be due to reduced ATM expression, as shown
128 by Western blotting (Fig. 2c). It was also not due to an artefact of expressing the V5 epitope tag
129 or transfection associated toxicity, since transfection with a V5 empty plasmid or GFP did not
130 result in defective ATM signalling (Supplementary Figure 1). Since the accumulation of
131 oxidative and protein-linked DNA breaks (PDBs) has been shown to cause neurodegeneration
132 in man¹⁵⁻¹⁷, we next examined if pATM foci would form normally following exposure to the
133 topoisomerase I poison camptothecin (CPT) or tert-butyl hydroperoxide (TBH); inducers of
134 PDBs and oxidative DNA breaks, respectively. Treatment with CPT or TBH led to prominent
135 increase in pATM foci in control cells (Fig. 2d). In a marked contrast, cells expressing 34, 69
136 DPRs (Fig. 2d) or 102 RREs (Fig. 2e) were unable to respond to DNA damage to the same
137 extent, consistently showing less pATM foci. We conclude from these experiments that the
138 expression of *C9orf72*-related DPRs and RREs impairs ATM activation.

139

140 An important consequence of defective ATM signalling is the defective accumulation of non-
141 homologous end-joining (NHEJ) repair factors, such as 53BP1, into nuclear foci¹⁸. As
142 expected, expression of 34 or 69 DPRs led to a marked reduction in 53BP1 recruitment to
143 nuclear foci when compared to control cells (Fig. 3a, $p=0.002$). As was the case for pATM, this
144 was not an artefact of expressing the V5 epitope-tag used to visualise DPRs, nor was it due to
145 transfection-related toxicity, as it was not observed in cells transfected with vectors encoding
146 V5 or GFP (Supplementary Figure 2a). Consistent with defective ATM in DPR expressing
147 cells, pre-treatment with the ATM inhibitor Ku55933 significantly impaired 53BP1 foci
148 formation in control cells but not in cells expressing 34 or 69 DPRs (Supplementary Figure
149 2c,d). As observed with DPRs, expression of 102 RREs also led to the suppression of 53BP1
150 foci formation, when compared to mock transfected cells and cells expressing 10 RREs (Fig.
151 3b, $p=0.035$). Strikingly, whilst CPT induced a significant increase in 53BP1 foci in control

152 cells (Fig. 3c, $p=0.009$), expression of 34 or 69 DPRs prevented the CPT-induced increase in
153 53BP1 foci formation (Fig. 3c, $p=0.294$). Similarly, expression of 102 RREs prevented the
154 CPT-induced increase in 53BP1 foci formation (Fig. 3d, $p=0.496$). To test if this was also true
155 in *bona fide* post-mitotic neurons, we transduced rat cortical neurons with AAV9 viral particles
156 expressing 34, 69 DPRs or 10, 102 RREs and examined 53BP1 foci formation following
157 exposure to CPT (Fig. 3e,f). Consistent with results using MRC5 cells, whilst control neurons
158 exhibited ~8 53BP1 foci, those expressing 69 DPRs or 102 RREs showed 3 or 1 foci per cell,
159 respectively (Fig. 3 e,f, $p<0.01$). Consistent with our transfection data, this was not an artefact
160 of viral transduction as cells transduced with AAV9 encoding V5 or GFP vectors possessed
161 normal level of 53BP1 foci (Supplementary Figure 2b). Together, these data demonstrate that
162 the expression of *C9orf72*-related RREs or DPRs suppress 53BP1 recruitment to DSBs. To test
163 whether the expression of *C9orf72*-related products results in defective phosphorylation of
164 downstream ATM-effector proteins, we next examined the phosphorylation of p53, a known
165 ATM target¹⁹. Whilst control cells displayed ~6-fold elevation of p-p53 levels in response to
166 CPT treatment (Fig. 3g, $p=0.037$), cells expressing 34 or 69 DPRs showed no significant
167 increase above baseline levels (Fig. 3g, $p>0.05$). Consistent with the 53BP1 data, we also
168 observed a defect in the recruitment of p-p53 to nuclear foci in DPR expressing cells, which
169 was mimicked in control cells by pre-treatment with the ATM inhibitor (Supplementary Figure
170 2e). Notably, ATM inhibition did not further attenuate p-p53 signalling in DPR-positive cells
171 ($p=0.059$ and $p=0.49$ for 34 and 69 DPRs, respectively). Similarly, cells containing 102 RREs
172 failed to induce p53 phosphorylation following CPT exposure (Fig. 3h, $p=0.536$), whilst mock-
173 transfected cells and those expressing the 10 RREs did show increased p-p53 response to CPT
174 treatment (Fig. 3h; $p=0.0143$ and $p=0.01$, respectively). Finally, in order to gain further
175 evidence for defective ATM signalling, we next examined the accumulation of topoisomerase I
176 (TOP1) – mediated PDBs, known as DNA covalent complexes (TOP1cc). ATM deficiency in

177 primary neural cultures and in rodent models has been shown to cause elevated levels of
178 TOP1cc, interfering with transcription and contributing to neuronal cell death^{20,21}. Consistent
179 with an ATM defect, cells expressing *C9orf72*-related DPRs or RREs exhibited ~4-fold
180 increase in TOP1cc levels compared to controls (Fig. 3i). Taken together, these experiments
181 reveal that *C9orf72*-related DPRs or RREs cause defective ATM signalling.

182

183 To further confirm ATM deficits in primary neural cultures we took advantage of the
184 observations in *Atm* knockout mice and Ataxia Telangiectasia (A-T) patient tissue in which
185 histone deacetylase 4 (HDAC4) is abnormally localised to the nucleus of neurons²². To remain
186 cytoplasmic, HDAC4 requires sustained phosphorylation, which is maintained by an intricate
187 balance between phosphorylation and dephosphorylation. The latter is conducted by protein
188 phosphatase 2A (PP2A), the activity of which is negatively regulated by ATM. Thus, ATM
189 deficiency results in increased PP2A activity, leading to HDAC4 hypophosphorylation and, in
190 turn, its re-localisation to the nucleus. HDAC4 knockdown/inhibition and cytoplasmic HDAC4
191 conferred therapeutic benefit *in vitro* and *in vivo* models of A-T²². To test if the ATM deficits
192 triggered by *C9orf72*-related DPRs would also cause nuclear accumulation of HDAC4, we
193 transduced rat cortical neurons with AAV9 viral particles expressing 69 DPRs and examined
194 the cellular localisation of HDAC4. Whilst HDAC4 appeared cytoplasmic in control neurons, it
195 showed marked nuclear enrichment in neurons expressing 69 DPRs with no detectable change
196 in neuronal morphology (Supplementary Figure 3a,c). Subsequent quantification revealed ~3
197 fold increase in nuclear HDAC4 driven by 69 DPR expression (Supplementary Figure 3b,
198 $p=0.001$). Nuclear accumulation of HDAC4 was not an artefact of viral transduction as neurons
199 transduced with AAV 9 encoding V5 epitope tag or GFP displayed normal cytoplasmic
200 HDAC4. Furthermore, pre-treatment of primary neurons with an ATM inhibitor (Ku 55933)
201 similarly caused nuclear HDAC4 re-localisation (Supplementary Figure 3d,e), suggesting that

202 the observed phenotype is due to ATM deficit.

203

204 We next examined if these observations were also true *in vivo*. To achieve this, AAV9 viral
205 vectors encoding 10, 102 RREs, or 0, 69 V5-tagged DPRs were delivered into the cerebrospinal
206 fluid via cisterna magna of wild type mice at post-natal day 1 (P1), which led to transgene
207 expression in multiple areas of the CNS (Supplementary Figure 4). As was the case *in vitro*,
208 expression of 102 RREs or 69 DPRs in mouse CNS tissue, led to HDAC4 nuclear re-
209 localisation and accumulation of DSBs as measured by increased γ H2AX foci (Fig. 4a-d,
210 $p<0.05$). We next tested if the pronounced DNA repair defect observed *in vivo* would lead to
211 neuronal cell death. Brain extract from mice injected with 69-V5 DPRs displayed ~2-fold
212 increase in PARP1 cleavage compared to mice injected with 0-V5 DPRs (Fig. 4 e, $p=0.045$).
213 To further confirm the loss of neuronal cells in CNS tissue derived from mice expression
214 DPRs, we used NeuN immunohistochemistry to quantify neurons from brainstem sections.
215 Consistent with our Western blotting data, we observed ~20% reduction in neurons from
216 brainstem sections derived from mice injected with 69 DPRs when compared to mice injected
217 with 0-V5 (Fig. 4f; $p=0.048$). Given the loss of neuronal cells, we anticipated that this would
218 translate into a gross functional motor deficit at the whole organismal level. Indeed, aged 6
219 months, DPR injected mice consistently displayed an aberrant stand, swing speed and stride
220 length, when analysed using a catwalk gait system (Fig. 4 g, $p<0.01$). At 12 months, we
221 assessed the hind limb splay of these mice using a manual scoring system. DPR injected mice
222 also displayed an increase in hind limb splay (Supplementary Figure 5, $p<0.01$). Taken
223 together, these data demonstrate the expression of *C9orf72*-related RREs or DPRs cause ATM
224 defects and DSBs in the murine nervous system, and that these observations are linked to a
225 neurodegenerative phenotype in mammals, which is also consistent with the neurological
226 deficits triggered by 50 poly-GA DPRs reported recently²³.

227

228 After showing that poly-GA expression causes ATM dysfunction and neurodegeneration in
229 mice, we next set out to get more insight into the molecular mechanism of the ATM defect. In
230 response to DNA damage, ATM signalling is mediated by the MRN complex, which forms foci
231 at sites of DSBs, facilitating interactions between ATM and Nbs1²⁴, and thereby enhancing
232 ATM phosphorylation. Thus, we speculated that defective ATM signalling might be a
233 consequence of upstream Nbs1 disruption. In contrast to 53BP1 and pATM, Nbs1 foci formed
234 normally in DPR expressing cells (Fig. 5a). Similar to γ H2AX, expression of 69 DPRs led to an
235 increase in Nbs1 foci formation (Fig 5a, $p=0.001$). These data argue against the failure of Nbs1
236 recruitment as a cause for the *C9orf72*-related ATM defect.

237

238 Central to efficient ATM mediated signalling is a cascade of post-translational histone
239 modifications required for efficient and sustained DSB repair. An important event is the
240 ubiquitylation of histone H2A by the E3 ubiquitin ligase RNF168, which plays an important
241 role for the recruitment of 53BP1 to DSBs²⁵⁻²⁹. Importantly, RNF168 mediated H2A
242 ubiquitylation, and subsequent 53BP1 recruitment, has also been shown to maintain efficient
243 ATM signalling^{30,31}. We thus examined if *C9orf72* expansions would impact H2A
244 ubiquitylation, which may explain the observed ATM signalling defect. Consistent with
245 previous reports²⁷⁻²⁹, Western blot analyses using H2A specific antibodies revealed multiple
246 ubiquitylated species (Fig 5b). The extent of H2A ubiquitylation, however, was attenuated in
247 DPR expressing cells, which showed \sim 2-fold less H2A ubiquitylation in comparison to
248 controls (Fig 5b, $p= 0.0042$). Consistent with the Western blotting data, immunostaining with
249 specific ubiquitylated-H2A (ub-H2A) antibodies also revealed reduced levels of ub-H2A foci
250 in cells expressing 69 DPRs, when compared to control cells (Fig. 5c, $p=0.013$). Interestingly,
251 although DPR-positive cells exhibited fewer ub-H2A foci, we noticed co-localisation between

252 DPRs and ub-H2A (Fig 5c, arrow). Since RNF168 is the key ubiquitin ligase driving H2A
253 ubiquitylation³², we wondered if the defective ATM signalling is due to decreased availability
254 of RNF168. Consistent with this idea, increasing the pool of RNF168 by overexpression of
255 GFP-RNF168 in DPR expressing cells restored 53BP1 (Fig 5d, $p=0.0003$) and increased
256 pATM (Fig 5e, $p=0.006$) foci formation. Moreover, these experiments revealed that RNF168
257 was also sequestered into DPRs (Fig 5d, arrow), explaining the unexpected co-localisation
258 between DPRs and ub-H2A.

259

260 Inspired by a recent seminal report showing that the ALS-associated autophagy protein,
261 p62/SQSTM1, perturbs RNF168 function and impairs H2A ubiquitylation-mediated DNA
262 repair³³, we wondered if the defective ATM signalling in *C9orf72* models may result from a
263 p62-mediated attenuation of H2A ubiquitylation. This hypothesis is particularly attractive since
264 p62 accumulation is a hallmark pathology of *C9orf72*-related disease^{34,35}. If this is true, we
265 predicted that depletion of p62 in DPR expressing cells would recapitulate the effect of
266 RNF168 overexpression. Indeed, p62 depletion with siRNA (Fig 5f) led to restoration of
267 53BP1 (Fig 5g, $p=0.047$) and pATM (Fig 5h, $p=0.003$) foci. These data are fully in-line with
268 reports that p62 accumulation impairs RNF168-mediated H2A ubiquitylation³³ and with studies
269 that highlight the role of H2A ubiquitylation in ATM signalling^{30,31}. Furthermore, the
270 restoration of ATM-mediated repair in DPR-positive cells by p62 depletion suppressed the
271 number of DSBs (Fig. 5i, $p=0.029$) and the aberrant accumulation of R-loops (Fig. 5j,
272 $p=0.037$). The latter is consistent with the recently reported reciprocal functional interaction
273 between R-loops and ATM signalling^{36,37}. The above data suggest that p62 accumulation and
274 the consequential defect in ATM signalling act together with the expansion-driven R-loops to
275 trigger genome instability, though it is not clear whether they constitute distinct or epistatic
276 pathways. To address this question, we examined whether SETX overexpression would further

277 suppress the elevated level of DSBs in p62-depleted cells. Consistent with our previous data
278 (Figure 1j and Figure 5i), overexpression of SETX or depletion of p62 was capable of reducing
279 γ H2AX foci in DPR-positive cells (Fig. 5k; $p=0.0013$ and 0.0146 , respectively). However, the
280 concomitant overexpression of SETX and depletion of p62 further reduced DSB levels
281 compared to levels observed by SETX overexpression or p62 depletion alone (Fig. 5k;
282 $p=0.0433$ and $p=0.0045$, respectively) Together, these data identify two separate arms that
283 drive genome instability in DPR expressing cells. One arm is driven by p62 accumulation,
284 defective H2A ubiquitylation and the subsequent ATM signalling defect; and the other is
285 driven by repeat-associated R-loop accumulation. Whilst the two arms are distinct, cross-talk
286 does exist due to the reciprocal functional interaction between R-loops and ATM signalling^{36,37}.
287

288 In addition to DNA damage, ATM can also become activated in the absence of DNA damage
289 by the drug chloroquine, a DNA intercalating agent³⁸. Since the accumulation of p62 and
290 inhibition of H2A ubiquitylation led to dysfunctional ATM signalling in *C9orf72* cells, we
291 reasoned that ATM activation by chloroquine, which is not linked to sensing DNA damage,
292 would not be defective. Thus, we tested whether chloroquine-induced chromatin relaxation
293 could restore ATM signalling in DPR-positive cells. As predicted, whilst DPR-positive cells
294 did not display any ATM phosphorylation under normal conditions, chloroquine treatment led
295 to pan-nuclear ATM phosphorylation (Supplementary Figure 6a, $p=0.004$), indicating that DPR
296 expressing cells are indeed responsive to chloroquine and, consequently, are able to activate
297 ATM. To test whether the ensuing ATM signalling was restored by chloroquine treatment in
298 *bona fide* primary neuronal cells, we next examined the localisation of HDAC4, which we
299 previously demonstrated is localised to the nucleus of neurons expressing DPRs. As predicted,
300 whilst DPR-positive neurons displayed nuclear HDAC4, the addition of chloroquine led to re-
301 localisation of HDAC4 to the cytoplasm (Supplementary Figure 6b). Whilst only ~40% of

302 DPR-positive neurons displayed cytoplasmic HDAC4 without chloroquine treatment, the
303 number of neurons with cytoplasmic HDAC4 increased to almost 100% after cell treatment
304 with chloroquine (Supplementary Figure 6b; $p=0.014$). These data demonstrate that ATM can
305 be activated by inducing chromatin relaxation, and are consistent with our previous data linking
306 dysfunctional ATM signalling to a defect in histone ubiquitylation.

307

308 Given that chromatin compaction has been reported in *C9orf72* samples³⁹, we speculated that
309 an increase in heterochromatin formation might be a hallmark of *C9orf72*-DPR expressing
310 cells. Using H3K9me3 as a marker for heterochromatin, we confirmed that DPR expression
311 increases heterochromatin formation, as measured by ~50% increase in H3K9me3 signal by
312 western blotting and immunocytochemistry, in comparison to empty vector control cells
313 (Supplementary Figure 7a,b, $p=0.002$). Since ATM has been linked to the repair of
314 heterochromatic DSBs⁴⁰, we wondered if promoting chromatin relaxation by Trichostatin A
315 (TSA)^{41,42} would reduce DSBs levels in DPR-positive cells. Indeed, treatment of DPR-positive
316 cells with TSA led to a reduction of H3K9me3 nuclear fluorescence (Supplementary Figure 7c,
317 $p=0.002$). Unlike chromatin relaxation by chloroquine, TSA treatment did appear to activate
318 ATM signalling (Supplementary Figure 7c). TSA treatment did, however, lead to a reduction in
319 γ H2AX – but not R-Loop - foci in DPR-expressing cells (Supplementary Figure 7d,e,
320 $p=0.001$), suggesting that chromatin relaxation reduces DSB levels by decreasing the
321 requirement for ATM signalling, rather than activating ATM directly. This is in-line with the
322 reported role of ATM during the repair of heterochromatic DSBs, in which the DNA repair
323 deficit caused by ATM inhibition was similarly overcome by chromatin relaxation⁴⁰.
324 Importantly, TSA treatment increased DSB levels in control cells, suggesting that whilst TSA
325 is beneficial in *C9orf72* models in which the chromatin is compact, it may confer sensitivity to
326 other types of DNA lesions in cells with non-pathological chromatin arrangements⁴¹. Finally,

327 we wondered whether TSA treatment would also have beneficial effects on cell survival in
328 *C9orf72* models, as a result of reduced DSBs. TSA treatment was able to rescue the cell
329 toxicity triggered by 69 DPR expression to a similar level that was observed in DMSO treated
330 control cells, as measured by both PARP1 cleavage (Supplementary Figure 7f, $p=0.0004$) and
331 trypan blue exclusion assays (Supplementary Figure 7g, $p=0.003$). These data suggest that
332 *C9orf72*-related DPRs drive the formation of heterochromatin, thereby exacerbating the ATM
333 defect and ultimately leading to cell death.

334

335 Next, we set out to understand why *C9orf72* cells display increased heterochromatin formation.

336 Whilst R-loops are generally associated with euchromatin, their progressive accumulation may
337 also drive heterochromatin formation^{43,44}. As such, we hypothesised that increased
338 heterochromatin might be, at least in part, caused by the expansion-driven R-loops. SETX
339 overexpression was able to reduce heterochromatin levels in DPR-positive cells
340 (Supplementary Figure 7h, $p=0.031$). Additionally, SETX overexpression did not further
341 reduce DSBs in TSA treated *C9orf72* cells (Supplementary Figure 7i). Since ATM signalling is
342 particularly important for the relaxation and subsequent repair of heterochromatic DNA³⁵, we
343 speculated that increased heterochromatin might also be a consequence of *C9orf72*-linked
344 ATM dysfunction. We observed a reduction in H3K9me3 signal after restoring ATM signalling
345 with p62 siRNA (Supplementary Figure 7j, $p=0.025$). Similarly, the reduction of DSBs via p62
346 depletion appeared to be epistatic with TSA treatment (Supplementary Figure 7k). Taken
347 together, these data demonstrate that chromatin relaxation via TSA attenuates heterochromatin
348 levels, genomic instability, and cellular toxicity in *C9orf72* models. We propose a model in
349 which *C9orf72* repeat expansion drive the formation of heterochromatin, placing greater
350 demand on ATM-mediated repair - a system that is already defective - and thereby creating a
351 vicious cycle that leads to neuronal cell death. Although this is an attractive model for *C9orf72*

352 pathologies, we note that TSA can enhance DNA damage sensitivity in *C9orf72* unrelated
353 models^{41,45}, suggesting that the extent of heterochromatin and genome-wide distribution of
354 breaks greatly influence the response of mammalian cells to TSA.

355

356 Finally, we examined if R-loops, DNA breaks and defective ATM signalling observed in
357 human cells, mouse neurons, and in mice were also present in *C9orf72*-ALS patient tissues. To
358 test this, we subjected post-mortem spinal cord sections from controls and *C9orf72*-ALS
359 patients to R-loop, γ H2AX, and HDAC4 immunohistochemistry analyses. Consistent with our
360 cellular data, we observed a significant increase in the number of R-loop (Figure 6a,b;
361 $p=0.0038$) and γ H2AX-positive motor neurons (Figure 6c,d; $p=0.0268$) in *C9orf72*-ALS
362 sections compared to controls. The specificity of S9.6 signal was confirmed by prior treatment
363 with RNase H1 (Supplementary Figure 8). In addition, we also observed an increase in the
364 percentage of motor neurons with nuclear HDAC4 staining in *C9orf72*-ALS sections compared
365 to controls (Figure 6e,f, $p=0.0463$), suggesting that ATM signalling is also dysregulated in
366 *C9orf72*-ALS patient motor neurons. We note that nuclear accumulation of HDAC4 could also
367 result from other factors, such as increased activity of the inhibitor of PP2A (I2PP2A), which
368 has been observed in sporadic-ALS patients⁴⁶. We also observed an increase in H3K9me3
369 nuclear fluorescence in *C9orf72*-ALS motor neurons (Supplementary Figure 9, $p=0.022$).
370 Together, our findings are consistent with a model whereby the expression of *C9orf72*-related
371 products drive R-loop-mediated DNA breakage that is further exacerbated by the suppression
372 of ATM-mediated DSB repair, resulting in disruption of co-transcriptional processing and
373 neurodegeneration (Supplementary Figure 10).

374

375 In summary, we identified an increase in R-Loops and a defect in ATM signalling using 4
376 model systems: human cells, rodent neuronal cultures, RNA and dipeptide mouse models of

377 *C9orf72*-related expansions, and in post-mortem ALS patient tissues. Consistent with an ATM
378 defect, we further report increased accumulation of TOP1 mediated protein-linked DNA breaks
379 (PDBs), an established marker for neurodegeneration^{20,21,47,48}, in cells expressing *C9orf72*-
380 related products. Furthermore, CNS tissues obtained from mice expressing *C9orf72*-related
381 DPRs and from *C9orf72*-ALS patients exhibit nuclear HDAC4 retention, which is consistent
382 with defective ATM signalling²². We further uncover an important link between pathological
383 accumulations of p62 - a hallmark of *C9orf72*-ALS - and defective H2A ubiquitylation,
384 dysfunctional ATM-mediated DNA repair, and increased genomic breaks. Our data reinforce
385 the pathophysiological significance of the recently reported mechanism by which p62 impedes
386 DNA damage repair³³. A length-dependent increase in R-loops was evident in cells expressing
387 RREs, likely driven by G-rich RNA, which is prone to template strand invasion. Genome-wide
388 analyses in yeast suggest that R-loops can regulate sense and antisense gene expression⁴⁹,
389 possibly explaining the aberrant expression of both transcript types in *C9orf72* patient
390 samples⁵⁰. Notably, suppressing R-loop levels by SETX overexpression was able to reduce
391 DSBs and cellular toxicity in both RRE and DPR models of *C9orf72*-related disease.

392

393 To conclude, we uncover two distinct but partially overlapping pathways by which *C9orf72*
394 repeat expansions lead to genomic instability. One arm is driven by p62 accumulation,
395 defective H2A ubiquitylation and defective ATM signalling, as measured by impaired 53BP1
396 foci formation and the phosphorylation of downstream target p53, the accumulation of TOP1
397 mediated PDBs, and nuclear accumulation of HDAC4. The second arm is driven by the
398 expansion led increase in R-loop formation. Increased R-Loops and defective ATM signalling
399 can account for multiple yet unexplained phenotypes of *C9orf72* repeat expansions: **(1)**
400 increased heterochromatin, **(2)** increased DSBs, and **(3)** the previously reported *C9orf72*-
401 associated splicing defects³⁷. Finally, this work suggests that targeted modulation of R-Loop

402 homeostasis by R-Loop specific helicases, or DSB repair kinetics by chromatin modulating
403 drugs, may offer new therapeutic opportunities for *C9orf72*-related neuropathologies.

404

405 **Acknowledgements**

406 M.A. and S.F.E-K made equal financial contribution to the work through the European
407 Research Council Award (294745) and the Wellcome Trust Investigator Award (103844),
408 respectively. S.F.E-K is funded by a Wellcome Trust Investigator Award (103844), a Lister
409 Institute of Preventative Medicine Fellowship, and a European Union British Council award.
410 M.A. is funded by the European Research Council (ERC Advanced Award no. 294745) and
411 MRC DPFS Award (129016/G1001492). C.L., S-C.C., M.J., S.R. and C.W. (50%), were
412 supported by the Wellcome Trust Investigator Award (103844). E.K. was supported by Eve
413 Davis Studentship. S.H.M., I.T, V.L., K.L., J.C. and C.W. (50%) were supported by the ERC
414 Advanced Award (294745). P.J.S and G.M.H are supported by the Medical Research Council
415 (MR/M010864/1) and P.J.S is supported as an NIHR Senior Investigator. K.D.V is supported
416 by the Thierry Latran Foundation (Project RoCIP) and Medical Research Council
417 (MR/K005146/1 and MR/M013251/1). We thank Adrian Isaacs for providing the original DPR
418 plasmids, Grant Stewart for providing the RNF168 plasmid and Lynne Baxter for the
419 immunohistochemistry of post-mortem sections.

420 **Author contributions**

421 C.W conducted the immunocytochemistry, immunohistochemistry and immunoblotting
422 experiments. S.H.M generated the mouse model and conducted behavioural experiments with
423 assistance from K.L., V.L., T.I., J.S.C, and I.C. E.K and S.H.M. generated neuronal cell
424 cultures. E.K, S-C.C, M.J, S.R., M.H. assisted with cell culture, imaging and DNA repair
425 assays. K.L., V.L., T.I., I.C assisted with mouse experiments. C.L purified and conducted
426 TOP1cc assays. K.D.V, I.T, J.C, A.H generated expansion constructs. I.T. and S.H.M.

427 generated and validated viral vector stocks. W.E. optimised and conducted the neutral comet
428 assays. P.J.S. and A.H. provided RRE constructs and I.T., J.S.C. and P.J.M. sub-cloned them
429 into viral vectors. P.J.S. and A.H. provided the C9orf72-ALS biosamples and expertise. S.F.E-
430 K, M.A, G.H, C.W, S.H.M analysed the data. S.F.E-K and C.W. wrote the manuscript with
431 help from M.A. and S.H.M. All authors contributed to the final manuscript. S.F.E-K and M.A.
432 conceived and co-led the project.

433

434 **Methods**

435 **Mammalian Cell Culture**

436 MRC5 cells were grown in Minimum Essential Media (MEM) (Sigma Aldrich) supplemented
437 with 10% Foetal Bovine Serum (FBS) (Sigma Aldrich), 2mM L-Glutamine (Sigma Aldrich)
438 and 1% Penicillin/Streptomycin (Sigma Aldrich). HEK 293T cells were grown in Dulbecco's
439 Modified Essential Media (Sigma Aldrich) supplemented with 10% Foetal Bovine Serum
440 (FBS), and 1% Penicillin/Streptomycin (Sigma Aldrich). MRC5 cells were seeded on
441 coverslips into the wells of a 24-well plate at a density of 3×10^5 cells/cm². For other
442 experiments, HEK 293T cells were seeded into the wells of a 12-well plate at a density of $3 \times$
443 10^5 cells/cm². The following day, cells were transfected with 250ng of DNA/cm² of each DNA
444 plasmid, using polyethylenimine at a molar concentration of 3:1 (PEI: DNA). 6-hours post-
445 transfection, the media was replaced and cells were incubated for a further 2 or 3 days, as
446 indicated in the figure legends. For p62 knockdown experiments, MRC5 cells were co-
447 transfected with cDNA plasmids using PEI (as described above) alongside control siRNA
448 particles or p62 targeting particles (Santa Cruz, sc-29679) at a molarity of 25nm. Dharmafect
449 was used as the transfection reagent for all siRNA experiments at a ratio of 1:1
450 (siRNA:Dharmafect). Recombinant adenovirus (Adenovirus-type 5 dE1/E3) encoding for
451 SETX or RFP were purchased from Vector Biolabs. Viral stocks (10^9) were diluted 1:5000 for
452 MRC5 and 1:20000 for HEK293T cells to give an approximate multiplicity of interest (MOI)
453 of 10 and 2.5 for MRC5 and HEK 293T cells, respectively. Virus containing media was added
454 for 2-3 hours, prior to transfection, and replaced with fresh. For Western blotting, cell death
455 assays, and COMET assay experiments, MRC5 cells and HEK 293T cells were grown for 72
456 hours post-transfection. For DNA repair immunocytochemistry assays, cells were grown for 48
457 hours post-transfection.

458

459 **Generation of Repeat Expansion Constructs**

460 Synthesised TCGAC(G4C2)₁₀ sense and ACGT(G2C4)₁₀ antisense ssDNA oligonucleotides
461 (Sigma-Aldrich) were designed with Sall/XhoI overhangs. The dsDNA oligos were generated
462 by denaturing oligos were denatured at 99°C for 30 min and then annealing by stepwise cooling
463 of 0.5°C/min. These (G4C2)₁₀ were ligated into Sall and XhoI digested pcDNA6.2-
464 GW/EmGFP-miR (Invitrogen), to generate pcDNA6.2-GW/EmGFP-(G4C2)₁₀. Further
465 (G4C2)₁₀₂ repeats were subcloned using the 3' XhoI site. pCMV-EmGFP-(G4C2)_n vectors
466 containing 10 and 102 repeats were generated via this method. EmGFP was subsequently
467 excised using the flanking DraI restriction site. The (G4C2)₁₀ and (G4C2)₁₀₂ constructs were
468 sub-cloned into pcDNA5/FRT/TO HIS (Addgene) using DraI and XhoI restriction sites.
469 Transformations of plasmids containing the (G4C2)_n repeat constructs were performed using
470 recombination-deficient β-10 E.coli (NEB) to minimise (G4C2)_n repeat shrinkage. To model
471 gain-of-function via RAN translation of C9orf72 repeat expansions we produced two
472 expression constructs coding for uninterrupted V5-tagged poly-GA DPRs, using an expandable
473 cloning strategy with AgeI and MreI as compatible enzymes⁵¹. We first constructed a 'start
474 acceptor' pCi-Neo vector (Promega) by cloning a V5-3xGA insert into the XhoI/NotI sites (ctc
475 gag gcc acc atg ggc aaa ccg att ccg aac ccg ctg ctg ggc ctg gat agc acc ggt gca ggt gct ggc gcc
476 ggc gga tcc gaa ttc tag ccg cgg ccg c) and a 'start donor' vector with a 14xGA insert (ctc gag
477 acc ggt gca ggt gct gga gct ggt gca ggt gct gga gca ggt gca ggt gct gga gct ggt gca ggt gct gga
478 gca ggt gct ggc gcc ggc gga tcc gaa ttc ccg cgg ccg c) in the XhoI/NotI sites and used these to
479 propagate the GA repeats as shown below to construct 34, 69 GA repeats. A V5 construct that
480 lacked DPRs was created by AgeI and NGOMIV digestion, which excised the DPR coding
481 region.

482 **Primary Cortical Neuron Cultures**

483 The cortex from the brains of E17.5 rat pups were harvested and stored in Hank's Balanced
484 Salt Solution (without calcium, without magnesium) (HBSS -/-) (Sigma Aldrich). The tissue
485 was washed with HBSS (-/-) and then incubated with 0.0035% Trypsin (Sigma Aldrich) for 15
486 minutes. DNase (10µg/mL)(Sigma Aldrich) was then added at a ratio of 1:1 (v/v), and the
487 tissue was re-suspended in 1mL triturating solution (1% Albumax, 25mg Trypsin Inhibitor,
488 10µg/mL DNase) (Sigma Aldrich). Neurobasal media (ThermoFisher), supplemented with 2
489 mM L-Glutamine (Sigma Aldrich), 1% Penicillin/Streptomycin (Sigma Aldrich) and 1 x B-27
490 (Sigma Aldrich) was then added at a ratio of 1:5 (Triturating solution:Neurobasal media). Cells
491 were then plated onto Poly-D-Lysiene coated coverslips, in the wells of a 24 well-plate at a
492 density of 9.365×10^4 cells/cm². 1.5×10^5 viral genomes (vg) per cell of AAV9 was added to
493 cultures after 5 days *in vitro* (DIV). Half of the culture media was replaced with fresh media
494 every 3 days. 7-days post-transduction (13 DIV), cells were treated with CPT (10µM) where
495 indicated, and fixed using 4% paraformaldehyde or methanol:acetone (50:50).

496 **Production of viral vectors**

497 Sixty 15 cm plates containing HEK 293T cells at a 80% confluence were transfected using
498 polyethylenimine (MW ~ 25K) with a mixture of three plasmids (at a molar ratio of 2:1:1 in
499 order as listed) required to generate an infectious AAV9 viral particle: (1) a plasmid providing
500 helper genes isolated from adenovirus that enhance viral infectivity (pHelper); (2) an ITR-
501 containing plasmid plasmid containing the 10, 102 RRE or 34, 69 DPR driven by the CMV
502 promoter; (3) a plasmid that carries the AAV Rep-Cap proteins (pAAV2/9); (3) A total of 52µg
503 of DNA was transfected per plate with pHelper:pAV2-CMV-GFP:pAAV2/9. For all
504 experiments, we used the pAV2-CMV-GFP consisting of two ITRs in a truncated genome that
505 resulted in a self-complementary AAV9 (scAAV9). Four days after transfection, the AAV
506 enriched media was collected, incubated at 37°C for 2 hours with 3,750 units of benzonase-

507 nuclease (Sigma, USA), filtered through a 0.22 µm filter, and concentrated to a volume of 1ml
508 using Amicon spin filter units (Millipore, USA). The virus was then washed with 50 ml of
509 phosphate buffered saline (PBS, pH 7.3) in the same Amicon spin filter units, and concentrated
510 to a final volume of 0.5 ml. The viral sample volume was expanded to 14ml with PBS and
511 separated through a discontinuous iodixanol (D1556, Sigma, USA) gradient (4ml of 54%, 9 ml
512 of 40%, 9 ml of 25%, 5 ml of 15%), and centrifuged at 69,000 rpm for 1.5 hours at 18°C. The
513 purified virus, which was found as a white layer between the 54% and 40% iodixanol gradient
514 was subsequently removed in 0.5ml fractions using a syringe, and 10µl of each fraction was
515 mixed at an equal ratio with a 2X reducing sample SDS-PAGE buffer, heated to 75°C for 20
516 minutes, separated on a 4-20% precast TGX mini-gel (Biorad, USA), and stained with Sypro-
517 Ruby according to the manufacturer's protocol (Life Technologies, USA). Fractions that showed
518 a pure virus composed solely of the VP1, VP2 and VP3 bands were combined, and washed
519 against 5 full volumes (15ml each) of PBS with an Amicon spin filter, before collecting in a
520 final volume of between 300-500µl. Concentrated viral stocks were stored at -80°C until usage.

521

522 Viral titers were determined with the Quantifast SyBR Green PCR Kit (Qiagen, Cat 204054) on
523 a BioRad CFX96 thermal cycler, following the manufacturer's instructions. The number of
524 GFP copies in three dilutions of a purified AAV9 virus (100x, 1000x, 10,000x) were compared
525 to a standard curve generated by serial dilutions of a linearized pAV2-CMV-GFP vector.
526 Primers used to quantify viral genomes were (Poly A, Forward: 5'-ATT TTA TGT TTC AGG
527 TTC AGG GGG AGG TG-3'), (PolyA, Reverse: 5'-GCG CAG AGA GGG AGT GGA CTA
528 GT-3'), (GFP, Forward: 5'- GAC GGC AAC ATC CTG GGG CAC AAG-3'), and (GFP,
529 Reverse: 5': CGG CGG CGG TCA CGA ACT C-3').

530

531 **RNA Fluorescent-In-Situ-Hybridisation (FISH):**

532 FISH was performed following a modification of the method described previously⁵². MRC5
533 cells or rat cortical neurons were fixed with 4% PFA for 10 minutes at room temperature. For
534 S9.6 staining however, fixation was performed using ice-cold methanol:acetone (50:50) for 10
535 minutes at -20°C. Cells were then incubated with pre-hybridisation buffer (50% formamide, 2X
536 saline sodium citrate (SSC), 100 mg/ml dextran sulphate, 50 mM sodium phosphate pH 7.0) for
537 1 hour at 66°C. Subsequently, cells were incubated with hybridisation buffer containing a 5'
538 TYE-563-labelled locked nucleic acid (16-mer fluorescent)-incorporated DNA probe against
539 the GGGGCC RNA hexanucleotide repeat (Exiqon, Inc., batch number 607323), at a
540 concentration of 400 ng/ml for 1 hour or overnight at 66°C. A 1 hour incubation period was
541 preferred for phospho-ATM staining, due to the loss of antigen signal after overnight treatment
542 in hybridisation buffer. After hybridization, slides were washed once in 2 X SSC with 0.1%
543 Tween-20 at room temperature and three times in 0.1 X SSC at 66°C. All washes were
544 performed for 10 minutes. All solutions were made with DEPC-treated water. Following the
545 completion of this FISH protocol, ICC was then performed as described below, though using
546 DEPC-treated PBS solutions.

547

548 **Immunocytochemistry (ICC):**

549 MRC5 cells or rat cortical neurons were fixed with 4% PFA for 10 minutes at room
550 temperature, or with ice-cold methanol:acetone (50:50) for 10 minutes at -20°C. Cells were
551 then washed 3 times with Phosphate Buffered Saline (PBS), incubated with 0.5% Triton-X (in
552 PBS) for 5 minutes, and washed a further 3 times with PBS. In order to confirm the specificity
553 of the S9.6 antibody, MRC5 cells, fixed with methanol:acetone, were incubated with RNASE
554 H enzyme (100units/mL) in 3% BSA in PBS overnight at 4°C before ICC/FISH-ICC. For ub-
555 H2A staining, cells were incubated with PBS containing 0.5% Triton-X for 2 minutes at room

556 temperature before fixation with 4% PFA. Subsequently, cells were incubated with 3% BSA
557 for 30 minutes, before being incubated with primary antibodies (in 3% BSA) for 1 hour (with
558 the exception of Ub-H2A, which was incubated overnight at 4°C). For DPR experiments, a V5
559 antibody was always used to detect DPR-positive cells. Cells were washed 3x with PBS and
560 incubated with fluorescent secondary antibodies (in 3% BSA) for 1 hour, before being washed
561 another 3x with PBS. Coverslips were mounted onto glass slides using Fluoromount™
562 Aqueous mounting medium (Sigma Aldrich).

563

564 Anti-V5 antibodies: mouse (Abcam, ab27671), and rabbit (Bethyl, A190-120A) were used at
565 1:1000. A mouse anti-RNA:DNA hybrid (S9.6) antibody (Kerafast, ENH001) and a mouse
566 anti-ub-H2a antibody (Merck Millipore, E6C5) were used used at a concentration of 1:500 for
567 immunocytochemistry, and at 1:5000 for FISH-ICC double staining. A mouse anti- γ H2AX
568 (Ser 139) antibody (Merck Millipore, JBW301), a rabbit anti-53BP1 antibody (Bethyl, A300-
569 272A), a rabbit anti-phospho-ATM antibody (Abcam, EP1890Y), a mouse anti-phospho-P53
570 antibody (Cell Signalling, 9286s), a mouse anti-cleaved PARP (Cell Signalling, 9548), and a
571 rabbit anti-Nbs1 antibody (Sigma, N3162) were all used at a concentration of 1:1000. A rabbit
572 anti-HDAC4 antibody (Abcam, ab1437) was used at a concentration of 1:250. A rabbit
573 H3K9me3 antibody (Abcam, ab8898) was used at a concentration of 1:2000. Subsequently,
574 cells were washed 3 times with PBS and incubated with the corresponding Alexa fluor
575 secondary antibodies (all purchased from Life Technologies and used at a concentration of
576 1:500) as well as DAPI for 1 hour. For FISH-IF double staining, an Alexa fluor 488 was used
577 in conjunction with the Cy3 fluorescent probe. Cells were washed a further 3 times with PBS
578 and coverslips were mounted onto glass slides using Fluoromount™ Aqueous mounting
579 medium (Sigma Aldrich).

580

581 **Immunohistochemistry**

582 Mouse brain and spinal cord sections were incubated with 0.5% Triton-X for 30 minutes,
583 followed by a 1 hour incubation with 3% BSA (with 0.2% Triton-X) for 1 hour. Subsequently,
584 sections were incubated with primary antibodies (in 3% BSA with 0.2% Triton-X) overnight at
585 4°C. The following day, sections were washed 3 times with PBS, and were incubated with
586 fluorescent secondary antibodies alongside DAPI for 1 hour at room temperature. For γ H2AX
587 and HDAC4 staining in DPR mice sections, a biotinylated secondary antibody was used in
588 conjunction with a tertiary anti-streptavidin Alexa Fluor 488 antibody, in order to enhance the
589 signal. Sections were washed a further 3 times, before mounting with Fluoromount™ Aqueous
590 mounting medium (Sigma Aldrich). For human spinal cord staining, 5 μ m paraffin embedded
591 spinal cord sections from *C9orf72*-ALS and non-ALS controls (Supplementary Table-1,) were
592 incubated with primary antibodies specific for HDAC4 (Abcam, ab1437), γ H2AX (R&D
593 systems, AF2288), S9.6 (Kerafast, ENH001), or H3K9me3 (Abcam, ab8898) at concentrations
594 of 1:250, 1:500, 1:1000 and 1:1000; respectively. The NeuN D3S3I antibody was from cell
595 signalling (Cat no 12943) and was used at a concentration of 1 in 500. Before primary antibody
596 incubation, antigen retrieval was performed in 10 mM Sodium citrate (pH6) or 10 mM Tris
597 Base for HDAC4, S9.6 and H3K9me3, or 1 mM EDTA (pH9) for γ H2AX. Antigen retrieval
598 was performed for 30 minutes in a pressure cooker. Immunohistochemistry was performed
599 using the IntelliPATH FLX™ Detection Kit, according to the manufacturer's protocol. Work
600 on human tissue was reviewed by the Sheffield Brain Tissue Bank (SBTB) Management Board
601 and approved to release tissue under REC 08/MRE00/103 was granted. All post-mortem tissue
602 stored at Sheffield Brain Tissue Bank was obtained with informed consent.

603

604 **Cell Lysis, SDS-PAGE and Western Blotting**

605 In order to collect whole-cell lysates, MRC5 cells, HEK 293T cells, or mice brain stem sections
606 were lysed in RIPA buffer (150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM
607 Tris, pH 8.0, supplemented with protease inhibitor cocktail, Sigma Aldrich) on ice for 30
608 minutes, before being sonicated in order to shear the DNA. For the isolation of chromatin
609 fractions, the cytoplasmic and nuclear soluble proteins were first removed by hypotonic and
610 hypertonic buffers. The remaining pellet was lysed with *nuclear insoluble buffer* (20mM Tris
611 pH8, 150mM NaCl, 1% SDS, and 1% NP-40) for 30 minutes on ice before being sonicated.
612 The protein concentration of each whole-cell lysate was estimated using a BCA assay
613 (Pierce™), and equal quantities of protein were mixed with a 2X reducing sample SDS-PAGE
614 buffer, heated to 95°C for 5 minutes, separated on a 4-20% precast TGX mini-gel (Biorad,
615 USA), and transferred onto a PVDF membrane (Millipore, USA). Membranes were blocked
616 with 3% BSA in TBS with 0.05% Tween (TBST) for 30 minutes, before incubating at either
617 room temperature for 2 hours or 4°C overnight with agitation with primary antibody in 3%
618 BSA in TBST. For Western Blotting, a mouse a rabbit ATM antibody (Abcam, ab82512) was
619 used at 1:2000, a rabbit H3K9me3 antibody (Abcam, ab8898) was used at 1:5000, mouse anti- α
620 Tubulin (Abcam, T9026) and mouse anti-GAPDH (Calbiochem, CB1001) were used at 1:5000.
621 Cleaved-PARP antibodies (Cell Signalling, 9548) were all used at 1:1000. Rabbit anti-H2A
622 antibodies (Abcam, ab18255) were used at 1:1000. Membranes were then washed 3 times for 5
623 minutes with TBST and incubated with either a HRP-linked secondary anti-mouse antibody
624 (Bio-Rad, 1721011) or an anti-rabbit antibody (Dako, D048701-2). Enhanced
625 ChemiLuminescence (ECL) substrate was then added to the membrane to enable detection, and
626 non-saturated images were acquired using a G:BOX EF machine (Syngene) and Snapgene
627 software (Syngene). **Supplementary Figures 11-14 contain raw files of the Western blots used**
628 **in this study.**

629

630 **Trypan Blue Cell Death Assay**

631 3 days after transfection, HEK 293T cells were washed once with PBS and treated with trypsin
632 until detached. Subsequently, 10% FBS-containing DMEM was added and cells were
633 resuspended gently by pipetting up and down 3 times. Cell suspensions were then mixed with
634 Trypan Blue (0.04%) at a ratio of 1:1. Immediately after, the percentage of cells that were
635 permeable to Trypan Blue was calculated using a haemocytometer and a brightfield
636 microscope. Scoring was performed under single-blinded conditions.

637

638 **Measurement of topoisomerase I cleavage complexes (TOP1cc)**

639 TOP1 protein–DNA complexes were purified using caesium chloride density gradients.
640 Approximately 2×10^6 cells were lysed in 1% sarcosyl, 8 M guanidine HCl, 30 mM Tris pH 7.5
641 and 10mM EDTA. Cell lysates were then incubated at 70°C for 15 minute to remove all non-
642 covalently bound proteins from DNA. Cell lysates were then loaded on a caesium chloride
643 density (CsCl) step gradient (5 ml total volume) and centrifuged at $75,600 \times g$ at 25°C for 24
644 hour to separate free proteins from DNA. Ten consecutive 0.5 ml fractions were collected and
645 slot blotted onto Hybond-C membrane (Amersham). To ensure equal DNA loading, the DNA
646 concentration in each extract was determined fluorimetrically using PicoGreen (Molecular
647 Probes/Invitrogen). Covalent TOP1–DNA complexes were then detected by immunoblotting
648 with anti-TOP1 polyclonal anti- bodies (sc-32736, Santa Cruz.) and visualised by
649 chemiluminescence.

650

651 **Neutral single-cell agarose gel electrophoresis (Comet) assays**

652 HEK293 cells at density (60,000 cell/24well plates) were seeded at 37°C overnight. In the
653 second day cells were transfected using polyethyleneimine (PEI) $1 \mu\text{g}/\mu\text{l}$ transfection reagent

654 with plasmid DNA 500 ng/well. Transfected cells were incubated at 37°C for 24h then the
655 complete media replaced with 1% FBS media and incubated for an additional 48h. Transfection
656 efficiency was assessed at ~75%. Cells were suspended in pre-chilled phosphate buffered saline
657 (PBS) and mixed with equal volume of low-gelling-temperature agarose 1.3% (Sigma, Type
658 VII) preserved at 42°C. Cell mixture was immediately spread onto pre-chilled frosted glass
659 slides (Fisher), pre-coated with 0.6% agarose. The slides were incubated at 4°C in the dark
660 until set, and for all further steps. Slides were incubated in pre-chilled lysis buffer (2.5 M NaCl,
661 10 mM Tris-base, 100mM EDTA (pH 8.0), 0.5% Triton X-100, 1% N-laurylsarcosine sodium
662 salt and 3% DMSO; pH9.5) for 2 h. After incubation time slides washed with pre-chilled
663 distilled H₂O (2-10 min), and immersed for 1h in pre-chilled electrophoresis buffer (300mM
664 sodium acetate, 100mM Tris-HCl, 1% DMSO, pH8.3). Then electrophoresis was conducted for
665 60 min at 1 V/cm, accompanied by washing 3 times with neutralization in 400 mM Tris- HCl
666 (pH 7.5) for 15min. Finally, slides was stained with DNA Sybr Green I nucleic (1:10000, in
667 PBS) (Sigma) for 30 min. Tail moments average from 100 cells per sample were counted using
668 Comet Assay IV software (Perceptive Instruments, UK).

669

670 **Generation of AAV9-mediated mice**

671 3×10^{10} vg of purified scAAV was injected into the cisterna magna of C57BL/6J P1 wild-type
672 mice under general anaesthesia. Postnatal day 1 (P1) Pups were placed over a red light torch in
673 prone position to enable visualisation of the injection site. scAAV was loaded into a 5ul syringe
674 and the virus were injected under a flow rate of 1µl/min. Animals were housed in groups of up
675 to 5 per cage. Gender splits for each group was as followed: 0-V5 (6M and 6F), 69-V5 (7M and
676 6F), 10-RRE (2M and 1F) and 102-RRE (2M and 1F). Six or 12 months after injection, mice
677 were sacrificed under terminal anaesthesia and transcardially perfused using a solution of PBS-
678 Heparin. Brain sections were then isolated and fixed using 4% PFA overnight at 4°C. After

679 fixation, tissue was washed with PBS, cryoprotected in 30% of sucrose at 4°C and embedded in
680 OCT (Cell Path®). 20 or 40µm brain coronal sections were derived using a cryostat, and then
681 analysed using immunohistochemistry. Alternatively, brainstem portions were snap frozen in
682 liquid nitrogen and were then lysed in RIPA buffer for Western Blotting analysis. All animal *in*
683 *vivo* experiments were approved by the University of Sheffield Ethical Review Committee and
684 performed according to the Animal (Scientific Procedures) Act 1986, under the Project License
685 40/3739. Animals were administered with AAV9 vectors at postnatal day 1 (P1). To mitigate
686 potential confounding of treatment with litter effects a randomised block allocation design was
687 used to ensure animals from a given litter were stratified across different treatment arms.
688 Where there is only a single viable P1 pup in the litter, this was not allocated to treatment.

689

690 **Whole animal neurological assays**

691 All our behavioural testing was performed the same time of day (10.00-12.00). Gait and
692 locomotion analysis at 6 months was performed by using the Catwalk system version 7.1.
693 Briefly, 6 month old mice (n=12, 13 for 0-V5, 69-V5; respectively) were placed on the Catwalk
694 machine and crossed a glass plate in darkness whilst footprints were captured and recorded
695 using the Catwalk 7.1 software. Each animal performed the Catwalk assay up to six times and
696 the best three runs were selected for analysis. Gait parameters (stand time, swing speed and
697 stride length) were calculated for each limb using the Catwalk 7.1 software. Power analysis
698 using GPower version 3.0.3 was used to determine sample sizes. Based on α of 0.05 and a
699 power of 80% ($\beta=0.8$) a sample size of 12 is required to detect a decrease in Catwalk
700 performance of 20% at 6 months of age. Neuroscoring analysis was performed in mice aged 12
701 months (0-V5, n=8; 69-V5, n=9). Briefly, mice were suspended by the tail and the splay defects
702 were observed and scored individually for right and left hind-limbs, using a scale described
703 previously: 0 normal splay; 1; mild defect; 2: moderate defect; 3: strong splay defect; 4

704 paralysed⁵³. All behavioural tests were performed in blinded conditions, though neuroscoring
705 splay analysis was performed under double-blinded conditions by a single observer.

706

707 **Image Acquisition and Analysis**

708 All representative images presented are Z stack projections acquired on a Leica SP5 confocal
709 microscope, using the 63x 1.20 lens. For imaging cell monolayers, each stack was performed at
710 0.5µm intervals, scanning the entire nucleus. In tissue sections images were acquired using
711 1µm Z stacks. Fluorescence intensity quantification was performed according to a method
712 described previously⁵⁴ with some modifications. Using Image J, the corrected nuclear
713 fluorescence value of the relevant channel was calculated using the formula: *Corrected Nuclear*
714 *Fluorescence (CNF) = Integrated Nuclear Density (IND) – Nuclear Area (NA) x Mean*
715 *Fluorescence of Background (MFB)*. For *in vitro* foci counting, quantification was performed
716 under the 100x lens of a Nikon Eclipse Ni microscope, with 20-100 cells/condition assessed (as
717 indicated in the respective figure legend). For quantification of mice brain sections, brainstem
718 or cerebellar sections were imaged on a confocal microscope, as described above, and analysed
719 manually, with one exception: For γH2AX analysis in mice expressing 0, 69 DPRs, foci
720 quantification was automated, using image J. For human spinal cord sections, images were
721 acquired using a Nikon Eclipse Ni microscope under the 100x and 20x objective lenses. Large
722 motor neurons located in the ventral horn of spinal cord sections were considered for analysis,
723 ~50 motor neurons (minimum 25) were analysed from 6 *C9orf72*-ALS samples and 6 non-ALS
724 samples. Motor neurons were considered R-Loop or γH2AX positive when the whole nucleus
725 was stained positive. We employed a double-blind randomization process in which
726 experimental groups and the protein being analysed were blinded to the person analysing the
727 data (e.g. counting nuclear foci or taking micrographs). For H3K9me3 analysis, 4 controls and
728 4 *C9orf72*-ALS were imaged using a Leica LP5 confocal microscope and the mean

729 fluorescence intensity (minus background signal) was calculated from 20 motor neurons per
730 case.

731 **Experimental repeats and Statistical analysis**

732 All data are presented as the means \pm standard errors of the mean (SEM) of 3 biological
733 replicates, unless otherwise stated. Statistical differences were analysed using Student's t-tests
734 for pair-wise comparisons or one-way ANOVA (with Tukey's correction) for comparing
735 groups more than 2 but less than 9. A p-value less than 0.05 was considered to be statistically
736 significant. Asterisks denote p values <0.05 , double-asterisks denote p values <0.01 , and triple
737 asterisks denote p values <0.001 , NS denotes p values >0.05 .

738

739 **Methods-only references**

740

- 741 51. McIntyre, G. J., Groneman, J. L., Tran, A. & Applegate, T. L. An infinitely expandable
742 cloning strategy plus repeat-proof PCR for working with multiple shRNA. *PLoS ONE* **3**,
743 e3827 (2008).
- 744 52. Cooper-Knock, J. *et al.* Sequestration of multiple RNA recognition motif-containing
745 proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 (2014).
- 746 53. Mead, R. J. *et al.* Optimised and rapid pre-clinical screening in the SOD1(G93A)
747 transgenic mouse model of amyotrophic lateral sclerosis (ALS). *PLoS ONE* **6**, e23244
748 (2011).
- 749 54. McCloy, R. A. *et al.* Partial inhibition of Cdk1 in G 2 phase overrides the SAC and
750 decouples mitotic events. *Cell Cycle* **13**, 1400–1412 (2014).
- 751

752 **Data availability**

753 The data that support the findings of this study are available from the corresponding authors
754 upon reasonable request.

755

756 **Competing financial interest**

757 The authors declare no competing financial interest

758 **References**

759

- 760 1. Schottlaender, L. V. *et al.* Analysis of C9orf72 repeat expansions in a large series of
761 clinically and pathologically diagnosed cases with atypical parkinsonism. *Neurobiol.*
762 *Aging* **36**, 1221.e1–6 (2015).
- 763 2. Rutherford, N. J. *et al.* Length of normal alleles of C9ORF72 GGGGCC repeat do not
764 influence disease phenotype. *Neurobiol. Aging* **33**, 2950.e5–7 (2012).
- 765 3. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding
766 region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256
767 (2011).
- 768 4. Groh, M. & Gromak, N. Out of Balance: R-loops in Human Disease. *PLoS Genet* **10**,
769 e1004630 (2014).
- 770 5. Roy, D., Yu, K. & Lieber, M. R. Mechanism of R-loop formation at immunoglobulin
771 class switch sequences. *Mol. Cell. Biol.* **28**, 50–60 (2008).
- 772 6. Aguilera, A. & García-Muse, T. R Loops: From Transcription Byproducts to Threats to
773 Genome Stability. *Molecular Cell* **46**, 115–124 (2012).
- 774 7. Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of
775 disease. *Nature* **507**, 195–200 (2014).
- 776 8. Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2
777 mRNA export factor PCID2. *Nature* **17**; 511(7509): 362-5 (2014).
- 778 9. Yuce-Petronczki, O. & West, S. C. Senataxin, defective in the neurodegenerative disorder
779 AOA-2, lies at the interface of transcription and the DNA damage response. *Mol. Cell.*
780 *Biol.* **33**(2): 406-17. (2012).
- 781 10. Lindquist, S. G. *et al.* Corticobasal and ataxia syndromes widen the spectrum of
782 C9ORF72 hexanucleotide expansion disease. *Clin. Genet.* **83**, 279–283 (2013).
- 783 11. Corcia, P. *et al.* Pure cerebellar ataxia linked to large C9orf72 repeat expansion.
784 *Amyotroph Lateral Scler Frontotemporal Degener.* **17**(3-4): 301-3 (2016).
- 785 12. McKinnon, P. J. ATM and the molecular pathogenesis of ataxia telangiectasia. *Annu.*
786 *Rev. Pathol. Mech. Dis.* **7**, 303–321 (2012).
- 787 13. Shiloh, Y. & Ziv, Y. The ATM protein kinase: regulating the cellular response to
788 genotoxic stress, and more. *Nature Reviews Molecular Cell Biology* **14**, 197–210 (2013).
- 789 14. Frappart, P.-O. *et al.* An essential function for NBS1 in the prevention of ataxia and
790 cerebellar defects. *Nat. Med.* **11**, 538–544 (2005).
- 791 15. Ashour, M. E., Atteya, R. & El-Khamisy, S. F. Topoisomerase-mediated chromosomal
792 break repair: an emerging player in many games. *Nat Rev Cancer* **15**, 137–151 (2015).
- 793 16. Gómez-Herreros, F. *et al.* TDP2 protects transcription from abortive topoisomerase
794 activity and is required for normal neural function. *Nat Genet* **46**, 516–521 (2014).
- 795 17. Chiang, S.-C. *et al.* Mitochondrial protein-linked DNA breaks perturb mitochondrial
796 gene transcription and trigger free radical-induced DNA damage. *Science Advances* **3**,
797 e1602506–16 (2017).
- 798 18. Harrigan, J. A. *et al.* Replication stress induces 53BP1-containing OPT domains in G1
799 cells. *The Journal of Cell Biology* **193**, 97–108 (2011).
- 800 19. Heo, J.-I. *et al.* ATM mediates interdependent activation of p53 and ERK through
801 formation of a ternary complex with p-p53 and p-ERK in response to DNA damage.
802 *Mol. Biol. Rep.* **39**, 8007–8014 (2012).
- 803 20. Katyal, S. *et al.* Aberrant topoisomerase-1 DNA lesions are pathogenic in
804 neurodegenerative genome instability syndromes. *Nat. Neurosci.* **17**, 813–821 (2014).
- 805 21. Alagoz, M., Chiang, S.-C., Sharma, A. & El-Khamisy, S. F. ATM deficiency results in
806 accumulation of DNA-topoisomerase I covalent intermediates in neural cells. *PLoS ONE*

- 807 **8**, e58239 (2013).
- 808 22. Li, J. *et al.* Nuclear accumulation of HDAC4 in ATM deficiency promotes
809 neurodegeneration in ataxia telangiectasia. *Nat. Med.* **18**, 783–790 (2012).
- 810 23. Zhang, Y.-J. *et al.* C9ORF72 poly(GA) aggregates sequester and impair HR23 and
811 nucleocytoplasmic transport proteins. *Nat. Neurosci.* **19**, 668–677 (2016).
- 812 24. Difilippantonio, S. *et al.* Role of Nbs1 in the activation of the Atm kinase revealed in
813 humanized mouse models. *Nat. Cell Biol.* **7**, 675–685 (2005).
- 814 25. Schwertman, P., Bekker-Jensen, S. & Mailand, N. Regulation of DNA double-strand
815 break repair by ubiquitin and ubiquitin-like modifiers. *Nature Reviews Molecular Cell*
816 *Biology* **17**, 379–394 (2016).
- 817 26. Shanbhag, N. M., Rafalska-Metcalf, I. U., Balane-Bolivar, C., Janicki, S. M. &
818 Greenberg, R. A. ATM-dependent chromatin changes silence transcription in cis to DNA
819 double-strand breaks. *Cell* **141**, 970–981 (2010).
- 820 27. Bohgaki, M. *et al.* RNF168 ubiquitylates 53BP1 and controls its response to DNA
821 double-strand breaks. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20982–20987 (2013).
- 822 28. Fradet-Turcotte, A. *et al.* 53BP1 is a reader of the DNA-damage-induced H2A Lys 15
823 ubiquitin mark. *Nature* **499**, 50–54 (2013).
- 824 29. Doil, C. *et al.* RNF168 binds and amplifies ubiquitin conjugates on damaged
825 chromosomes to allow accumulation of repair proteins. *Cell* **136**, 435–446 (2009).
- 826 30. Noon, A. T. *et al.* 53BP1-dependent robust localized KAP-1 phosphorylation is essential
827 for heterochromatic DNA double-strand break repair. *Nat Cell Biol* **12**, 177–184 (2010).
- 828 31. Lee, J.-H., Goodarzi, A. A., Jeggo, P. A. & Paull, T. T. 53BP1 promotes ATM activity
829 through direct interactions with the MRN complex. *EMBO J.* **29**, 574–585 (2010).
- 830 32. Stewart, G. *et al.* The RIDDLE syndrome protein mediates a ubiquitin-dependent
831 signaling cascade at sites of DNA damage. *Cell* **136**(3): 420-34 (2009).
- 832 33. Wang, Y. *et al.* Autophagy Regulates Chromatin Ubiquitination in DNA Damage
833 Response through Elimination of SQSTM1/p62. *Molecular Cell* **63**, 34–48 (2016).
- 834 34. Mann, D. M. A. *et al.* Dipeptide repeat proteins are present in the p62 positive inclusions
835 in patients with frontotemporal lobar degeneration and motor neurone disease associated
836 with expansions in C9ORF72. *Acta Neuropathol Commun* **1**, 68 (2013).
- 837 35. Davidson, Y. S. *et al.* Brain distribution of dipeptide repeat proteins in frontotemporal
838 lobar degeneration and motor neurone disease associated with expansions in C9ORF72.
839 *Acta Neuropathol Commun* **2**, 70 (2014).
- 840 36. Tresini, M., Marteijn, J. A. & Vermeulen, W. Bidirectional coupling of splicing and
841 ATM signaling in response to transcription-blocking DNA damage. *mbiology* **13**, 1–8
842 (2017).
- 843 37. Tresini, M. *et al.* The core spliceosome as target and effector of non-canonical ATM
844 signalling. *Nature* **523**(7558): 53-8 (2015).
- 845 38. Bakkenist, C. J. & Kastan, M. B. DNA damage activates ATM through intermolecular
846 autophosphorylation and dimer dissociation. *Nature* **421**, 499–506 (2003).
- 847 39. Belzil, V. V. *et al.* Reduced C9orf72 gene expression in c9FTD/ALS is caused by
848 histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* **126**,
849 895–905 (2013).
- 850 40. Goodarzi, A. A. *et al.* ATM signaling facilitates repair of DNA double-strand breaks
851 associated with heterochromatin. *Mol. Cell* **31**, 167–177 (2008).
- 852 41. Meisenberg, C. *et al.*, Epigenetic changes in histone acetylation underpin resistance to
853 the topoisomerase I inhibitor irinotecan. *Nucleic Acids Res.* **45**(3): 1159-1176 (2017).
- 854 42. Felisbino, M. B., Gatti, M. S. V. & Mello, M. L. S. Changes in chromatin structure in
855 NIH 3T3 cells induced by valproic acid and trichostatin A. *J Cell Biochem* **115**, 1937–
856 1947 (2014).

- 857 43. Sanz, L.A.1 *et al.* Prevalent, dynamic, and conserved R-Loop structures associate with
858 specific epigenomic signatures in mammal. *Molecular Cell* **63(1)**: 167-78 (2016).
- 859 44. Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and
860 chromatin condensation. *Molecular Cell* **52**, 583–590 (2013).
- 861 45. Zhang, Y. *et al.* Attenuated DNA damage repair by trichostatin A through BRCA1
862 suppression. *Radiat. Res.* **168**, 115–124 (2007).
- 863 46. Wang, X. *et al.* Alzheimer disease and amyotrophic lateral sclerosis: an etiopathogenic
864 connection. *Acta Neuropathol.* **127**, 243–256 (2014).
- 865 47. El-Khamisy, S. F. *et al.* Defective DNA single-strand break repair in spinocerebellar
866 ataxia with axonal neuropathy-1. *Nature* **434**, 108–113 (2005).
- 867 48. El-Khamisy, S. F. To live or to die: a matter of processing damaged DNA termini in
868 neurons. *EMBO Mol Med* **3**, 78–88 (2011).
- 869 49. Chan, Y. A. *et al.* Genome-wide profiling of yeast DNA:RNA hybrid prone sites with
870 DRIP-chip. *PLoS Genet* **10**, e1004288 (2014).
- 871 50. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS
872 and frontotemporal dementia. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4968–77 (2013).
- 873
- 874
- 875

Figure 1

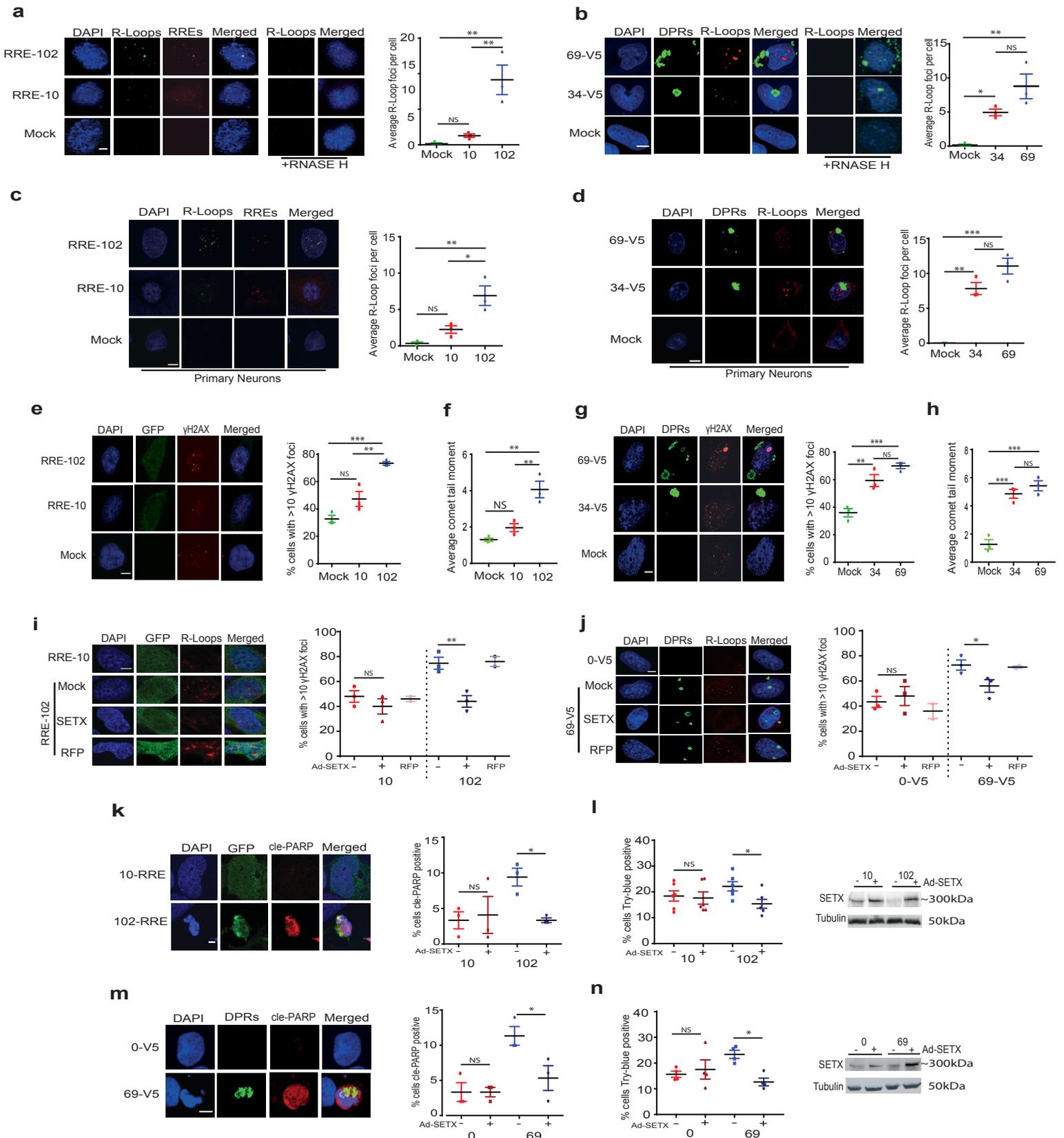


Figure 1. Expression of C9orf72 expansions leads to R-loop-driven DSBs and cellular toxicity. (a) MRC5 cells mock transfected or transfected with 10 or 102 RREs. FISH-IF was performed using a G4C2 fluorescent probe 'RNA' and S9.6 antibodies 'R-Loops'. Cells were treated with RNase H1 '+RNASE H'. *Left*, Representative images shown, scale bar 5 μm. *Right*, The average (± SEM) number of nuclear S9.6 foci per cell was quantified from 3 cell culture replicates, 50 cells each. Significance assessed using a one-way ANOVA. (b) MRC5 cells mock transfected or transfected with 34 or 69 poly-GA DPRs. Cells examined by immunocytochemistry using anti-V5 'DPRs' and S9.6 antibodies 'R-Loops'. *Left*, Representative images are shown, scale bar 5 μm. *Right*, S9.6 foci was quantified, presented, and analysed as described for (a). (c,d) Rat cortical neurons transfected with AAV9 viral-vectors encoding 10,102 RREs (c) or 34, 69 DPRs (d) were processed with FISH-IF double staining (c) or with immunocytochemistry (d), as described for (a,b). (e,g) MRC5 cells mock transfected or transfected with 10,102 RREs (with GFP) (e) or 34, 69 DPRs (g). Cells were immunostained with anti-γH2AX antibodies 'γH2AX', with GFP (e) or anti-V5 antibodies 'DPRs' (g). *Left*, Representative images are shown, scale bar 5 μm. *Right*, The percentage of cells with 10 or more foci was quantified, presented and analysed as described for (a). (f,h) HEK 293T cells mock transfected, transfected with 10,102 RREs (f), or 34, 69 DPRs (h). Neutral comet tail moments were quantified, 100 cells each, presented, and analysed as described for (a). (i-j) MRC5 cells mock transfected or transfected with adenoviral vectors encoding for SETX or RFP and then transfected with 10 or 102 RREs (with GFP) (i) or with 0, 69 DPRs (j). *Left*, Cells were immunostained with S9.6 antibodies 'R-Loops' alongside GFP (i) or alongside anti-V5 'DPRs' antibodies (j). Representative images are shown, scale bar 5 μm. *Right*, Cells were immunostained with anti-γH2AX antibodies as described for panels (e,f), and the average (± SEM) percentage of cells exhibiting 10 or more γH2AX foci was quantified, 25 cells each, and analysed using Student's t-test. (k,m) MRC5 cells transfected with adenoviral vector particles encoding for SETX or mock transfected and transfected with constructs encoding 10, 102 RREs (k) or 0 or 69 DPRs (m). Cells examined by immunocytochemistry using cleaved-PARP (Cell Signalling, 9548) 'cle-PARP' antibodies alongside GFP (k) or anti-V5 (Bethyl, A190-120A) 'DPRs' antibodies (m). *Left*, Representative images of cle-PARP-positive and -negative cells shown, scale bar 5 μm. *Right*, the percentage of cells cleaved-PARP-positive was quantified, 50-100 cells each, presented and analysed as described for (i,j). (l,n) HEK 293T cells were mock transfected or transfected with adenoviral vector particles encoding for SETX and transfected with 10,102 RREs (l) or 0 or 69 DPRs (n). *Left*, Cells were analysed using Trypan blue exclusion assays, and the % of cells Trypan-permeable was quantified from 6 (l) and 4 (n) cell culture replicates, ~200 cells each, and was presented and analysed as described for (i,j). *Right*, Whole cell lysates from samples used in (l) and (n) were analysed by western blotting, using senataxin and anti-α-tubulin antibodies.

Figure 2

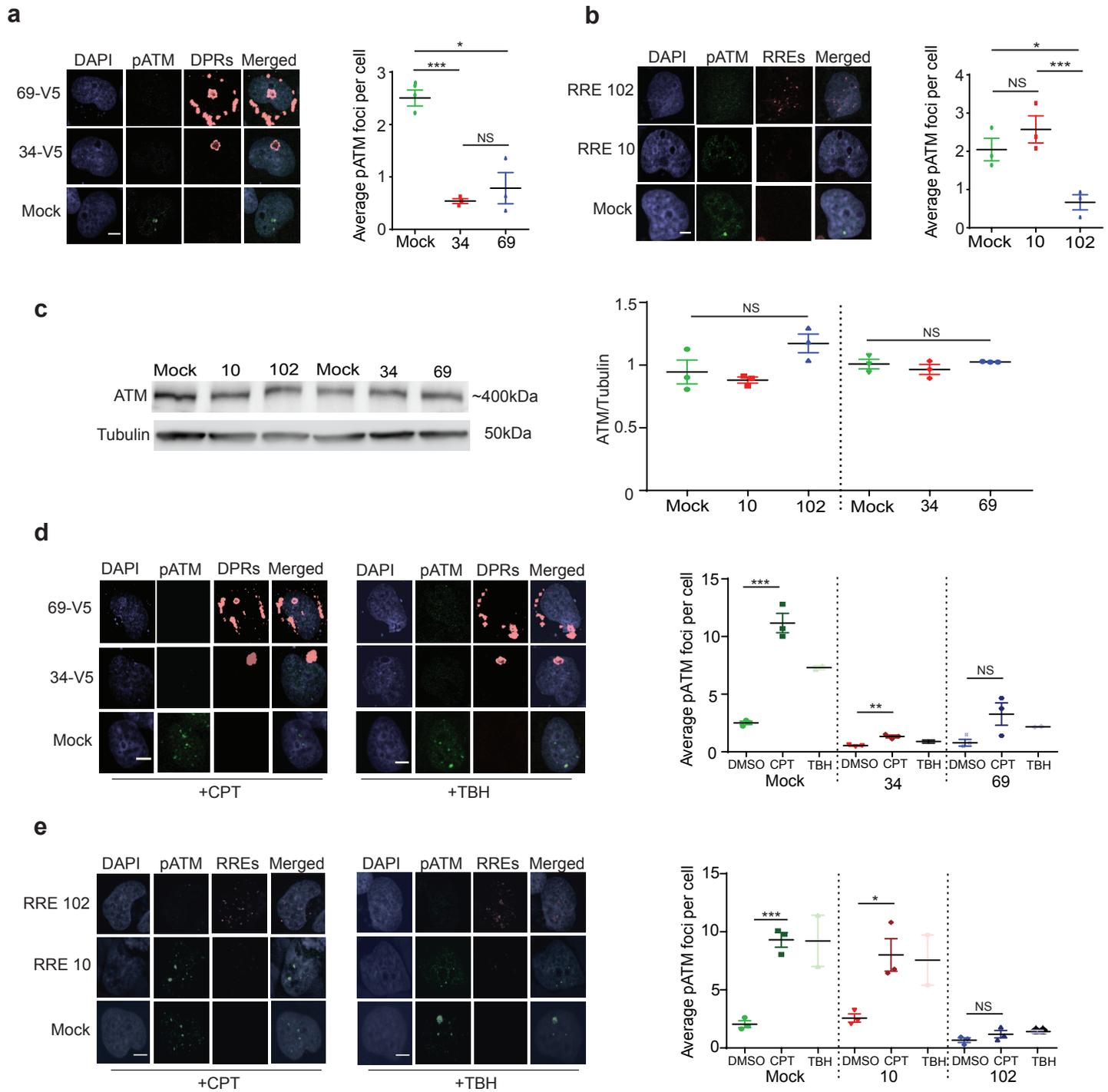


Figure 2: Expression of C9orf72 expansions leads to defective ATM activation. (a,b) MRC5 cells mock transfected, transfected with 34, 69 DPRs (a), or 10, 102 RREs (b). Cells analysed using immunocytochemistry with anti-phosphorylated ATM 'pATM' and anti-V5 'DPRs' antibodies (a) or with FISH-IF (b). *Left*, Representative images shown, scale bar 5 μ m. *Right*, The average (\pm SEM) number of pATM foci per cell was quantified from 3 cell culture replicates, 50 cells each, and analysed using a one-way ANOVA. (c) *Left*, HEK 293T cells were mock transfected 'M' or were transfected constructs encoding 10 or 102 RREs '10, 102' or 34, 69 DPRs '34, 69'. Whole cell lysates were analysed using anti-ATM and α -tubulin antibodies. *Right*, ATM protein expression (normalised to α -tubulin) is presented as average \pm SEM from 3 cell culture replicates, and analysed using a one-way ANOVA. (d,e) MRC5 cells mock transfected, transfected with constructs encoding 34, 69 DPRs (d), or transfected with constructs encoding for 10, 102 RREs (e). Cells were incubated with 10 μ M CPT, 0.037% TBH, or DMSO for 1 hour, and analysed by immunocytochemistry as described for (a,b). *Left*, Representative images of 3 cell culture replicates are shown, scale bar 5 μ m. *Right*, pATM foci were quantified as described for (a,b).

Figure 3

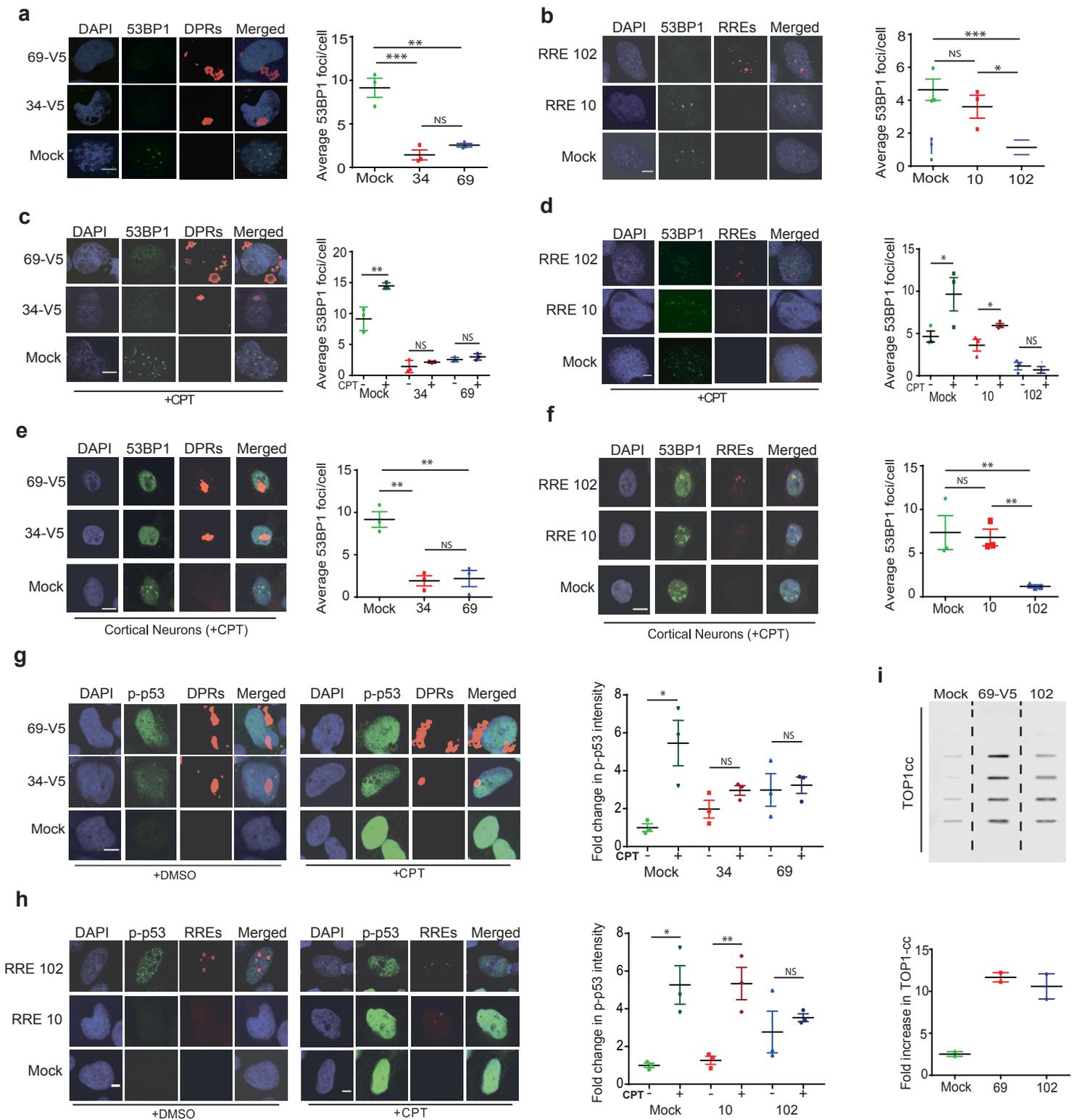


Figure 3: Expression of C9orf72 expansions leads to defective ATM signalling. (a,b) MRC5 cells mock transfected or transfected with 34 or 69 DPRs (a) or 10, 102 RREs (b), and immunostained with anti-53BP1 antibodies, alongside anti-V5 'DPRs' antibodies or with FISH-IF double-staining 'RNA' (b). *Left*, Representative images are shown, scale bar 5µm. *Right*, The average (± SEM) number of 53BP1 foci per cell was quantified from 3 cell culture replicates, 50 cells each, and analysed using a one-way ANOVA. (c,d) MRC5 cells were mock transfected or transfected with constructs encoding 34 or 69 DPRs (c), or 10, 102 RREs (d). Cells were incubated with 10µM CPT or DMSO for 1hour, and immunostained as described for (a,b). *Left*, Representative images are shown, scale bar 5µm. *Right*, 53BP1 foci was quantified as described above and analysed using a Student's t-test. (e,f) Rat cortical neurons mock transduced or transduced with AAV9 viral-vectors expressing 34 or 69 DPRs (e) or with 10 or 102 RREs (f). Neurons treated with 10µM CPT for 1 hour and analysed by immunocytochemistry as described for (a,b). *Left*, Representative images are shown, scale bar 5µm. *Right*, 53BP1 foci was quantified as described for (a,b), 20 neurons each, and analysed using a one-way ANOVA. (g,h) MRC5 cells were mock transfected or transfected with constructs encoding 34 or 69 DPRs (g), or 10, 102 RREs (h) and were then treated with DMSO or with 10µM CPT. Cells were then immunostained with anti-phosphorylated p53 antibodies, alongside anti-V5 'DPRs' antibodies (g) or with FISH-IF double-staining 'RNA' (h). *Left*, Representative images are shown, scale bar 5µm. *Right*, The nuclear fluorescence value for 50 nuclei was quantified from 3 cell culture replicates, and presented as the average (± SEM) fold change in nuclear intensity (relative to control cells), and analysed using Student's t-test. (i) *Top*, HEK 293T cells were mock transfected or were transfected with 69 DPRs or 102 RREs. Cells were treated 10µM CPT for 40 min, subjected to CsCl step gradients, and fractions slot blotted with anti-TOP1 antibodies. *Bottom*, The fold increase in TOP1-ccs, normalised to mock was calculated and presented as the average from 2 cell culture replicates ± range.

Figure 4

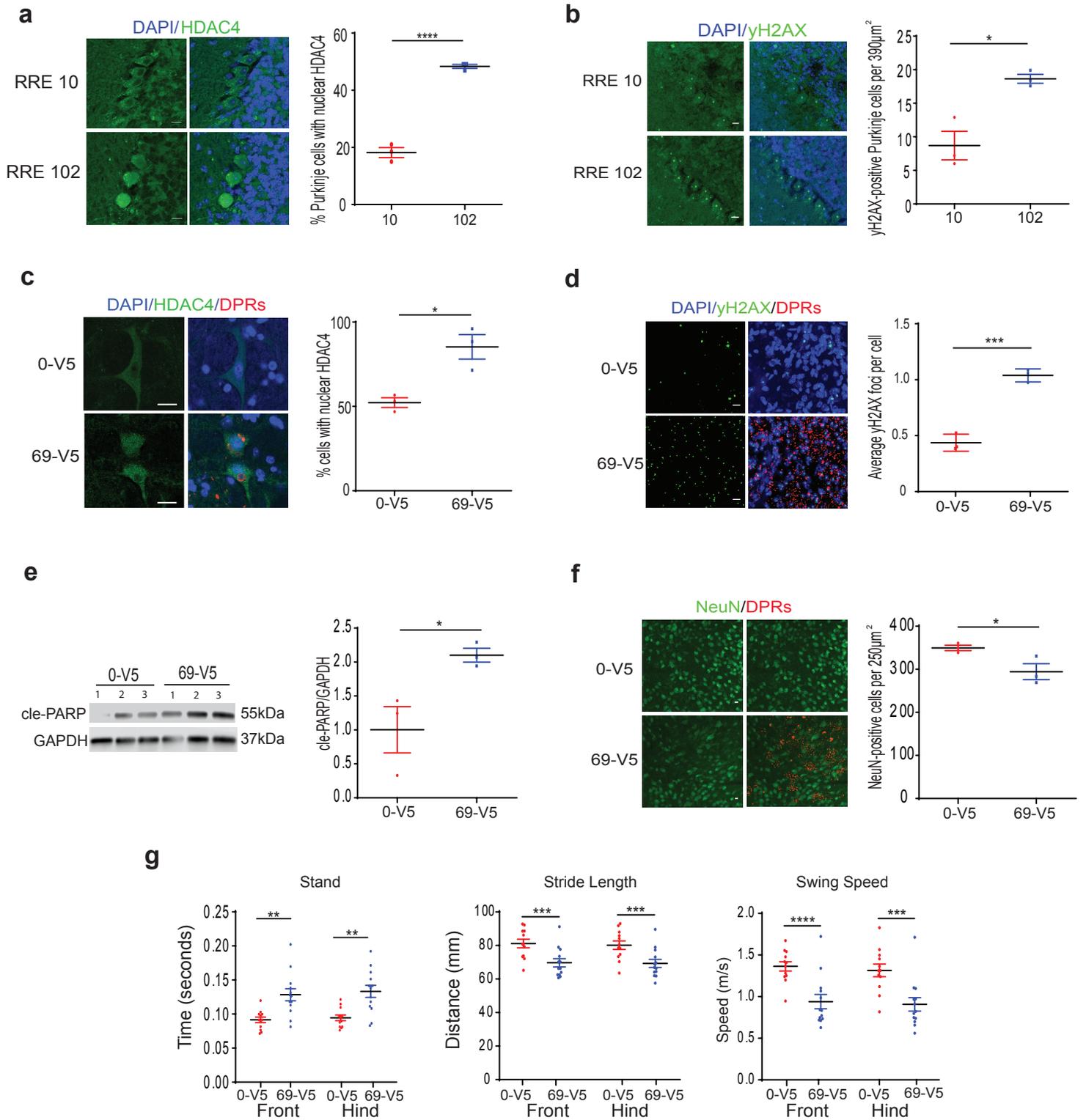


Figure 4: The expression of C9orf72 expansions in the murine CNS leads to DSBs, nuclear HDAC4, and neurodegeneration. (a) Cerebellar sections from mice injected with AAV9-10 or -102 RREs subjected to immunohistochemistry using anti-HDAC4 antibodies. *Left*, Representative images shown, scale bar 10 μm . *Right*, The average (\pm SEM) percentage of Purkinje cells displaying nuclear HDAC4 was calculated for 3 animals per group, 50 Purkinje cells per animal, and analysed using a Student's t-test. (b) Cerebellar sections from mice injected with AAV9-10 or -102 RREs subjected to immunohistochemistry using anti-yH2AX antibodies. *Left*, Representative images shown, scale bar 10 μm . *Right*, The average (\pm SEM) number of yH2AX-positive Purkinje cells was calculated from 3 animals per group, 10 images each, and analysed using a Student's t-test. (c) Brainstem sections from mice injected with AAV9-0 or -69 poly-GA DPRs were subjected to immunohistochemistry using anti-HDAC4 and anti-V5 antibodies. *Left*, Representative images shown, scale bar 10 μm . *Right*, The average (\pm SEM) percentage of brainstem cells displaying nuclear HDAC4 was calculated for 3 animals per group, 30 HDAC4-positive cells per animal, and analysed using a Student's t-test. (d) Brainstem sections from mice injected with AAV9-0 or -69 poly-GA DPRs were subjected to immunohistochemistry using anti-yH2AX and anti-V5 antibodies. *Left*, Representative images shown, scale bar 10 μm . *Right*, The average (\pm SEM) number of yH2AX foci per cell calculated from 3 animals per group, \sim 1000 cells per animal, and analysed using Student's t-test. (e) *Left*, Brainstem tissue harvested from mice injected with AAV9-0 or -69 poly-GA DPRs were analysed using Western blotting, with anti-GAPDH and anti-cleaved PARP antibodies. *Right*, cleaved-PARP was quantified and normalised to GAPDH, presented as the average intensity \pm SEM from 3 animals per group, and analysed using a Student's t-test. (f) *Left*, Brainstem sections from mice injected with AAV9-0 or -69 poly-GA DPRs were subjected to immunohistochemistry using anti-NeuN and anti-V5 antibodies. The average (\pm SEM) number of NeuN-positive cells within the periaqueductal gray region of the brainstem was quantified from 3 animals per group, and analysed using Student's t-test. (g) Catwalk analysis was performed in animals injected with AAV9-0 or -69 poly-GA DPRs, aged 6 months. Stand intensity, stride length, and swing speed were quantified ($n=12/13$ for 0-V5/69-V5), presented as average \pm SEM, analysed using Student's t-test.

Figure 5

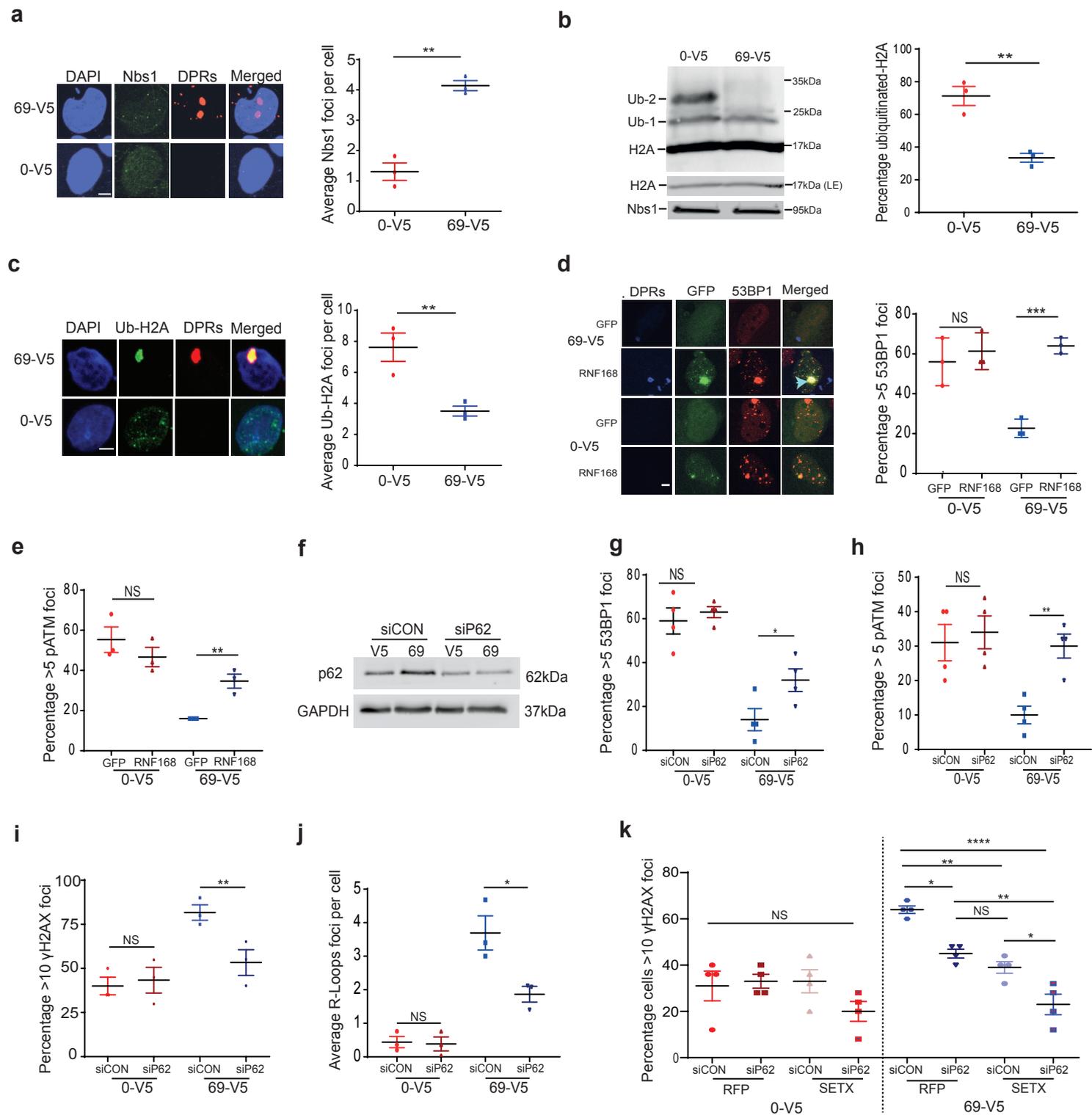


Figure 5: Defective ATM-mediated DNA repair can be restored by RNF168 overexpression or p62 depletion. (a) MRC5 cells transfected with 0 or 69 DPRs, and immunostained with anti-V5 'DPRs' and anti-Nbs1 antibodies. *Left*, Representative images are shown, scale bar 5µm. *Right*, The average (± SEM) number of Nbs1 foci per cell was quantified and from 3 cell culture replicates, 50 cells each, and analysed using a Student's t-test. (b) *Left*, Chromatin fractions from MRC5 cells transfected with 0 or 69 DPRs were separated and analysed with Western blotting using antibodies specific to H2A. Low exposure (LE) H2A and Nbs1 bands show equal loading. *Right*, The average (± SEM) percentage of H2A that was ubiquitinated was quantified from 3 cell culture replicates and analysed using Student's t-test. (c) *Left*, MRC5 cells transfected with constructs encoding 0 or 69 DPRs, and immunostained with anti-V5 'DPRs' and anti-Ubiquitinated-H2A 'Ub-H2A' antibodies. Representative images are shown, scale bar 5µm. *Right*, Ub-H2A foci was quantified, 25 cells each, presented and analysed as described for (a). (d) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, and with control-GFP or RNF168-GFP, were immunostained with anti-V5 and anti-53BP1 antibodies. *Left*, Representative images are shown. *Right*, the percentage of cells with 5 or more 53BP1 foci was quantified (25 cells each), presented, and analysed as described for (a). (e) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, and with control-GFP or RNF168-GFP plasmids, were immunostained with anti-V5 and anti-pATM antibodies. The percentage of cells with 5 or more pATM foci was quantified (25 cells each), presented and analysed as described for (a). (f) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, with either control siRNA particles or p62 siRNA particles, were analysed with Western blotting using antibodies specific to p62 and GAPDH. (g) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, and with control siRNA particles or p62 siRNA particles, were immunostained with anti-V5 and anti-53BP1 antibodies. The percentage of cells with 5 or more 53BP1 foci was quantified from 4 cell culture replicates (25 cells each), presented and analysed as described for (a). (h) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, and with control siRNA particles or p62 siRNA particles, were immunostained with anti-V5 and anti-pATM antibodies. The percentage of cells with 5 or more pATM foci was quantified from 4 cell culture replicates (25 cells each), presented and analysed as described for (a). (i) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, with control siRNA particles or p62 siRNA particles. Cells were immunostained with anti-V5 'DPRs' and anti-γH2AX antibodies. The percentage of cells with 10 or more γH2AX foci was quantified (25 cells each), presented, and analysed as described for (a). (j) MRC5 cells transfected with constructs encoding 0 or 69 DPRs, with control siRNA particles or p62 siRNA particles, were immunostained with anti-V5 'DPRs' and anti-S9.6 antibodies. R-Loop foci were quantified, presented and analysed as described for (a). (k) MRC5 cells transduced with adenoviral vectors encoding for SETX or RFP, transfected with constructs encoding 0 or 69 DPRs, with control siRNA or p62 siRNA particles, were immunostained with anti-V5 and anti-γH2AX antibodies. Nuclei were counterstained with DAPI. The average (± SEM) percentage of cells with 10 or more γH2AX foci was quantified from 4 cell culture replicates (25 cells each), and analysed using a one-way ANOVA.

Figure 6

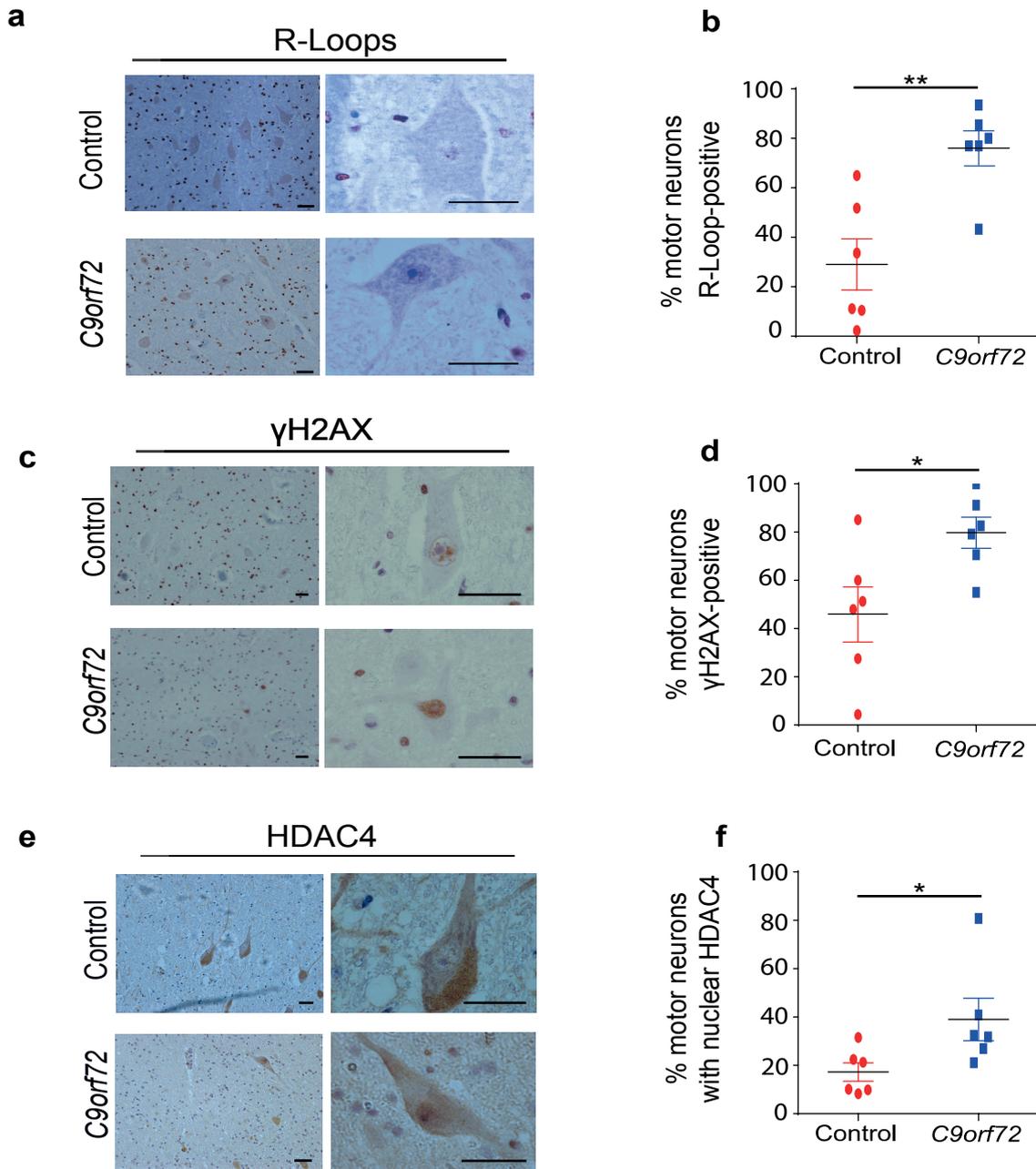


Figure 6. Spinal Cord Motor Neurons from *C9orf72*-ALS post-mortem show elevated levels of R-Loops, DSBs, and nuclear HDAC4. (a) Human spinal cord sections were analysed by immunohistochemistry using S9.6 antibodies. Representative images are presented, scale bar 5μm. (b) The average (± SEM) percentage of R-Loop-positive motor neurons was quantified from 6 *C9orf72* patient and 6 control sections, ~50 cells each, and analysed with Student's t-test. (c) Human spinal cord sections were analysed by immunohistochemistry using anti-γH2AX antibodies. Representative images are presented, scale bar 5 μm. (d) The % of γH2AX-positive motor neurons was quantified, presented and analysed as described for (b). (e) Human spinal cord sections were analysed by immunohistochemistry using anti-HDAC4 antibodies. Representative images are presented, scale bar 5 μm. (f) The % of motor neurons displaying HDAC4 enrichment in the nucleus was quantified, presented and analysed as described for (b).