SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage

Mohini Jangi[a,1], Christina Fleet[a], Patrick Cullen[a], Shipra V. Gupta[a], Shila Mekhoubad[b], Eric Chiao[b,2], Norm Allaire[a], C. Frank Bennett[c], Frank Rigo[c], Adrian R. Krainer[d], Jessica A. Hurt[a], John P. Carulli[a,3], and John F. Staropoli[e,3]

[a]Computational Biology & Genomics, Biogen, Cambridge, MA 02142; [b]Stem Cell Research, Biogen, Cambridge, MA 02142; [c]Neuroscience Drug Discovery, Ionis Pharmaceuticals, Carlsbad, CA 92008; [d]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; and [e]Rare Disease, Biogen, Cambridge, MA 02142

Edited by James L. Manley, Columbia University, New York, NY, and approved February 7, 2017 (received for review August 8, 2016)

Spinal muscular atrophy (SMA), an autosomal recessive neuromuscular disease, is the leading monogenic cause of infant mortality. Homozygous loss of the gene survival of motor neuron 1 (*SMN1*) causes the selective degeneration of lower motor neurons and subsequent atrophy of proximal skeletal muscles. The *SMN1* protein product, survival of motor neuron (SMN), is ubiquitously expressed and is a key factor in the assembly of the core splicing machinery. The molecular mechanisms by which disruption of the broad functions of SMN leads to neurodegeneration remain unclear. We used an antisense oligonucleotide (ASO)-based inducible mouse model of SMA to investigate the SMN-specific transcriptome changes associated with neurodegeneration. We found evidence of widespread intron retention, particularly of minor U12 introns, in the spinal cord of mice 30 d after SMA induction, which was then rescued by a therapeutic ASO. Intron retention was concomitant with a strong induction of the p53 pathway and DNA damage response, manifesting as γ-H2A.X positivity in neurons of the spinal cord and brain. Widespread intron retention and markers of the DNA damage response were also observed with SMN depletion in human SH-SY5Y neuroblastoma cells and human induced pluripotent stem cell-derived motor neurons. We also found that retained introns, high in GC content, served as substrates for the formation of transcriptional R-loops. We propose that defects in intron removal in SMA promote DNA damage in part through the formation of RNA:DNA hybrid structures, leading to motor neuron death.

SMA | SMN | DNA damage | neurodegeneration | splicing

Spinal muscular atrophy (SMA) is a devastating autosomal recessive neuromuscular disease characterized by the progressive loss of α-motor neurons from the medial ventral horn of the spinal cord (1). The most common form of SMA, type I, manifests by age 6 mo as an advancing proximal paralysis that leads to respiratory distress, and affected children generally require permanent ventilation or succumb to the disease by age 2 y (2). Consequently, SMA is the most frequent monogenic cause of infant mortality, with a prevalence of approximately 1 in 10,000 live births (1). In recent years, much progress has been made in understanding the genetic and molecular underpinnings of the disease.

SMA is caused by homozygous mutation or deletion of the survival of motor neuron 1 (*SMN1*) gene (3). In humans, the paralogous *SMN2* gene arose from a gene duplication event and is located centromeric to the *SMN1* locus. The two paralogs are nearly identical but differ by several key nucleotides within and flanking exon 7 (4). As a result, transcripts arising from the *SMN2* locus predominantly splice out exon 7, producing a protein that is rapidly degraded and thought to be nonfunctional; however, a small fraction of transcripts include exon 7 and encode a protein identical to SMN1 (5). Importantly, *SMN2* copy number varies across individuals and is a modifier of disease severity, with a higher copy number leading to reduced severity. We previously characterized the function of an antisense oligonucleotide (ASO) that binds and blocks an exonic splicing silencer in exon 7 of SMN2 pre-mRNA, enhancing the inclusion of exon 7 and the production of full-length protein when delivered to the cerebrospinal fluid in an inducible mouse model of type I SMA (6, 7). We found that early treatment with an ASO that promoted SMN2 exon 7 inclusion prevented and reversed the gene expression changes that occurred on disease induction (6). Recently, nusinersen, an ASO drug that promotes exon 7 inclusion, was approved to treat all forms of SMA.

With the rapid approval of nusinersen for a broad patient population, a molecular understanding of why SMN loss results in motor neuron death is critical for monitoring therapeutic efficacy and enabling further progress. SMN RNA and protein are widely expressed during development and in adulthood (4, 8, 9); however, particularly high expression in ventral motor neurons of the spinal cord from the second trimester of life into adulthood may imply an increased demand for SMN protein in this cell population (10). SMN has been shown to play both "housekeeping" and cell type-specific roles in

Significance

Spinal muscular atrophy is the leading monogenic cause of infant mortality and is caused by homozygous loss of the survival of motor neuron 1 (*SMN1*) gene. We investigated global transcriptome changes in the spinal cord of inducible SMA mice. SMN depletion caused widespread retention of introns with weak splice sites or belonging to the minor (U12) class. In addition, DNA double strand breaks accumulated in the spinal cord of SMA mice and in human SMA cell culture models. DNA damage was partially rescued by suppressing the formation of R-loops, which accumulated over retained introns. We propose that instead of single gene effects, pervasive splicing defects caused by severe SMN deficiency trigger a global DNA damage and stress response, thus compromising motor neuron survival.

Author contributions: M.J., J.P.C., and J.F.S. designed research; M.J. and C.F. performed research; P.C., S.V.G., S.M., E.C., N.A., C.F.B., F.R., A.R.K., and J.A.H. contributed new reagents/analytic tools; M.J., J.A.H., J.P.C., and J.F.S. analyzed data; and M.J. wrote the paper.

Conflict of interest statement: C.F., P.C., S.V.G., S.M., N.A., J.A.H., J.P.C., and J.F.S. are employees of Biogen. M.J. and E.C. were employed at Biogen at the time the work was performed, but are now employed at Grail and Regeneron, respectively. C.F.B. and F.R. are employees of Ionis Pharmaceuticals. A.R.K. serves as a consultant to Ionis Pharmaceuticals.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE87281).

[1]Present address: GRAIL, Inc., Menlo Park, CA 94025.

[2]Present address: Cell Technologies, Regeneron Pharmaceuticals, Tarrytown, NY 10591

[3]To whom correspondence may be addressed. Email: john.carulli@biogen.com or john.staropoli@biogen.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1613181114/-/DCSupplemental.

ribonucleoprotein (RNP) assembly and RNA metabolism. The most well-understood role of SMN in the cell is in the assembly of the spliceosomal small nuclear ribonucleoproteins (snRNPs), which are required for the catalysis of intron removal during pre-mRNA splicing (11, 12). In vivo, SMN forms a complex composed of Gemins 2–8 and Unr-interacting protein (Unrip) that can be found localized to nuclear bodies called "gems" at steady state (13, 14). In its active role in the cytoplasm, the SMN complex assembles a heptameric ring of Sm proteins around each U-rich small nuclear RNA (snRNA) (U1, U2, U4, U4atac, U5, U11, or U12) to form a mature snRNP (15). These assembled snRNPs can then catalyze splicing on translocation to the nucleus. Several groups have suggested a function for SMN beyond a role in splicing, in the assembly of axonal granules and mRNA transport within axons (16–18). Although this may partially explain the neuronal bias of SMA pathology, conclusive evidence of how SMN-dependent axonal transport contributes specifically to the SMA motor neuron phenotype remains to be determined.

Previous studies have focused on the link between snRNP assembly and splicing defects, and how missplicing of specific transcripts from SMN loss can lead to motor neuron dysfunction in SMA. Consistent with a fundamental role for SMN in snRNP assembly, Dreyfuss et al. (19) identified distinct patterns of aberrant splicing by exon array analysis across several tissues of a severe mouse model of SMA and attributed these defects to tissue-specific alterations in snRNA levels. Presymptomatic SMA mice exhibit motor neuron-specific skipping of the agrin Z exons, which are involved in postsynaptic organization of acetylcholine receptors at the neuromuscular junction and may contribute to the NMJ defects seen in human patients (20). Other models have suggested that locomotor defects arising from SMN loss are independent of defective snRNP biogenesis, and that developmental delays in SMN-deficient animals are responsible for such apparent splicing differences (21, 22).

Inconsistencies in transcriptome changes between models, as well as evidence that many of these splicing changes are secondary to SMN loss (23), warrant a closer examination of the nature of the transcriptomic defects occurring in SMA and their relation to the SMA phenotype. Several previous studies, including ours, have reported strong up-regulation of Cdkn1a/p21, a cyclin-dependent kinase inhibitor and mediator of cell cycle arrest downstream of p53, as well as other regulators of cell cycle and DNA repair in SMA models (6, 19, 23–25). Whereas the Cdkn1a up-regulation has been attributed to a p53-independent stabilization of Cdkn1a transcript in the absence of SMN protein (26, 27), few other gene expression changes can be explained by such posttranscriptional regulation.

In an effort to understand the link between RNA processing changes and gene-level changes in SMA and to further investigate the Cdkn1a activation, we characterized spinal cord transcriptome changes by RNAseq at high depth in an ASO-mediated inducible mouse model of type I SMA, which bypasses confounding developmental changes (6, 7). Expanding on the aforementioned studies showing snRNP assembly defects and splicing changes, we found that thousands of introns in functionally diverse genes were retained in a manner correlating with splice site strength, such that introns with weaker splicing signals were more sensitive to SMN depletion. Importantly, we found that the p21 activation was part of a larger p53 response to DNA double-strand breaks that occurred in the absence of overt apoptosis in both the mouse model and in two human cell culture systems.

Here we propose a unifying model whereby defects in RNA processing, particularly in intron retention, trigger DNA damage and subsequent neurodegeneration through the formation of transcriptional R-loops. We found increased R-loops over introns retained on SMN knockdown in SH-SY5Y cells, as well as downstream of transcription termination sites, extending a recent study suggesting a direct role for SMN in R-loop resolution (28). Our results argue for a broad role for SMN in the prevention of genome instability and neurodegeneration through the proper control of RNA processing.

Results

SMN Depletion Induces a Specific Subset of Splicing Changes. The demonstrated role of SMN in the assembly of spliceosomal snRNPs suggests that splicing dysregulation may be an early hallmark of

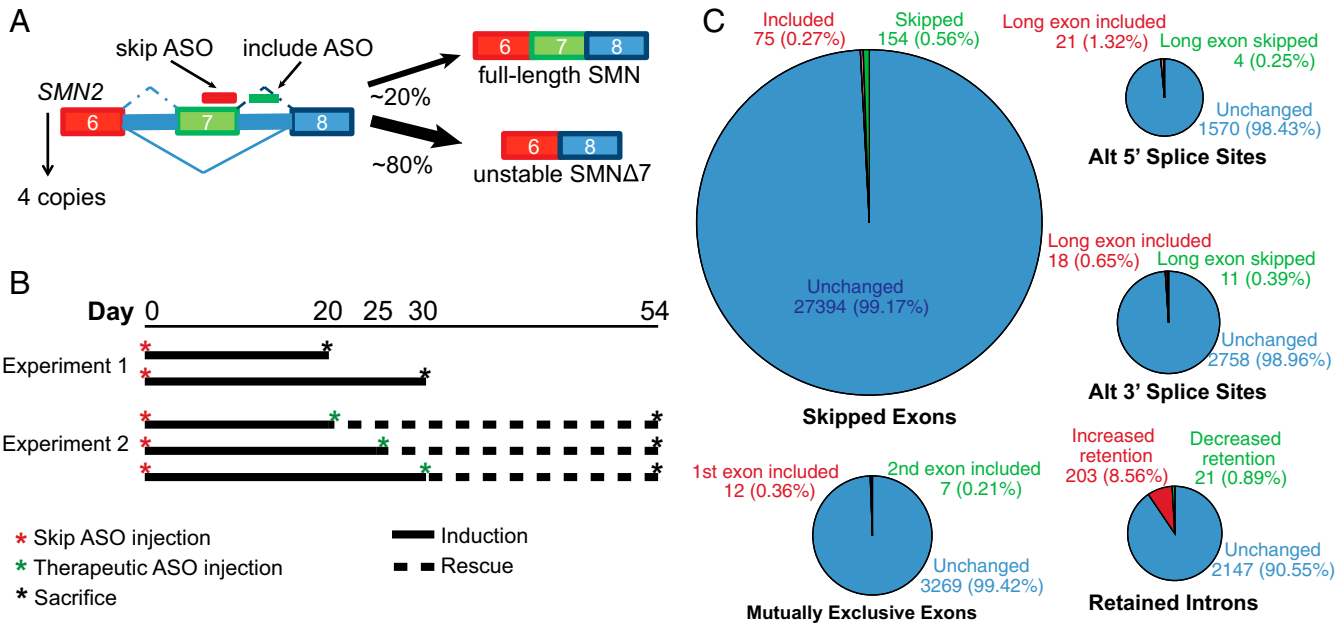

Fig. 1. Splicing changes in inducible SMA mouse spinal cords. (*A*) Schematic of ASO targeting of SMN2 locus. Administration of the skip ASO (red) leads to a shift to ~5% full-length isoform production and ~95% unstable isoform production, whereas additional administration of the include ASO (green) reverts the ratio to >20% full-length isoforms and <80% unstable isoforms. (*B*) Schematic of mouse cohorts with timing of ASO administration and sacrifice. (*C*) Fraction of each category of Gencode-annotated splicing events showing increased or decreased alternative isoform use across each cohort of mice at 30 d post-SMN depletion. FDR <0.05.

SMA. We previously described a mouse model of SMA in which severe SMA is induced in adult *Smn*$^{-/-}$ mice expressing four copies of human *SMN2* by intracerebroventricular administration of an ASO mediating the skipping of exon 7 from the human *SMN2* transgene (6, 7). In the model, 8-wk-old mice were treated with a control oligonucleotide or the disease-inducing ASO ("control" or "skip" ASO) for 20 or 30 d. A second cohort of 8-wk-old mice were injected with the control or skip ASO, then injected with a therapeutic ASO to reverse exon 7 skipping at day 20, 25, or 30 and killed at day 54 (Fig. 1 *A* and *B*). RNA was extracted from whole spinal cord from each cohort and analyzed by exon array as in our previous study (6).

To obtain higher-resolution splicing and gene expression data in this model, we generated poly(A)-mRNA libraries from the same spinal cord samples and sequenced 50-bp paired-end reads to a depth of ∼40 M fragments per sample. The direction of changes in gene expression was concordant between the RNAseq and exon arrays, although RNAseq provided a wider dynamic range of expression changes (Fig. S1*A*). We analyzed transcriptome-wide splicing changes between control and skip ASO at days 20 and 30 postinjection using the rMATS algorithm (29) and Gencode version M4 transcript annotations. Of the splicing categories measured, the largest number of changes occurred in skipped exons (229 significant changes at day 30), followed by retained introns (224 changes); however, the proportion of detectable alternative events that was significantly altered in each class was much greater for retained introns than for skipped exons (9.45% vs. 0.83%) (Fig. 1*C*).

Pervasive Retention of Major- and Minor-Class Introns. Owing to the large proportion of introns showing significant retention at day 30, we next focused on characterizing this subset of splicing events. We first quantified intron retention transcriptome-wide, not limited to annotated retained introns, for major- and minor-class introns. Novel annotations for rMATS were constructed from Gencode vM4 transcript annotations, allowing the possible retention of all unique annotated introns. Annotations of mouse U12 introns were generated similarly from the U12DB (30). Using these annotations, we found that nearly 4% of all introns were significantly retained at 30 d after injection of the skip ASO (Fig. 2*A*). Global intron retention was also confirmed using a second splicing algorithm, MISO (31), with 1,935 introns (1.2% of detected introns) showing significantly increased retention on SMN depletion at day 30 and 211 introns showing increased splicing. Along with driving the assembly of the major spliceosomal snRNPs, SMN also has been implicated in assembling the snRNPs that compose the minor spliceosome, including U11 and U12 (19). Nearly one-third of all U12 introns in the spinal cord of day 30 SMA mice showed significantly increased levels of retention, consistent with a role for SMN in regulating the minor spliceosome (Fig. 2*A*). Importantly, intron retention was reversed or prevented by delivery of the therapeutic ASO, with earlier administration resulting in a greater degree of proper splicing (Fig. 2*B*).

We also performed a parallel induction experiment to understand the time course of splicing and gene expression changes using 15-wk-old mice that were injected with the control or skip oligo and aged for 10, 20, or 30 d. Spinal cords were then harvested and processed for RNA analysis as described previously (6). As expected, the levels of intron retention increased between 10 d and 30 d postinduction (Fig. S2 *A* and *B*). Taken together with the findings from the rescue experiment, these data demonstrate that intron removal is directly correlated with SMN expression.

To understand whether intron retention is a primary effect of SMN loss, we analyzed various characteristics of introns impacted by SMN depletion. We hypothesized that direct mRNA targets arising from splicing dysregulation would be less likely to belong to specific biological functions and would be structurally similar, whereas indirect targets might fall within particular pathways and impact gene expression as a result of activation of a developmental process or stress response (32–35). Analysis of splice site strength showed that introns with greater degrees of retention had weaker 5′ and 3′ splice sites, as determined by consensus to canonical 5′ and 3′ splice site sequences (36), and retained introns were significantly higher in GC content compared with expression- and length-matched control introns (Fig. 2 *C* and *D*).

Although genes containing significantly retained introns were on average more highly expressed than control genes (Fig. S3*A*), they were not enriched for any gene ontology terms and showed similar fold changes in gene expression between control and skip ASO conditions at day 30 as seen in genes containing control

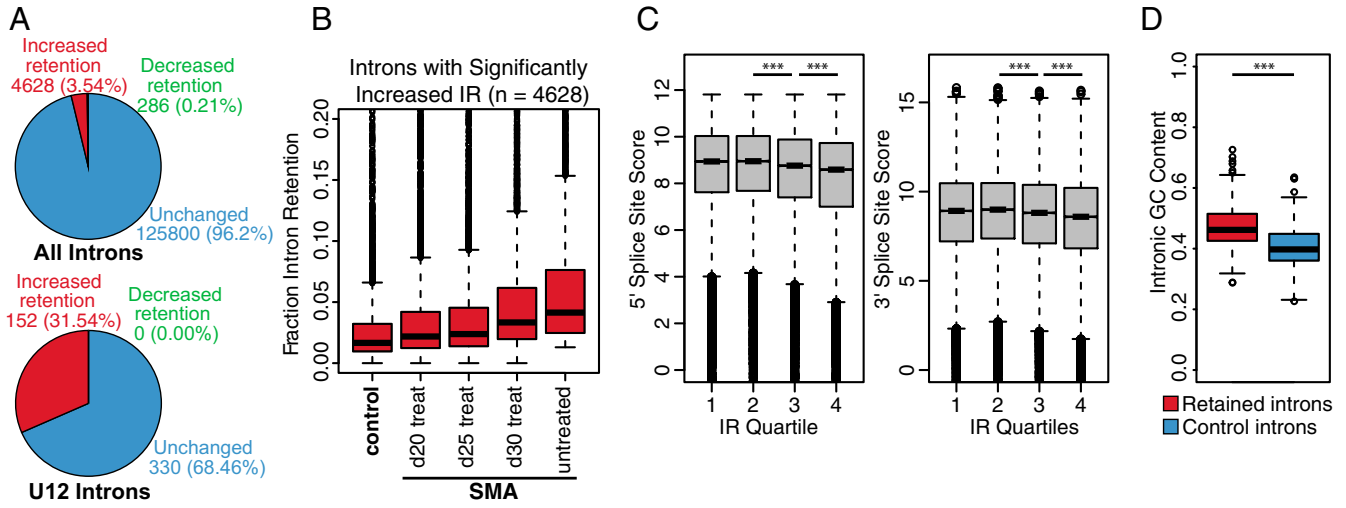

Fig. 2. Increased retention of major and minor class introns on SMN depletion. (*A*) Percentage of all introns and U12 (minor) introns showing increased, decreased, or unchanged levels of intron retention at 30 d. (*B*) Median fraction intron retention for each cohort of mice for 4,628 introns showing significantly increased retention at 30 d postinduction. All pairwise comparisons are significant at $P < 0.05$ by the paired Wilcoxon rank-sum test. (*C*) Median 5′ (*Left*) and 3′ (*Right*) splice site scores for all introns binned into quartiles by change in intron retention on SMN depletion at day 30. Quartile 1 has the smallest or negative change in intron retention; quartile 4 has largest increase in intron retention. ***$P <$ 2.2e-16, Wilcoxon rank-sum test. (*D*) Median intronic GC content for SMN-regulated or length- and expression-matched control introns. ***$P <$ 2.2e-16, Wilcoxon rank-sum test. IR, intron retention.

introns (Fig. S3*B*). This finding suggests that intron retention on SMN depletion does not directly function as a coordinated gene expression response. Furthermore, whereas intron retention is frequently a characteristic of last introns owing to the kinetics of splicing and polyadenylation (35), the distribution of SMN-dependent retained introns was not biased toward any particular location within transcripts (Fig. S3*C*). Taken together, these results suggest that intron retention arising from SMN depletion is likely to be a direct consequence of disrupted spliceosome function rather than a result of secondary activation of a gene regulatory pathway in response to SMN loss.

p53 Activation Occurs Early After SMN Depletion. Gene ontology analysis using Ingenuity Pathway Analysis (Qiagen; qiagenbioinformatics.com) of genes with significant expression changes at 20 d postinjection revealed significant activation of the p53 pathway (Fig. 3*A*). The p53 signature was driven in part by a 5.71-fold increase in expression of Cdkn1a, or p21, a cyclin-dependent kinase inhibitor that mediates cell cycle inhibition. This is in agreement with the Cdkn1a up-regulation previously found in microarray analysis of spinal cord RNA from this model (6) (Fig. S1*B*). This signature persisted to day 30 postinduction (Fig. 3*B*) and was accompanied by up-regulation of early apoptotic markers, including Fas and Pidd1 (Fig. S4*A*), suggesting that p53 induction was preceding, and likely triggering, the activation of apoptosis. In addition to the p53 response, components of the complement cascade, such as C1Qa/b/c, C3AR1, and C5AR1, were up-regulated at days 20 and 30, as reflected in the significant enrichment of the

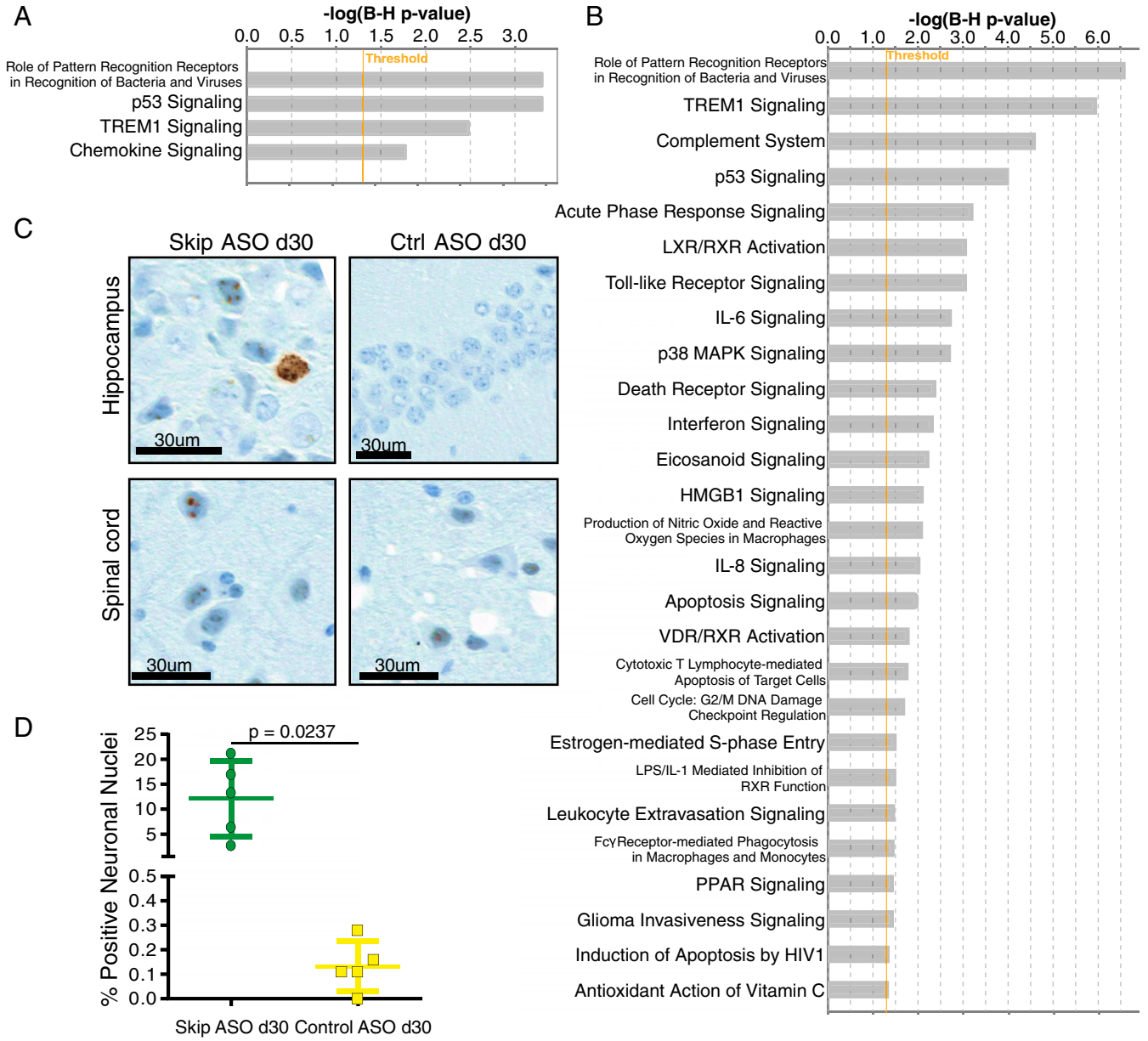

Fig. 3. p53 pathway activation and immune clearance signals over time. (*A* and *B*) Enriched gene ontology terms at day 20 (*A*) and day 30 (*B*). Benjamini–Hochberg corrected *P* value < 0.05; activation/repression absolute *z*-score >0.5. (*C*) Brain and spinal cord of mice treated with skip or control (Ctrl) ASO were harvested at day 30 and fixed and sectioned for immunohistochemistry analysis. Sections were stained with an antibody against γ-H2A.X, a marker for DNA double-strand breaks. Punctate staining in skip day 30 samples represents positive γ-H2A.X loci. (*D*) Quantification of % pyramidal neurons in the hippocampus staining positive for γ-H2A.X.

"pattern recognition receptors" ontology class. Although activation of complement was previously reported in motor neurons of presymptomatic severe SMA mice and was attributed to aberrant synaptic pruning (20), our results do not preclude a contribution of complement secreted from glial cells within the spinal cord, along with the possibility of endogenous neuronal complement causing overpruning.

DNA Double-Strand Breaks in Spinal Cord and Brain of SMA Mice. We next sought to address the nature of p53 activation by assessing the various pathways feeding into and resulting from the p53 response. Intracellular and extracellular stresses, such as hypoxia, oncogenic stress, and DNA damage, can lead to phosphorylation of p53 and up-regulation of its downstream transcriptional targets. Depending on the nature of the insult and degree of cellular damage, this can result in cell cycle arrest or apoptosis. We first looked for markers of DNA damage, cell cycle arrest, and apoptosis in the CNS of SMA mice at 30 d after injection of ASO using immunohistochemistry. Surprisingly, we found evidence for DNA double-strand breaks as measured by γ-H2A.X foci throughout the hippocampus and in interneurons in the spinal cord (Fig. 3*C*). Within the hippocampus, 12.11% of pyramidal neurons showed positive γ-H2A.X staining in SMN ASO-treated mice at day 30 postinjection, whereas only 0.13% of pyramidal neurons stained positive in control ASO-treated mice (Fig. 3*D*). Cyclin D1, the activation of which is required for progression through the G1/S phase transition of the cell cycle and has been postulated to be reactivated in postmitotic neurons in response to DNA damage (37), showed a similar staining pattern in the brain and spinal cord (Fig. S4*B*).

Despite this DNA damage and the up-regulation of apoptotic markers at the RNA level, we did not detect any evidence of apoptosis in either of these tissues, with virtually no cells in brain or spinal cord staining positive for the apoptotic marker cleaved caspase-3. This suggests that at day 30, neurons in the brain and spinal cord had accumulated DNA damage but were not actively undergoing apoptosis.

Human Cell Culture Models of Acute SMN Depletion Recapitulate Splicing Defects and DNA Damage. To determine whether the intron retention and DNA double-strand breaks seen in our SMA mouse model can be translated to humans, we established two human cell culture models to assess the effects of acute SMN loss. We stably infected SH-SY5Y human neuroblastoma cells and human induced pluripotent stem cell (hiPSC)-derived motor neurons (hiPSC-MNs) with lentiviral vectors expressing short-hairpin RNA (shRNA) targeting SMN1/2 under the control of a doxycycline-inducible promoter. After 7 and 5 d of shRNA induction, respectively, both SH-SY5Y cells and hiPSC-MNs showed appreciable depletion of SMN protein compared with a nontargeting control shRNA (Fig. 4*A* and Fig. S5*A*). SH-SY5Y cells also showed an accumulation in the G0/G1 phase of the cell cycle that correlated with activation of p21 (Fig. S4*C*).

We next performed mRNAseq on these samples in duplicate and analyzed gene expression and splicing changes between control and SMN knockdown conditions in both cell lines. Pathway analysis by Ingenuity Pathway Analysis on significantly changing genes revealed that for SH-SY5Y and hiPSC-MNs, transient knockdown of SMN resulted in activation of a DNA double-strand break response, indicated by activation of targets of BRCA1, a key player in double-strand break repair (Fig. 4*B* and Fig. S5*B*). Interestingly, we failed to observe the up-regulation of complement components that had been detected in the SMA mouse model. This suggests that complement activation is a non–cell-autonomous effect of the interaction between neuronal cells and glia and might not be intrinsic to motor neurons, and that it may be an indirect effect of SMN depletion.

We next quantified splicing changes, specifically global intron retention, in SH-SY5Y cells and hiPSC-MNs depleted of SMN. Splicing and gene expression changes were of smaller magnitude

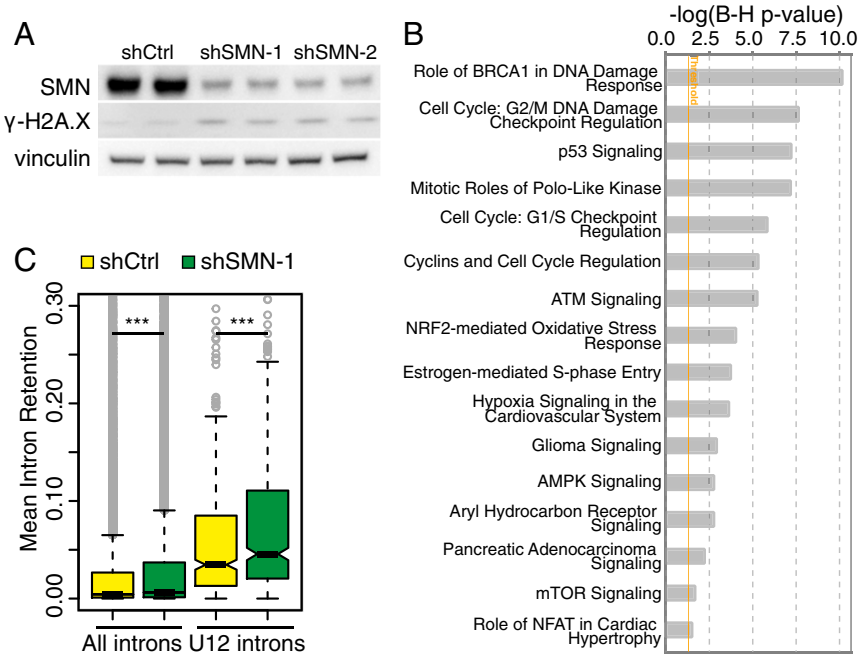

Fig. 4. DNA damage and intron retention in the human neuroblastoma cell culture model. SHSY5Y cells were lentivirally transduced with a doxycycline-inducible shRNA targeting SMN1/2. RNA and protein were harvested at 7 d after shRNA induction with 1 μg/mL doxycycline. (*A*) SMN knockdown leads to activation of γ-H2A.X in SHSY5Y cells. (*B*) Enriched gene ontology terms in SHSY5Y cells after SMN knockdown shows induction of DNA damage response pathways. Benjamini–Hochberg corrected *P* value < 0.05; activation/repression absolute z-score >0.5. (*C*) Median retention of all introns and U12 introns is significantly higher on SMN knockdown. ***$P < 0.001$. Ctrl, control.

in the iPSC-MNs than in the SH-SY5Y cells, perhaps because of the shorter time point of SMN depletion and the lack of cell division in the iPSC-MNs. Despite this, consistent with our observations in the SMA mouse model, all introns, as well as the U12 subset of introns, demonstrated significantly increased average retention on SMN knockdown in both cell lines (Fig. 4*C* and Fig. S5*C*). Furthermore, in support of a direct effect of SMN loss on minor intron retention arising from defects in snRNP assembly, knockdown of SMN in SH-SY5Y cells resulted in a significant decrease in minor spliceosomal snRNA U11 (Fig. S6).

Of note, only eight of the significantly retained introns in mouse spinal cord and human SH-SY5Y cells were shared between both datasets, implying that the conserved DNA damage effect is not dependent on the identity of the retained introns, but rather that the global stress of widespread intron retention may be the driving force in this signature. Considered together, these results suggest that preventing DNA damage by promoting proper splicing is likely to be a conserved function of SMN that is disrupted in SMA.

R-Loop Induction on SMN Depletion. Several recent studies have proposed that defects in mRNA processing can contribute to genome instability (14, 28, 38). One proposed mechanism for this effect is through the failure to resolve transcriptional R-loops, transient RNA:DNA hybrid structures that form during transcription that displace the nontemplate DNA strand and generate susceptibility to DNA damage if stabilized. R-loop formation is limited by the rapid association of RNA-binding proteins with the nascent RNA during transcription and splicing, preventing hybridization to the DNA template, and is actively resolved by RNA:DNA helicase activity and RNase H1/H2 activity (39, 40). The presence of G-rich nontemplate strands can stabilize R-loops by folding into G-quadruplexes and further contribute to double-strand break formation (39, 41). Knockdown of RNA splicing and transport factors has been associated with increased R-loop formation and DNA damage (42, 43). Similarly, we found that splicing inhibition by the small molecule pladienolide B caused accumulation of γ-H2A.X in a time-dependent manner (Fig. S7*A*). Based on this finding, we hypothesized that intron retention arising from SMN depletion may lead to DNA damage through the formation of R-loop intermediates at retained intron loci, which we had found to have increased GC content.

To test this model, we assayed R-loop formation genome-wide in SH-SY5Y cells expressing the control or SMN shRNA using DNA-RNA immunoprecipitation (IP) sequencing (DRIPseq), which enriches for hybrid structures in the chromatin using the R-loop–specific antibody S9.6 (44). We also performed DRIPseq on samples immunoprecipitated with mouse IgG as a control for specific enrichment. Normalizing to this negative control, we identified 6,315 regions as significantly enriched across two replicates in nontargeting control (shCtrl) cells, representing the basal level of R-loop formation in SH-SY5Y cells. In shSMN1 cells, 10,163 regions were significantly enriched in both replicates, showing that loss of SMN resulted in a 1.6-fold increase in global R-loop formation (Fig. 5*A*). Comparing the distribution of R-loops across genomic regions, we found that R-loops were significantly enriched in gene-associated regions in SMN knockdown cells compared with control knockdown cells ($P <$ 2.2e-16, χ^2 test) (Fig. S7*B*).

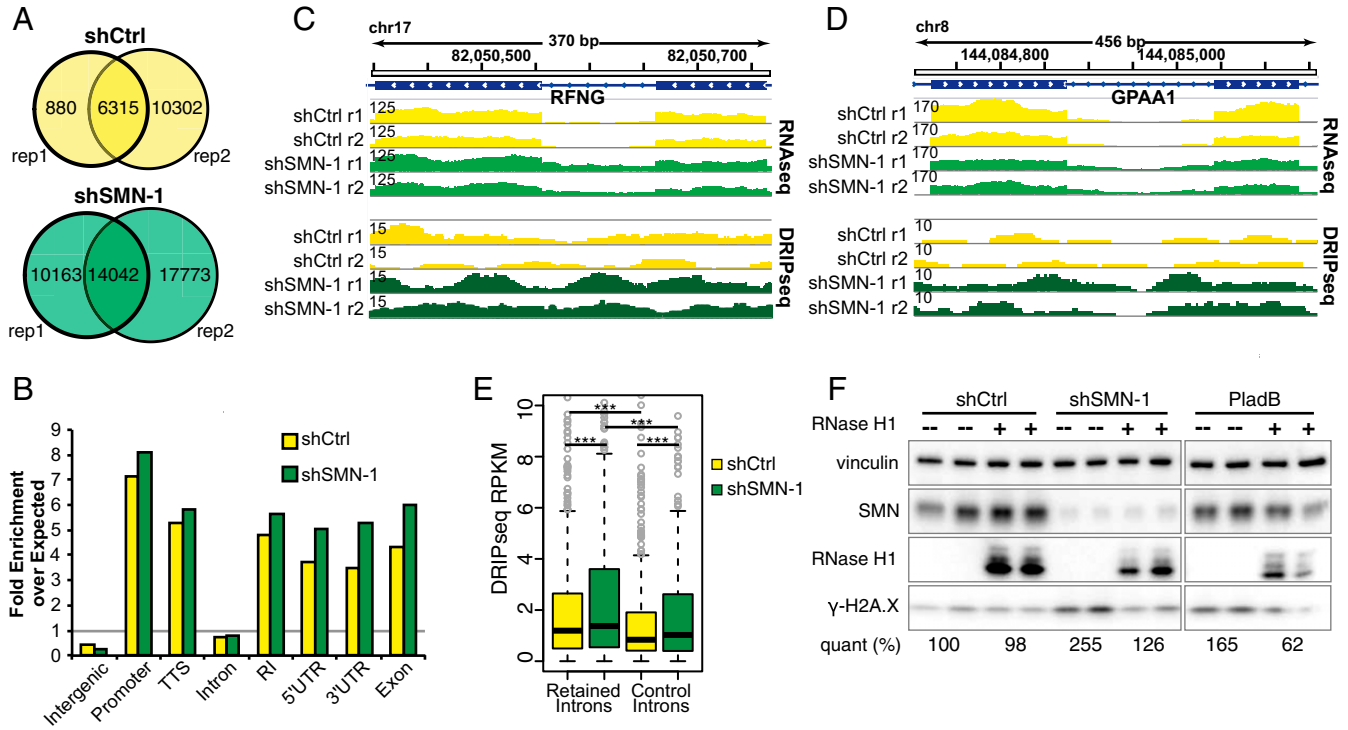

Fig. 5. Increased R-loop induction on SMN depletion. (*A*) Venn diagrams of shCtrl and shSMN1 replicate S9.6 IPs enriched over IgG IPs. Intersection represents peak summits overlapping within a 250-bp window. (*B*) Fold enrichment of peak distribution across genomic regions using overlapping peaks determined from *A*, normalized to the genomic distribution of a randomly shuffled set of peaks. RI, retained intron; TTS, transcription termination site; UTR, untranslated region. (*C* and *D*) IGV browser images of a major-class intron in the RFNG gene (*C*) and a minor-class intron in the GPAA1 gene (*D*) show increased accumulation of RNAseq and DRIPseq reads in SMN knockdown cells compared with controls. (*E*) Median coverage of normalized R-loop DRIPseq reads across introns with increased retention on SMN depletion at day 30 (*Left*) or a length- and expression-matched control set of introns (*Right*). ***$P < 0.001$. (*F*) Representative Western blot of SH-SY5Y cells induced to express control or SMN1 shRNA for 4 d, then transfected with empty vector or Myc-tagged RNase H1 for 48 h. PladB, SH-SY5Y cells exposed to 1 μM pladienolide B for 4 h before harvesting. Average normalized γ-H2A.X quantification across replicates is shown below the blot.

We next normalized DRIPseq peak distributions to a set of randomly distributed peak regions. Although the DRIPseq peaks were depleted across all introns, in retained introns they were 4.8-fold enriched over the random distribution in shCtrl cells and 5.6-fold enriched in shSMN1 cells, indicating that R-loops are specifically enriched over retained introns (Fig. 5*B*). Because this difference in enriched peaks occurring in retained introns between control and SMN knockdown cells is not statistically significant, this perhaps reflects a high basal level of R-loop formation in these regions. However, examination of individual examples of major and minor introns with significant retention revealed accumulation of DRIPseq reads in SMN knockdown cells compared with control cells (Fig. 5 *C* and *D*). Furthermore, when we specifically measured normalized DRIPseq read coverage across regions of interest, we detected significantly higher DRIPseq read coverage across introns that were retained on SMN depletion compared with a control set of length- and expression-matched introns, suggesting that intron retention does indeed correlate with increased propensity to form R-loops, and that SMN depletion leads to increased R-loop formation (Fig. 5*E*).

It has been observed that R-loops are involved in transcript termination (45), and a recent study proposed that SMN recruitment of the RNA:DNA helicase senataxin may enhance R-loop resolution at transcription termination sites (28). Consistent with these findings, a significantly larger fraction of peaks in SMN knockdown cells were located within a region 1 kb downstream of genes compared with peaks in control cells (19.8% vs. 17.9%; $P = 0.002$) (Fig. S7*B*). Specific enrichment flanking transcription termination sites on SMN depletion was also verified through metagene analysis (Fig. S7*C*). Motif analysis of enriched DRIPseq regions in shSMN cells revealed T-rich motifs, which are found near transcript termination sites (46). We also detected a G-rich motif consistent with G-quadruplex–forming regions, suggesting that these regions are in fact more likely to form stable R-loops through the formation of G-quadruplexes (Fig. S7*D*).

We next asked whether the DNA damage signature detectable in SH-SY5Y cells arises from R-loops. To test this, we sought to rescue DNA double-strand breaks, measured by H2A.X phosphorylation, by repairing R-loops with RNase H1 overexpression. RNase H1 has been demonstrated to be sufficient to repair R-loops in cell culture (42). We overexpressed tagged human RNase H1 in control or SMN knockdown cells or in cells treated with the splicing inhibitor pladienolide B and assayed γ-H2A.X levels. Indeed, we found that RNase H1 overexpression partially suppressed the increased γ-H2A.X signal seen in SMN knockdown cells, suggesting at least part of the DNA damage signal arises from the R-loops (Fig. 5*F*). Acute treatment with pladienolide B also resulted in activation of γ-H2A.X, and this was prevented or suppressed by RNase H1 overexpression. These results support our hypothesis that decreases in splicing efficiency, such as those caused by SMN deficiency or direct splicing inhibition, may cause genome instability and activation of a DNA damage response through R-loop formation.

Discussion

Here we have characterized a prevalent intron retention and DNA damage signature in an inducible mouse model of SMA, as well as in human SMA cell culture systems. As an alternative to the paradigm in the field attributing motor neuron deficits in type I SMA to alternative splicing changes in specific target transcripts, we present a unifying model in which inefficiencies in RNA processing, manifested primarily as intron retention, lead to DNA damage through the formation of RNA:DNA hybrids (Fig. 6). Although we cannot formally rule out the possibility that the splicing changes that we observed are indirect effects of SMN loss (23), our computational and experimental analyses suggest that SMN depletion causes systematic defects in intron removal. We found that both U2 and U12 introns were retained in mouse and cell culture models of SMA, with a bias toward retention of U12 introns, and that minor snRNAs showed decreased expression on SMN depletion. This bias may be explained by the observation that snRNAs of the minor spliceosome are two orders of magnitude less abundant than snRNAs of the major spliceosome (47) and may be disproportionately affected by SMN depletion because of the stoichiometry of major and minor spliceosome components. Introns preferentially retained on SMN depletion in SH-SY5Y cells were enriched in R-loops compared with spliced introns. RNA and protein signatures of DNA double-strand breaks were also detectable in SMA mice and in the human cell lines, and overexpression of RNase H1, which resolves R-loops, partially rescued DNA breaks in SH-SY5Y cells depleted of SMN. Based on this evidence, we propose that DNA damage in type I SMA is triggered by R-loop formation resulting from defects in intron removal. The extent of intron retention and DNA damage is likely correlated with SMN expression levels, and as a result these observations may be relevant in intermediate models of type II or type III SMA.

Several previous studies have focused on individual splicing events as drivers of SMA pathogenesis in motor neurons (19, 20, 48). Although it is likely that the cohort of tissue-specific alternative splicing changes in SMA contributes to the sensitivity of motor neurons to cell death (19), we propose that the phenomenon of global intron retention is a distinct cause of cellular stress proximal to SMN loss. We observed that intron retention was rescued by the therapeutic ASO, indicating that these splicing changes are directly dependent on SMN expression levels. Unlike retained introns that trigger NMD on export of the transcript to the cytoplasm (34), SMN-dependent retained introns in our mouse model did not have a significant effect on expression of the genes containing them. This finding suggests that they are ultimately spliced out in the nucleus, perhaps more slowly than other introns owing to their weaker splicing signals, yet their reduced splicing kinetics in the absence of SMN may be sufficient to trigger cellular stress. There was also minimal overlap between the set of retained introns in the mouse model and the human cell culture models despite striking concordance in DNA damage response gene expression signatures. These observations argue for a more critical role in the process of stochastic intron retention than in the specific identity of retained introns. Interestingly, Ackerman et al. (49) recently demonstrated that mutation of one of the multicopy U2 snRNA genes resulted in retention of short introns across broad gene ontologies and subsequent neurodegeneration, further implicating intron retention in neuronal stress.

Numerous hypotheses have been put forth to explain the activation of the DNA damage response in neurodegeneration. In some cases, mutations in DNA damage response and repair

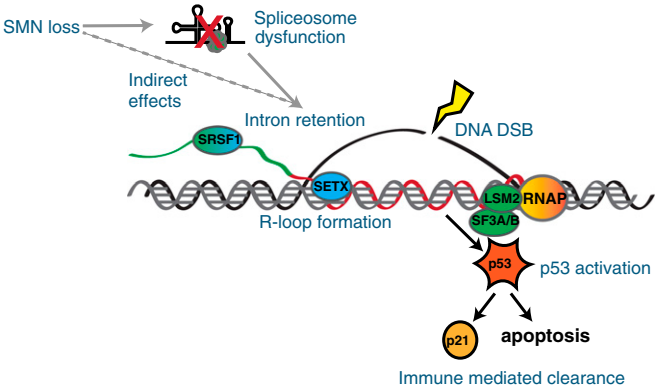

Fig. 6. Intron retention as a mechanism for DNA damage induction. Model for splicing dysregulation as a modulator of R-loop formation and DNA damage. Depletion of proteins in green induces DNA double-strand breaks, and depletion of proteins in blue induces R-loop formation.

components themselves are the causative factors in DNA damage accumulation (50). For example, the neurologic disorder ataxia-telangiectasia is caused by mutations in *ATM,* the serine-threonine kinase responsible for coordinating the double-strand break response (51). Similarly, mutations in the single-strand break repair gene aprataxin (*APTX*) cause one of the most common forms of spinocerebellar ataxia, ataxia with oculomotor apraxia 1 (52, 53). Notably, despite being germline mutations in ubiquitous DNA damage response genes, both diseases display primarily neurologic deficits, suggesting a particular sensitivity of the nervous system to defects in the DNA damage response.

Other sources of DNA damage, particularly in age-related neurodegeneration, can be attributed to oxidative stress. The high rate of oxygen consumption by the brain results in increased production of reactive oxygen species with age, and oxidative DNA damage is emerging as a hallmark of both Alzheimer's disease and Parkinson's disease (54, 55). Induction of oxidative stress results in activation of p38 MAPK and initiates an oxidative stress transcriptional response. We failed to detect activation of p38 in our human cell culture models of SMA and detected only a subtle transcriptional signature of p38 activation in the mouse spinal cord at day 30, suggesting that oxidative stress is not a prominent player in these systems and is unlikely to be the cause of DNA damage.

More recently, evidence has been accumulating for the involvement of cotranscriptional R-loops and DNA damage in several neurodegenerative diseases, perhaps most strikingly in motor neuron diseases. R-loops form normally in many circumstances, including Ig class switch recombination and transcription termination, and are actively resolved by RNA:DNA helicases and RNase H1 and H2 (39, 56). Persistent R-loops can become deleterious, however. These aberrant R-loops are generally stabilized by the formation of G-quadruplexes in the displaced, G-rich nontemplate strand, which can trigger gene down-regulation or DNA double-strand breaks through as-yet unclear mechanisms (39, 41). To date, two distinct mechanisms for R-loop formation in neurodegeneration have been proposed, depending on the genetic insult. In the case of repeat expansion diseases, such as the *c9orf72* G_4C_2 hexanucleotide repeat that is the most common genetic lesion in both sporadic and familial amyotrophic lateral sclerosis (ALS), the repeat itself forms R-loops when transcribed, causing abortive transcription at the *c9orf72* locus (57). The GAA repeats in *frataxin* that cause Friedreich ataxia have also been shown to form R-loops in vitro, and it is likely that other G-rich repeat expansions also cause gene silencing and potentially form focal DNA breaks in an R-loop–dependent manner (58).

The other proposed mechanism is widespread aberrant R-loop formation owing to mutations in components of the R-loop formation or resolution pathways. It was recently shown that SMN recruits senataxin, an RNA:DNA helicase involved in R-loop resolution, to transcription termination sites (28). Intriguingly, senataxin itself is mutated in a juvenile-onset form of ALS, designated ALS4, that bears a striking resemblance to SMA and results in loss of motor neurons in early childhood (59). An additional RNA:DNA helicase, IGHMBP2, was initially characterized in the context of R-loop–dependent class-switch recombination, but it has since been found that loss-of-function mutations in IGHMBP2 cause spinal muscular atrophy type I with respiratory distress (60, 61). These two independent lines of genetic evidence, in addition to our results, provide strong support for the causative role of unresolved R-loops in motor neuron disease.

The conspicuous enrichment of mutations in RNA-binding proteins in neuromuscular disease begs the question of whether there is a common defect in RNA metabolism across this spectrum of illnesses. Mutations in the RNA-binding protein genes *TDP43*, *FUS*, and *ATXN2* have been associated with sporadic and familial forms of ALS. These proteins shuttle between the nucleus and cytoplasm and have multiple roles in RNA metabolism, including transcription, splicing, and mRNA transport. In addition, all of these factors, as well as SMN, have been found in cytoplasmic stress granules, dynamic sites of translational repression that form rapidly during cellular stress (62, 63). A hallmark of ALS pathology is the appearance of cytoplasmic ubiquitinated aggregates of TDP-43 in the brain and spinal cord. These aggregates occasionally contain FUS protein as well, and *FUS* mutations can independently lead to FUS aggregate formation. The identity of these aggregates, and whether they are distinct from stress granules, remain to be determined (64). In addition, mutant FUS sequesters SMN in the cytoplasm and causes altered snRNA levels and splicing changes, some of which are consistent with SMN loss on its own (65). SMN itself has been shown to induce stress granule assembly on overexpression (66), and Ataxin2 also interacts with several RNA-binding proteins, including FUS and TDP-43, and localizes to stress granules (67).

One potential model for a shared mechanism across these genetic defects is that the induction of aggregates or stress granules sequesters these shuttling proteins away from their nuclear functions. Consequently, defects in splicing, RNA processing, and transport would be expected. Perhaps consistent with this idea, FUS and TDP-43 were recently implicated in the prevention of R-loop–mediated transcription-associated DNA damage in models of ALS (68). As we have demonstrated here, changes in RNA processing due to either to cytoplasmic aggregation or genetic deletion of splicing factors could initiate the formation of DNA double-strand breaks in part through R-loops, leading ultimately to the degeneration of motor neurons in both ALS and SMA.

Several open questions stemming from our observations and findings of others remain to be addressed. One outstanding issue is the specificity of the pathology to motor neurons in the face of mutations in ubiquitously expressed genes. Although the sensitivity of this cell type to genomic instability owing to inefficient repair machinery is a tantalizing explanation, it is possible that additional functions of SMN contribute to the motor neuron specificity. This possibility has been suggested by SMA models of SMN point mutations that recapitulate motor neuron degeneration in the absence of defects in snRNP assembly (69). Alternatively, there likely is less prominent pathology in other cell types (70).

Second, the relative contributions of intron retention, transcription termination defects, and other RNA processing defects to R-loop formation in SMA remain unclear. SMN may directly mediate R-loop resolution through the recruitment of senataxin, as demonstrated previously (28). Alternatively, we present a model wherein G-rich retained introns arising indirectly from SMN loss are putative substrates for R-loop formation. Further study is needed to elucidate whether intron retention synergizes with other defects in RNA processing to initiate DNA damage when SMN is deficient. Most importantly, it will be necessary to further dissect whether manipulation of R-loop formation and subsequent DNA damage can be a viable therapeutic option in SMA. Targeting this pathway in combination with SMN restoration may provide additional benefit to patients beyond the benefit from SMN restoration alone. If indeed SMA pathogenesis and ALS pathogenesis share this mode of DNA damage induction and neurodegeneration, then such therapeutic modalities would have broad implications for motor neuron disease.

Materials and Methods

Western Blot Analysis. Protein from SH-SY5Y cells was extracted with RIPA buffer containing protease and phosphatase inhibitors, resolved on 4–12% Bis-Tris gels (Novex), and probed with the following antibodies: mouse anti-vinculin (clone VIN-11–5; Sigma-Aldrich, V4505), mouse anti-SMN (BD Biosciences, 610647), rabbit anti–γ-H2A.X (Abcam, ab11174), mouse anti-RNase H1 (Abcam, ab56560), rabbit anti-p21 (Cell Signaling Technology, 2947), and rabbit anti-p53 (Cell Signaling Technology, 9282).

Immunohistochemistry. Spinal cord and brain from control and SMA mice were fixed with 4% (vol/vol) formaldehyde in PBS overnight. Fixed tissue was soaked in 30% (wt/vol) sucrose overnight at 4 °C and embedded in frozen

medium. The Ventana XT slide staining system was used to stain 10-μm sections. For detection of γ-H2A.X, rabbit anti–γ-H2A.X (S139) (R&D Systems, AF2288) was used at 0.25 μg/mL. Secondary staining was done using OmniMap anti-Rb HRP (Roche, 760–4311).

Lentiviral Transduction. SMARTchoice lentiviral particles of doxycycline-inducible shRNAs targeting SMN1 [catalog no. VSH6376: V2IHSPGG_743120 (shSMN1), V2IHSPGG_743128 (shSMN2)] or a nontargeting control [VSC6580 (shCtrl)] under the PGK promoter and constitutively expressing GFP were purchased from Dharmacon. SH-SY5Y cells were incubated in concentrated viral supernatant with 4 μg/mL polybrene (EMD Millipore, TR-1003-G) for 5 h, then viral supernatant was diluted fourfold in fresh Complete Medium. At 24 h postinfection, medium was changed to fresh medium containing 1 μg/mL puromycin. Selection was continued for 5 d. Cells were expanded and frozen as pools of infected cells. For hiPSC-MNs, cells were differentiated in a 12-well format, then infected overnight without polybrene. The medium was changed daily for 3 d without puromycin selection. For both SH-SY5Y and hiPSC-MNs, fluorescence microscopy indicated nearly 100% infection efficiency as assessed by GFP expression.

mRNA Sequencing Sample Isolation and Preparation. Total RNA from spinal cord, lentivirally infected SH-SY5Y cells induced for 7 d with doxycycline, or lentivirally infected hiPSC-MN cells induced for 5 d with doxycycline was extracted using TRIzol (Thermo Fisher Scientific, 15596026) and the Direct-zol RNA MiniPrep Kit (Zymo Research, R2050). Then 50–500 ng of DNase-treated total RNA, with a Universal Human Reference RNA as control, were used to construct libraries using the TruSeq Stranded mRNA HT Sample Prep Kit (Illumina, RS-122-2103) with barcodes for sample multiplexing. The products were purified and enriched with PCR for 13 cycles to create final cDNA libraries. Libraries were quantified using the HT DNA Reagent Kit (Caliper Life Sciences, 760435) on a LabChip GX automated gel electrophoresis system (PerkinElmer). Multiplexed libraries were equimolarly pooled, diluted to a 2 nM pool for final analysis on Agilent High-Sensitivity DNA Kit to run on one rapid flow cell per lane, and sequenced using 2 × 50-nucleotide paired-end runs of the Illumina HiSeq 2500 platform.

Reverse-Transcription and qPCR Measurement of snRNAs. Total RNA was extracted from lentivirally infected SH-SY5Y cells induced for 7 d with doxycycline and DNase treated using the Directzol RNA MiniPrep Plus Kit (Zymo Research). snRNAs were quantified as in Zhang et al. (19). In brief, cDNAs were synthesized from each sample using gene-specific reverse primers and 100 ng of total RNA (375 ng of total RNA for minor snRNAs). Real-time qPCR was performed for each snRNA and normalized to 5S rRNA using SYBR Green chemistry on an Applied Biosystem QuantStudio 12K Flex Real-Time PCR system. Triplicate total RNA samples were harvested from each knockdown sample and propagated through cDNA preparation and qPCR assays. Details of primer sets are presented in Table S1.

Transfections. SH-SY5Y cells expressing lentiviral shRNA against control or SMN cells were induced for 4 d with 1 μg/mL doxycycline. Cells were transfected with pCMV6-Entry (OriGene, PS100001) or FLAG-Myc–tagged human RNase H1 (OriGene, RC200595) using Lipofectamine 3000 (Thermo Fisher Scientific, L3000008). Protein was harvested after 48 h using RIPA buffer and analyzed by Western blot analysis.

DRIPseq. DRIPseq was performed as described (71) with the following changes. In brief, nucleic acids isolated from ∼2 × 10^6 cells were digested with a restriction enzyme mixture (50 units each of BsrGI, XbaI, EcoRI, HindIII, and SspI) overnight at 37 °C in 1× CutSmart Buffer (New England BioLabs). Digests were purified by phenol/chloroform extraction, and 8.8 μg of digested DNA per cell culture condition was treated with 3 μL of RNase H overnight at 37 °C. For each condition, 4.4 μg of RNase H-treated or untreated DNA was immunoprecipitated with 10 μg of S9.6 antibody (Kerafast, ENH001), washed three times with IP buffer, eluted, treated with Proteinase K and phenol/chloroform-extracted. Resulting samples were sonicated using a Bioruptor Pico sonicator (Diagenode), with 11 cycles of 15 s on/90 s off. Fragments were prepared for sequencing using the NEBNext Ultra II Library Prep Kit for Illumina (New England BioLabs) with barcodes for sample multiplexing. Pooled libraries were sequenced on the Illumina NextSeq platform.

Derivation of hiPSCs and Differentiation into Motor Neuron Cultures. Human iPSCs were derived from lymphoblastoid cell lines (Coriell Biorepository) with the Epi5 Episomal iPSC Reprogramming Kit (Life Technologies) and propagated in mTeSR Medium (Stem Cell Technologies) on Matrigel (Life Technologies). The pluripotency of the hiPSCs was confirmed by staining for pluripotency markers and in vitro differentiation into three germ layers. Differentiation of hiPSCs into motor neuron cultures was carried out as described previously (72), with slight modifications. In brief, hiPSC colonies were dissociated with EDTA and placed in suspension in low-adhesion flasks (Corning) to form embryoid bodies in medium consisting of DMEM/F12/ Neurobasal medium (Life Technologies), N2 supplement (Life Technologies), B27 supplement (Life Technologies), ascorbic acid (20 μM), and β-mercaptoethanol (0.055 mM). The medium was changed every other day, and small molecules were added for neuronal patterning [0.2 μM LDN193189, 40 μM SB431542, and 3 μM CHIR99021); motor neuron induction (1 μM retinoic acid, 0.5 μM agonist of hedgehog signaling (SAG), and 10 μM of the drug DAPT]. On day 14, EBs were dissociated into single cells with papain and plated on Matrigel or poly D-lysin/laminin-coated plates in medium containing neurotrophic factors [20 ng/mL brain-derived neurotrophic factor (BDNF), 10 ng/mL glial cell-derived neurotrophic factor (GDNF), and 10 ng/mL ciliary neurotrophic factor (CNTF)].

RNAseq Gene Expression and Splicing Analysis. RNAseq reads were mapped with the STAR algorithm, aligning to the Gencode version M4 transcript annotation for mouse spinal cord RNA and the hg38 annotation for SH-SY5Y cells and hiPSC-MNs. We used parameters–alignIntronMin 20–alignIntronMax 1000000–outFilterMismatchNmax 999–outFilterMismatchNoverLmax 0.05–outFilterMultimapNmax 20–alignSJoverhangMin 8–alignSJDBoverhangMin 1–alignMatesGapMax 1000000–sjdbOverhang 49. Aligned reads were assigned to transcripts from the aforementioned annotations using RSEM, and resulting raw read counts were input into DESeq2 for differential gene expression analysis using default parameters.

For general splicing analysis, BAM files produced from STAR mapping were input into rMATS, using the Gencode version M4 transcript annotation for mouse spinal cord RNA and the hg38 annotation for SH-SY5Y and hiPSC-MN RNA samples. For detection of intron retention, human and mouse annotations were generated containing all consecutive spliced and unspliced exon-intron-exon triads from hg38 and Gencode vM4. To measure the retention of U12 introns, we generated a second set of annotations by filtering this novel set of annotations for introns that overlapped exactly with human and mouse U12 introns extracted from U12DB. These annotations were used to call intron retention in mouse spinal cord, SH-SY5Y, and hiPSC-MN RNAseq datasets using rMATS. A false discovery rate (FDR) cutoff of 0.05 was used to define significantly differentially spliced regions.

Bioinformatic Analyses. Custom Python and R scripts were used to determine intronic GC enrichment, expression of genes containing retained introns, genic distribution of retained introns, and DRIPseq metagene plots. BedTools (73) was used to determine the genomic distribution of DRIPseq regions.

Cell Cycle Analysis. SH-SY5Y cells stably transduced with a nontargeting shRNA or shRNAs targeting SMN1/2 were induced to express shRNA for 7 d with 1 μg/mL doxycycline. Cells were then treated for 4 h with DMSO or 50 μM etoposide. A total of 1 × 10^5 cells per sample were trypsinized, fixed in cold 70% EtOH, and incubated in 0.1% Triton X-100 with 200 μg/mL RNase A and 20 μg/mL propidium iodide. Cells were analyzed by flow cytometry to determine the proportions of cell populations in distinct stages of the cell cycle based on DNA content.

ACKNOWLEDGMENTS. We thank Christopher Henderson for providing helpful guidance and discussion, and staff at Biogen Translational Pathology for conducting the immunohistochemistry studies. This work was supported in part by the St. Giles Foundation and the SMA Foundation (A.R.K.). Biogen and Ionis Pharmaceuticals provided funding for the acquisition, analysis, and interpretation of data. Biogen reviewed and provided feedback on the manuscript to the authors, and all authors provided their approval of the manuscript.

1. Talbot K, Davies KE (2001) Spinal muscular atrophy. *Semin Neurol* 21(2):189–197.
2. Birnkrant DJ, Pope JF, Martin JE, Repucci AH, Eiben RM (1998) Treatment of type I spinal muscular atrophy with noninvasive ventilation and gastrostomy feeding. *Pediatr Neurol* 18(5):407–410.
3. Campbell L, Potter A, Ignatius J, Dubowitz V, Davies K (1997) Genomic variation and gene conversion in spinal muscular atrophy: Implications for disease process and clinical phenotype. *Am J Hum Genet* 61(1):40–50.
4. Lefebvre S, et al. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80(1):155–165.
5. Burnett BG, et al. (2009) Regulation of SMN protein stability. *Mol Cell Biol* 29(5):1107–1115.
6. Staropoli JF, et al. (2015) Rescue of gene-expression changes in an induced mouse model of spinal muscular atrophy by an antisense oligonucleotide that promotes inclusion of SMN2 exon 7. *Genomics* 105(4):220–228.

7. Sahashi K, et al. (2013) Pathological impact of SMN2 mis-splicing in adult SMA mice. *EMBO Mol Med* 5(10):1586–1601.
8. Coovert DD, et al. (1997) The survival motor neuron protein in spinal muscular atrophy. *Hum Mol Genet* 6(8):1205–1214.
9. Lefebvre S, et al. (1997) Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16(3):265–269.
10. Tizzano EF, Cabot C, Baiget M (1998) Cell-specific survival motor neuron gene expression during human development of the central nervous system: Implications for the pathogenesis of spinal muscular atrophy. *Am J Pathol* 153(2):355–361.
11. Wahl MC, Will CL, Lührmann R (2009) The spliceosome: Design principles of a dynamic RNP machine. *Cell* 136(4):701–718.
12. Burghes AH, Beattie CE (2009) Spinal muscular atrophy: Why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci* 10(8):597–609.
13. Pellizzoni L (2007) Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* 8(4):340–345.
14. Li DK, Tisdale S, Lotti F, Pellizzoni L (2014) SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Semin Cell Dev Biol* 32:22–29.
15. Pellizzoni L, Yong J, Dreyfuss G (2002) Essential role for the SMN complex in the specificity of snRNP assembly. *Science* 298(5599):1775–1779.
16. Fallini C, et al. (2011) The survival of motor neuron (SMN) protein interacts with the mRNA-binding protein HuD and regulates localization of poly(A) mRNA in primary motor neuron axons. *J Neurosci* 31(10):3914–3925.
17. Zhang HL, et al. (2003) Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization. *J Neurosci* 23(16):6627–6637.
18. Fallini C, Bassell GJ, Rossoll W (2012) Spinal muscular atrophy: The role of SMN in axonal mRNA regulation. *Brain Res* 1462:81–92.
19. Zhang Z, et al. (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* 133(4):585–600.
20. Zhang Z, et al. (2013) Dysregulation of synaptogenesis genes antecedes motor neuron pathology in spinal muscular atrophy. *Proc Natl Acad Sci USA* 110(48):19348–19353.
21. Garcia EL, Lu Z, Meers MP, Praveen K, Matera AG (2013) Developmental arrest of *Drosophila* survival motor neuron (Smn) mutants accounts for differences in expression of minor intron-containing genes. *RNA* 19(11):1510–1516.
22. Praveen K, Wen Y, Matera AG (2012) A *Drosophila* model of spinal muscular atrophy uncouples snRNP biogenesis functions of survival motor neuron from locomotion and viability defects. *Cell Reports* 1(6):624–631.
23. Bäumer D, et al. (2009) Alternative splicing events are a late feature of pathology in a mouse model of spinal muscular atrophy. *PLoS Genet* 5(12):e1000773.
24. Corti S, et al. (2008) Neural stem cell transplantation can ameliorate the phenotype of a mouse model of spinal muscular atrophy. *J Clin Invest* 118(10):3316–3330.
25. Ruggiu M, et al. (2012) A role for SMN exon 7 splicing in the selective vulnerability of motor neurons in spinal muscular atrophy. *Mol Cell Biol* 32(1):126–138.
26. Hubers L, et al. (2011) HuD interacts with survival motor neuron protein and can rescue spinal muscular atrophy-like neuronal defects. *Hum Mol Genet* 20(3):553–579.
27. Tadesse H, Deschênes-Furry J, Boisvenue S, Côté J (2008) KH-type splicing regulatory protein interacts with survival motor neuron protein and is misregulated in spinal muscular atrophy. *Hum Mol Genet* 17(4):506–524.
28. Zhao DY, et al. (2016) SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature* 529(7584):48–53.
29. Shen S, et al. (2014) rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci USA* 111(51):E5593–E5601.
30. Alioto TS (2007) U12DB: A database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res* 35(Database issue):D110–D115.
31. Katz Y, Wang ET, Airoldi EM, Burge CB (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 7(12):1009–1015.
32. Braunschweig U, et al. (2014) Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* 24(11):1774–1786.
33. Shalgi R, Hurt JA, Lindquist S, Burge CB (2014) Widespread inhibition of post-transcriptional splicing shapes the cellular transcriptome following heat shock. *Cell Reports* 7(5):1362–1370.
34. Wong JJ, et al. (2013) Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* 154(3):583–595.
35. Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV (2012) Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev* 26(11):1209–1223.
36. Yeo G, Burge CB (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 11(2-3):377–394.
37. Park DS, et al. (1998) Cyclin-dependent kinases participate in death of neurons evoked by DNA-damaging agents. *J Cell Biol* 143(2):457–467.
38. Paulsen RD, et al. (2009) A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell* 35(2):228–239.
39. Aguilera A, García-Muse T (2012) R loops: From transcription byproducts to threats to genome stability. *Mol Cell* 46(2):115–124.
40. Li X, Manley JL (2006) Cotranscriptional processes and their influence on genome stability. *Genes Dev* 20(14):1838–1847.
41. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 18(13):1618–1629.
42. Denis MM, et al. (2005) Escaping the nuclear confines: Signal-dependent pre-mRNA splicing in anucleate platelets. *Cell* 122(3):379–391.
43. Domínguez-Sánchez MS, Barroso S, Gómez-González B, Luna R, Aguilera A (2011) Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet* 7(12):e1002386.
44. Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* 45(6):814–825.
45. Skourti-Stathaki K, Proudfoot NJ, Gromak N (2011) Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* 42(6):794–805.
46. Chou ZF, Chen F, Wilusz J (1994) Sequence and position requirements for uridylate-rich downstream elements of polyadenylation signals. *Nucleic Acids Res* 22(13):2525–2531.
47. Montzka KA, Steitz JA (1988) Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc Natl Acad Sci USA* 85(23):8885–8889.
48. Lotti F, et al. (2012) An SMN-dependent U12 splicing event essential for motor circuit function. *Cell* 151(2):440–454.
49. Jia Y, Mu JC, Ackerman SL (2012) Mutation of a U2 snRNA gene causes global disruption of alternative splicing and neurodegeneration. *Cell* 148(1-2):296–308.
50. Madabhushi R, Pan L, Tsai LH (2014) DNA damage and its links to neurodegeneration. *Neuron* 83(2):266–282.
51. Shiloh Y, Ziv Y (2013) The ATM protein kinase: Regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* 14(4):197–210.
52. Date H, et al. (2001) Early-onset ataxia with ocular motor apraxia and hypoalbuminemia is caused by mutations in a new HIT superfamily gene. *Nat Genet* 29(2):184–188.
53. Moreira MC, et al. (2001) The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. *Nat Genet* 29(2):189–193.
54. Sanders LH, et al. (2014) Mitochondrial DNA damage: Molecular marker of vulnerable nigral neurons in Parkinson's disease. *Neurobiol Dis* 70:214–223.
55. Lovell MA, Markesbery WR (2007) Oxidative DNA damage in mild cognitive impairment and late-stage Alzheimer's disease. *Nucleic Acids Res* 35(22):7497–7504.
56. Wahba L, Amon JD, Koshland D, Vuica-Ross M (2011) RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol Cell* 44(6):978–988.
57. Haeusler AR, et al. (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507(7491):195–200.
58. Grabczyk E, Mancuso M, Sammarco MC (2007) A persistent RNA.DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP in vitro. *Nucleic Acids Res* 35(16):5351–5359.
59. Chen YZ, et al. (2004) DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet* 74(6):1128–1135.
60. Grohmann K, et al. (2001) Mutations in the gene encoding immunoglobulin mu-binding protein 2 cause spinal muscular atrophy with respiratory distress type 1. *Nat Genet* 29(1):75–77.
61. Fukita Y, et al. (1993) The human S mu bp-2, a DNA-binding protein specific to the single-stranded guanine-rich sequence related to the immunoglobulin mu chain switch region. *J Biol Chem* 268(23):17463–17470.
62. Ling SC, Polymenidou M, Cleveland DW (2013) Converging mechanisms in ALS and FTD: Disrupted RNA and protein homeostasis. *Neuron* 79(3):416–438.
63. Lagier-Tourenne C, Cleveland DW (2009) Rethinking ALS: The FUS about TDP-43. *Cell* 136(6):1001–1004.
64. Han H, et al. (2013) MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* 498(7453):241–245.
65. Sun S, et al. (2015) ALS-causative mutations in FUS/TLS confer gain and loss of function by altered association with SMN and U1-snRNP. *Nat Commun* 6:6171.
66. Hua Y, Zhou J (2004) Survival motor neuron protein facilitates assembly of stress granules. *FEBS Lett* 572(1-3):69–74.
67. Salvi JS, Mekhail K (2015) R-loops highlight the nucleus in ALS. *Nucleus* 6(1):23–29.
68. Hill SJ, et al. (2016) Two familial ALS proteins function in prevention/repair of transcription-associated DNA damage. *Proc Natl Acad Sci USA* 113(48):E7701–E7709.
69. Garcia EL, Wen Y, Praveen K, Matera AG (2016) Transcriptomic comparison of *Drosophila* snRNP biogenesis mutants reveals mutant-specific changes in pre-mRNA processing: Implications for spinal muscular atrophy. *RNA* 22(8):1215–1227.
70. Fayzullina S, Martin LJ (2014) Skeletal muscle DNA damage precedes spinal motor neuron DNA damage in a mouse model of spinal muscular atrophy (SMA). *PLoS One* 9(3):e93329.
71. Loomis EW, Sanz LA, Chedin F, Hagerman PJ (2014) Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet* 10(4):e1004294.
72. Maury Y, et al. (2014) Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* 33(1):89–96.
73. Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.