UC Irvine
UC Irvine Previously Published Works

Title

Altered H3 histone acetylation impairs high-fidelity DNA repair to promote cerebellar degeneration in spinocerebellar ataxia type 7

Permalink

https://escholarship.org/uc/item/4043s7nk

Journal

Cell Reports, 37(9)

ISSN

2639-1856

Authors

Switonski, Pawel M
Delaney, Joe R
Bartelt, Luke C
et al.

Publication Date
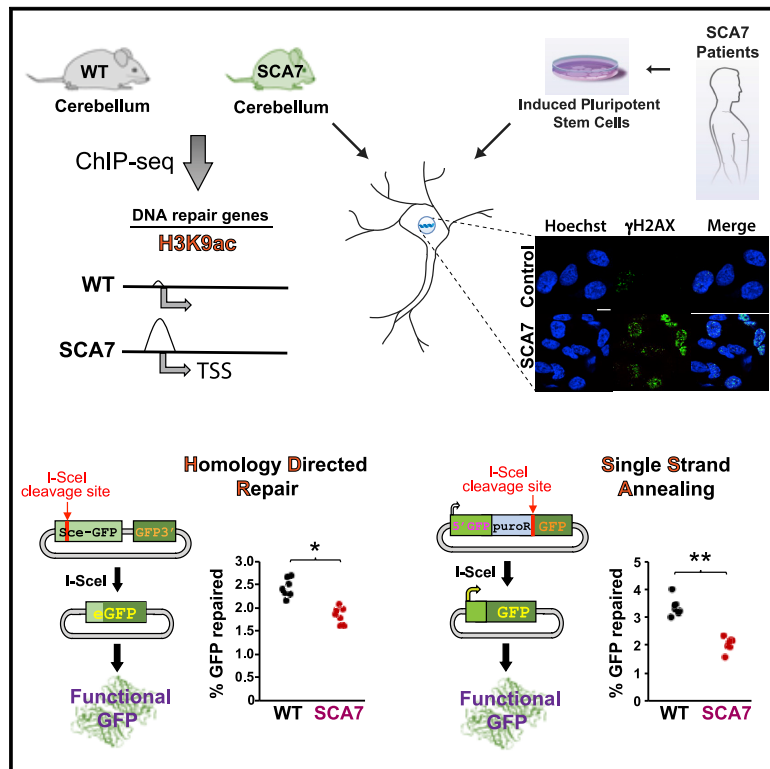
2021-11-01

DOI

10.1016/j.celrep.2021.110062

Copyright Information

Peer reviewed

Altered H3 histone acetylation impairs high-fidelity DNA repair to promote cerebellar degeneration in spinocerebellar ataxia type 7

Graphical abstract

Authors

Pawel M. Switonski, Joe R. Delaney, Luke C. Bartelt, ..., Jaidev Bapat, Bryce L. Sopher, Albert R. La Spada

Correspondence

alaspada@uci.edu

In brief

Switonski et al. performed ChIP-seq on cerebellar DNA from SCA7 mice and detect increased histone H3-promoter acetylation in DNA repair genes. They document DNA damage in SCA7 models and patient stem-cell-derived neurons. Using *in vitro* assays and genome-wide translocation sequencing, they observe altered DNA repair in SCA7.

Highlights

- SCA7 mice display increased histone H3 promoter Ac in DNA repair genes in cerebellum

- DNA damage is elevated in SCA7 cells, primary neurons, and patient stem cell neurons

- DNA repair assays in cells expressing polyQ-ataxin-7 show deficits in specific pathways

- LAM-HTGTS identifies altered repair processes on endogenous DNA in SCA7 cell models

CellPress

Cell Reports

Report

Altered H3 histone acetylation impairs high-fidelity DNA repair to promote cerebellar degeneration in spinocerebellar ataxia type 7

Pawel M. Switonski,[1,2,3,9] Joe R. Delaney,[2,8,9] Luke C. Bartelt,[2] Chenchen Niu,[1] Maria Ramos-Zapatero,[4] Nathanael J. Spann,[5] Akshay Alaghatta,[2] Toby Chen,[2] Emily N. Griffin,[5] Jaidev Bapat,[4] Bryce L. Sopher,[6] and Albert R. La Spada[1,2,7,*]

[1]Departments of Pathology & Laboratory Medicine, Neurology, and Biological Chemistry, University of California, Irvine, Irvine, CA 92697, USA
[2]Department of Neurology, Duke University School of Medicine, Durham, NC 27710, USA
[3]Department of Medical Biotechnology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14 Str., 61-704 Poznan, Poland
[4]Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093, USA
[5]Department of Cellular & Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA
[6]Department of Neurology, University of Washington Medical Center, Seattle, WA 98195, USA
[7]UCI Institute for Neurotherapeutics, University of California, Irvine, Irvine, CA 92697, USA
[8]Present address: Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC 29425, USA
[9]These authors contributed equally
*Correspondence: alaspada@uci.edu
Lead contact
https://doi.org/10.1016/j.celrep.2021.110062

SUMMARY

A common mechanism in inherited ataxia is a vulnerability of DNA damage. Spinocerebellar ataxia type 7 (SCA7) is a CAG-polyglutamine-repeat disorder characterized by cerebellar and retinal degeneration. Polyglutamine-expanded ataxin-7 protein incorporates into STAGA co-activator complex and interferes with transcription by altering histone acetylation. We performed chromatic immunoprecipitation sequencing ChIP-seq on cerebellum from SCA7 mice and observed increased H3K9-promoter acetylation in DNA repair genes, resulting in increased expression. After detecting increased DNA damage in SCA7 cells, mouse primary cerebellar neurons, and patient stem-cell-derived neurons, we documented reduced homology-directed repair (HDR) and single-strand annealing (SSA). To evaluate repair at endogenous DNA in native chromosome context, we modified linear amplification-mediated high-throughput genome-wide translocation sequencing and found that DNA translocations are less frequent in SCA7 models, consistent with decreased HDR and SSA. Altered DNA repair function in SCA7 may predispose the subject to excessive DNA damage, leading to neuron demise and highlights DNA repair as a therapy target.

INTRODUCTION

Spinocerebellar ataxia type 7 (SCA7) is an inherited neurological disorder characterized by cerebellar and retinal degeneration (Garden and La Spada, 2008). SCA7 is one of nine inherited disorders caused by CAG-polyglutamine repeats that expand to cause disease (Zoghbi and Orr, 2000). Ataxin-7, the causal disease protein in SCA7, contains a polyglutamine tract in the amino-terminus of ataxin-7, beginning at position 30. Ataxin-7 allele sizes range from 4–35 CAGs in unaffected individuals, but expand to 37 to >400 CAGs in affected patients (David et al., 1997; Stevanin et al., 2000). SCA7, thus, belongs to the CAG-polyglutamine-repeat disease category, which includes spinobulbar muscular atrophy (SBMA), Huntington disease (HD), dentatorubral-pallidoluysian atrophy (DRPLA), and five other forms of spinocerebellar ataxia (SCA1, 2, 3, 6, and 17).

Numerous lines of investigation suggest that polyglutamine tract expansion is a gain-of-function mutation and initiates a cascade of events culminating in the dysfunction, degeneration, and, ultimately, death of neurons. As transition from a normal length glutamine tract to a polyglutamine expansion results in a shift to an altered protein conformation (Paulson et al., 2000; Ross, 1997), SCA7 and the other eight polyglutamine diseases share a common pathogenic basis with Alzheimer disease, Parkinson disease, amyotrophic lateral sclerosis, and the tauopathies.

In 2001, we created a representative mouse model for SCA7 retinal degeneration and documented that SCA7 cone-rod dystrophy resulted from transcription interference in retinal photoreceptors (La Spada et al., 2001). Our discovery of an interaction between normal ataxin-7 and cone-rod homeobox protein (CRX), together with our realization that ataxin-7 possesses a functional nuclear localization signal (Chen et al., 2004),

suggested that ataxin-7 is a transcription factor. When studies of Sgf73, the yeast ortholog of ataxin-7, indicated that Sgf73 is part of the SAGA co-activator complex (Sanders et al., 2002), we hypothesized that ataxin-7 could be a component of the analogous mammalian co-activator complex, which is known as the STAGA (for Spt3-Taf9-Ada-Gcn5-acetyltransferase) complex (Martinez et al., 1998). The first indication that ataxin-7 is indeed a member of the STAGA complex came when mass spectrometry analysis of a 110-kDa protein, which we purified by immunoprecipitation of a FLAG-tagged STAGA complex protein, yielded a peptide sequence matching ataxin-7 (Martinez et al., 1998). Independent biochemical studies of the STAGA complex confirmed that ataxin-7 is a core component of STAGA (Helmlinger et al., 2004; Palhan et al., 2005). Although the exact function of ataxin-7 in STAGA is unknown, polyglutamine expansion in ataxin-7 can disrupt the histone acetyltransferase activity of SAGA in yeast and of STAGA in retinal photoreceptor cells (Helmlinger et al., 2006; McMahon et al., 2005; Palhan et al., 2005), indicating that epigenetic dysregulation is likely a central feature of SCA7 disease pathogenesis.

To determine whether the STAGA complex dysfunction affects chromatin remodeling in SCA7 cerebellar degeneration, we pursued unbiased chromatin immunoprecipitation sequencing (ChIP-seq) analysis of histone 3 lysine 9 acetylation (H3K9ac) because that histone mark is an established indicator of open chromatin in the transcription regulatory regions of most genes (Jenuwein and Allis, 2001). We studied SCA7 266Q knockin mice, which faithfully recapitulate a severe, rapidly progressive, neurodegenerative phenotype involving the cerebellum (Yoo et al., 2003). We detected altered H3K9ac in the promoter regions of 777 genes in the cerebellum of SCA7 266Q knockin mice and wild-type (WT) littermate control mice and noted that most chromatin alterations corresponded to markedly increased H3K9 promoter acetylation. Gene Ontology (GO) analysis revealed an enrichment for DNA repair genes, and validation studies confirmed both altered chromatin remodeling in SCA7 cerebellar DNA repair genes and increased gene expression. Because DNA repair gene mutations are a major cause of inherited ataxia syndromes (Paulson and Miller, 2005) and SCA7 cerebellar degeneration has been linked to PARP1 hyperactivation in association with DNA damage, resulting in depletion of nicotinamide adenine dinucleotide, which contributes to neuron demise by impairing the function of the neuroprotective, anti-aging regulator sirtuin-1 (Stoyas et al., 2020; Ward et al., 2019), we quantified DNA damage in SCA7 model systems and noted marked increases in SCA7 cell lines, mouse primary neurons, and human-patient stem-cell-derived neurons. We examined DNA repair function using reporter assays and unbiased high-throughput next-generation sequencing and documented specific repair-pathway abnormalities, indicating a role for abnormal DNA repair in the pathogenesis of SCA7.

RESULTS

ChIP-seq analysis reveals a preponderance of genes with increased promoter histone acetylation in the SCA7 cerebellum

To examine the role of altered STAGA complex function in SCA7-cerebellar neurodegeneration, we performed unbiased ChIP-

seq analysis of histone acetylation on the cerebella isolated from SCA7 266Q knockin mice and WT littermate controls, using a well validated antibody directed against H3K9ac (O'Geen et al., 2011). The resulting sequence FASTQ files were then mapped to the mouse reference genome (mm10), and peaks were called by model-based analysis of ChIP-seq (MACs) (Feng et al., 2012) (Figure S1A). To determine whether this ChIP-seq experiment had successfully enriched for chromatin near the transcription start site (TSS), we derived a randomized distribution of peak locations relative to the TSS, which we compared to the relative TSS peak locations obtained for SCA7 and control samples (Figure S1B). We found that most of our ChIP-seq reads for both SCA7 and control samples mapped to within 1 kb of the TSSs, with only a small fraction of ChIP-seq reads mapping randomly. As H3K9ac marks the promoter regions of active genes (Bártová et al., 2008), we matched peaks to their adjacent genes and then compared differential promoter H3K9ac peaks between the SCA7 cerebellum and the WT cerebellum. Although most gene-promoter H3K9ac peaks were shared between SCA7 and WT mice, we identified 777 H3K9ac sites that were enriched adjacent to genes in either SCA7 or WT cerebellum chromatin. Of the 777 H3K9ac differentially enriched promoter peaks, 746 H3K9ac gene promoter peaks were increased in SCA7 mice, with only 31 H3K9ac gene promoter peaks increased in WT mice (Figure S1C, Data S1). Previous studies have found that incorporation of polyglutamine-expanded ataxin-7 into the STAGA complex can diminish histone acetyltransferase activity at the promoters of certain genes (Helmlinger et al., 2004; Palhan et al., 2005) but can also result in increased histone acetyltransferase activity (Helmlinger et al., 2006). Our results indicate that, in the degenerating cerebellum of affected SCA7 266Q knockin mice, STAGA complex dysfunction leads to enhanced H3K9ac marks in the promoters of more than 700 genes.

Epigenetic dysregulation of DNA repair genes occurs in SCA7 cerebellar degeneration

To identify pathways in the cerebellum preferentially affected by altered STAGA complex histone acetyltransferase activity in SCA7 neurodegeneration, we performed GO analysis with two different software algorithms, GOrilla (Eden et al., 2009) and ShinyGO (Ge et al., 2020). Using GOrilla, we uncovered 26 pathways, and using ShinyGO, we uncovered eight pathways, all selectively enriched for the 746 SCA7 genes with altered H3K9ac promoter marks at a false-discovery rate of <0.001 (Table S1). Many of these pathways involved metabolic processes; however, the GO terms "DNA repair" and "Cellular response to DNA damage stimulus" appeared using both software algorithms and were ranked highest by the ShinyGO analysis. As we recently implicated PARP1 hyperactivation in response to increased DNA damage as a factor in SCA7 cerebellar degeneration (Stoyas et al., 2020), our ChIP-seq findings suggest that epigenetic dysregulation of DNA repair gene expression could be contributing to the propensity for DNA damage in SCA7 cerebellar degeneration. Indeed, ChIP-seq analysis of SCA7 cerebellum revealed 33 DNA repair genes with significantly increased histone H3 promoter acetylation, and their gene products are involved in a many essential DNA repair pathways (Table 1). To validate our ChIP-seq findings, we considered that the results

Table 1. DNA repair genes with increased H3K9 acetylation at promoter regions in SCA7 266Q knockin mice

Gene symbol	Gene name	DNA repair pathway
ACTL6A	Actin-like 6a	HR
ACTR8	Arp8 actin-related protein 8	HR
AP5Z1	Adaptor-related protein complex 5, zeta 1 subunit	HR
BRCA2	Breast cancer 2	HDR
CDK2	Cyclin-dependent kinase 2	HR, cell cycle arrest
CHAF1A	Chromatin assembly factor 1, subunit a (p150)	NER
DCLRE1B	DNA cross-link repair 1b, pso2 homolog (*S. cerevisiae*)	NHEJ, FA pathway
DCAF8	DDB1 and CUL4 associated factor 8	NER
DDB2	Damage-specific DNA binding protein 2	NER
EEPD1	Endonuclease/exonuclease/phosphatase family domain containing 1	HR
ERCC5	Excision repair cross-complementing, complementation group 5	NER
EXO5	Exonuclease 5	HDR
FANCI	Fanconi anemia, complementation group i	FA pathway
FBXO18	F-box protein 18	HR, FA pathway
GTF2H2	General transcription factor ii-h, polypeptide 2	NER
MAD2L2	Mad2 mitotic arrest deficient-like 2	FA pathway
MLH1	Mut l homolog 1 (*E. coli*)	BER
NTHL1	Nth (endonuclease iii)-like 1 (*E. coli*)	BER
OGG1	8-oxoguanine DNA-glycosylase 1	BER
OTUB1	Otu domain, ubiquitin aldehyde binding 1	HR
PNKP	Polynucleotide kinase 3'- phosphatase	SSBR/BER, NHEJ, TCR
POLD2	Polymerase (DNA-directed) delta 2, regulatory subunit	MMR, NER, FA pathway
POLI	Polymerase (DNA-directed) iota	FA pathway
POLM	Polymerase (DNA-directed) mu	NHEJ
PRPF19	Prp19/Pso4 pre-mRNA processing factor 19 homolog (*S. cerevisiae*)	HR, NHEJ
RAD9A	Rad9 homolog a	BER, HR
REXO4	Rex4 homolog, 3'–5' exonuclease	HDR
RFC1	Replication factor c (activator 1) 1	HR, FA pathway
RFWD3	Ring finger and WD repeat domain 3	HR, FA pathway
RNF168	Ring finger protein 168	FA pathway
SLX1b	Slx1 structure-specific endonuclease subunit homolog b (*S. cerevisiae*)	HR
TRP53	Transformation-related protein 53	cell cycle arrest
XRCC5	X-ray repair cross-complementing 5	HR, NHEJ

Abbreviations: BER, base excision repair; HDR, homology directed repair; HR, homologous recombination; FA, Fanconi anemia; MMR, mismatch repair; NER, nucleotide excision repair; NHEJ, non-homologous end joining; SSBR, single-strand break repair; TCR, transcription-coupled repair

of model-based analysis of ChIP-seq (MACS) promoter peak calls for a subset of eight "hits" selected from four different, broadly representative DNA repair pathways, including non-homologous end-joining (NHEJ), the Fanconi anemia (FA) repair pathway, nucleotide excision repair (NER), and homology-directed repair (HDR), and for *Trp53* because it is central to the DNA damage response (Figure 1A). We also evaluated the *Igfbp5* gene for purposes of quality control and confirmed reduced H3K9ac in the SCA7 cerebellum, in agreement with a previous publication (Gatchel et al., 2008). We then performed direct ChIP analysis on cerebellar samples obtained from SCA7 266Q knockin mice and WT littermate controls for all eight DNA repair/damage-response genes, the cell cycle arrest gene *Cdk2*, and the gene for polynucleotide-kinase-3'-phosphatase (PNKP) because PNKP has been implicated in altered DNA repair in polyglutamine neurodegeneration (Gao et al., 2019). We confirmed significant increases in promoter H3K9ac for eight of the 10 genes in the cerebellum of SCA7 mice, and we noted a strong trend for *Pnkp* (Figure 1B). To determine whether increased promoter H3K9ac affects the expression of these genes, we performed RT-PCR analysis and observed markedly increased expression for eight of these 10 genes in cerebellar RNAs from SCA7 266Q knockin mice, with six genes showing both significantly increased promoter H3K9ac and significantly increased RNA expression (Figure 1C). These findings confirm that epigenetic dysregulation of DNA repair genes could be involved in SCA7 cerebellar degeneration. Because SCA7 266Q knockin mice also recapitulate the retinal degeneration phenotype observed in human SCA7 patients (Yoo et al., 2003), we obtained retinal RNAs from SCA7 mice and WT
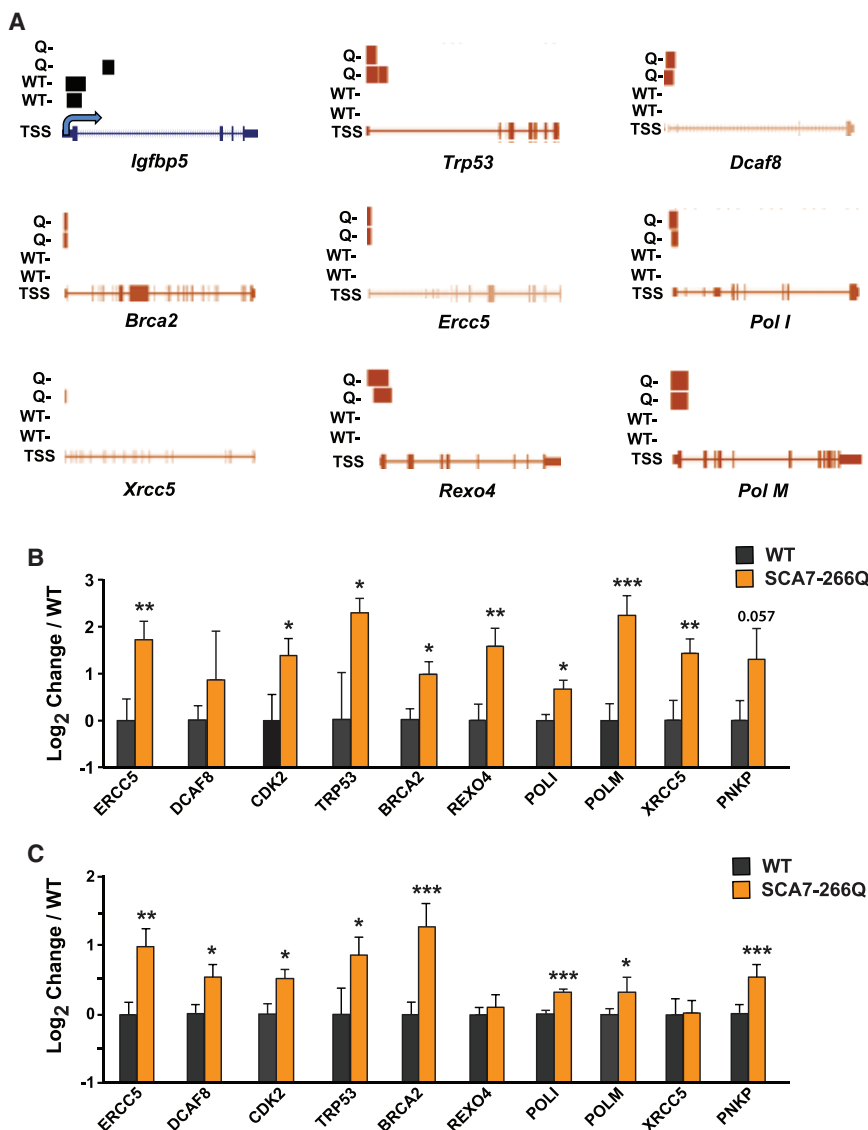
Figure 1. DNA repair gene expression is upregulated in the cerebellum of SCA7 mice

(A) Processed data from ChIP-seq analysis of SCA7 266Q knockin mice (Q) and WT control cerebellum. Gene maps from the UCSC genome browser display the TSS on the left and show all exons as boxes, up to the final exon on the right. MACS peaks are plotted as horizontal bars above each gene map, with the thickness of the bar corresponding to the height of the peak. Results for *IGFBP5* are included for purposes of quality control because a previous study reported dramatic downregulation of *IGFBP5* expression in SCA7 266Q knockin cerebellum.

(B) ChIP analysis of H3K9ac enrichment in SCA7 266Q knockin mice. ChIP was performed on cerebellar DNA from SCA7 and WT 9-week-old mice (n = 4/group) with an antibody against H3K9ac for promoter regions of the indicated genes. Results for qPCR analysis are for IP/input DNA, with H3K9ac occupancy for WT arbitrarily set to 1. Two-tailed t test; *p < 0.05, **p < 0.01. Error bars: SEM. The exact p value for the *Pnkp* gene promoter region is shown to indicate a strong trend.

(C) RT-PCR analysis of cerebellar RNAs from 9-week-old SCA7 266Q knockin mice and WT littermate controls (n = 3/group). Results for qRT-PCR analysis are normalized to *Tubg1*, with expression for WT arbitrarily set to 1. Two-tailed t test; *p < 0.05, **p < 0.01. Error bars: SEM.

littermate controls, and we measured the expression levels of five DNA repair genes, but we did not detect any significant alterations in DNA repair gene expression in the SCA7 retina (Figure S2).

SCA7 epigenetic dysregulation in cerebellum affects major DNA repair pathways

Repair of DNA double-strand breaks can occur through a variety of different pathways. When we surveyed H3K9ac promoter peaks for key DNA repair regulatory genes, we noted that multiple DNA repair factors simultaneously exhibit altered epigenetic regulation for certain pathways (Figure S3A), NHEJ, the FA repair pathway, base excision repair (BER), NER, HDR, and mismatch repair (MMR). For NHEJ, the FA pathway, and NER, increased H3K9ac peaks occur for genes whose products all promote the function of the DNA repair pathway; however, in other cases, such as for HDR, genes encoding both

positive and negative regulators display increased H3K9 promoter acetylation. Although most DNA repair genes exhibit increased H3K9 promoter acetylation, there are a few rare instances in which the promoter H3K9ac was decreased. To fully appreciate the scope of DNA repair gene epigenetic dysregulation in the cerebellum of SCA7-model mice, we generated a protein-protein interaction network for DNA repair from interaction data in the BioGRID resource (Chatr-Aryamontri et al., 2017). Of 1,022 interactions that occur between various DNA repair proteins, 303 of these interactions (i.e., 29.6%) involve proteins encoded by genes found to display altered promoter H3K9ac in the cerebellum of SCA7 mice (Figure S3B), with the protein products of epigenetically dysregulated DNA repair genes linked to at least six major DNA repair pathways.

Polyglutamine-expanded ataxin-7 promotes DNA damage in cell culture models, mouse primary cerebellar neurons, and human-patient stem-cell-derived neurons

To determine whether expression of polyglutamine-expanded ataxin-7 protein is sufficient to induce DNA damage, we initially sought evidence for DNA damage in cell culture models of SCA7 by measuring levels of γH2AX, which is a sensitive and specific marker of DNA double-strand breaks
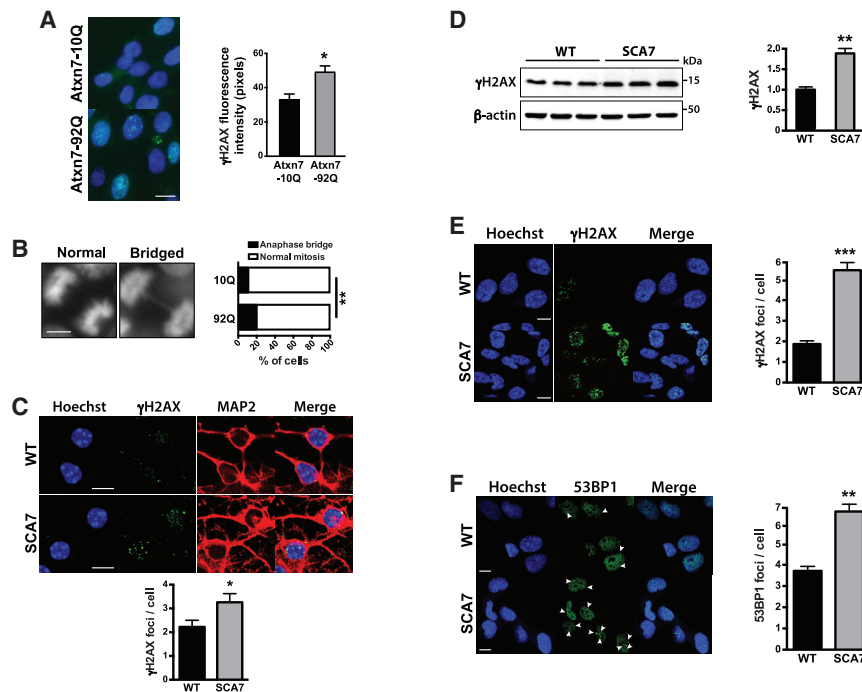
Figure 2. DNA damage is increased in cell culture and mouse models of SCA7

(A) We transfected HEK293T cells with either an ataxin-7 10Q or ataxin-7 92Q expression construct, fixed cells after 48 h, immunostained with anti-γH2AX antibody (green), and counterstained with Hoechst 33342 (blue). Quantification of γH2AX immunostaining intensity is shown to the right for n = 4 independent experiments. Two-tailed t test; *p < 0.05. Scale bar: 10 μM. Error bars: SEM.

(B) We transfected HEK293T cells with either an ataxin-7 10Q or ataxin-7 92Q expression construct, fixed cells after 48 hr, and stained with DAPI to visualize anaphase bridge formation; n = 3 independent experiments. Fisher's exact test; **p < 0.01. Scale bar: 10 μM

(C) We cultured primary cerebellar granule cell neurons from SCA7 r210 knockin mice and WT littermate controls, and then immunostained DIV10 neurons with anti-γH2AX antibody (green), anti-MAP2 antibody (red), counterstaining with Hoechst 33342 (blue). Quantification of γH2AX foci formation is shown to the right for n = 3 independent experiments. Two-tailed t test; *p < 0.05. Scale bar: 10 μM. Error bars: SEM.

(D) We prepared protein lysates from the cerebellum of 8-week-old SCA7 266Q knockin mice and WT littermate controls (n = 3/genotype) and immunoblotted with anti-γH2AX antibody and

anti-β-actin antibody. Quantification of γH2AX normalized to β-actin is shown to the right. Two-tailed t test; **p < 0.01. Error bars: SEM.

(E) We derived neural progenitor cells (NPCs) from patients with SCA7 and unaffected first-degree relatives as controls, immunostained with anti-γH2AX antibody (green), and counterstained with Hoechst 33342 (blue). Quantification of γH2AX foci formation is shown to the right for n = 2 patients/genotype, n = 2 independent experiments. Two-tailed t test; ***p < 0.001. Scale bar: 10 μM. Error bars: SEM.

(F) We derived neural progenitor cells (NPCs) from patients with SCA7 and unaffected first-degree relatives as controls, immunostained with anti-53BP1 antibody (green), and counterstained with Hoechst 33342 (blue). Quantification of 53BP1 puncta formation is shown to the right for n = 2 patients/genotype, n = 2 independent experiments. Two-tailed t test; **p < 0.01. Scale bar: 10 μM. Error bars: SEM.

(Kuo and Yang, 2008). We detected significantly increased γH2AX immunoreactivity in HEK293T cells expressing ataxin-7 92Q in comparison to HEK293T cells expressing normal ataxin-7 (Figure 2A). One focus of prolonged damage repair signaling for a DNA break is at unprotected telomeres (d'Adda di Fagagna et al., 2003). In dividing cells, such as HEK293T cells, when DNA breakage at telomeres cannot be repaired, the cell will recruit the NHEJ machinery to fuse the broken telomeres to other sites within the genome (Terasawa et al., 2014), which often occurs when the cell is not in mitosis (Smogorzewska et al., 2002). Typically, this fusion point is on another chromosome, resulting in a dicentric chromosome. When the cell then enters mitosis, such dicentric chromosomes cannot separate during anaphase, resulting in the formation of an anaphase bridge (Stimpson et al., 2010). When we quantified the anaphase bridge formation in HEK293T cells expressing ataxin-92Q or ataxin-7 10Q, we observed a near doubling of anaphase bridges in HEK293T cells expressing ataxin-7 92Q (Figure 2B). To determine whether DNA damage is increased in cerebellar neurons expressing polyglutamine-expanded ataxin-7 protein, we cultured primary cerebellar granule cell neurons (GCNs) from SCA7 mice and WT littermate controls, and upon immunostaining for γH2AX, we noted a significant increase in γH2AX foci in SCA7 cerebellar GCNs (Figure 2C). To independently corroborate those findings, we

performed immunoblot analysis of γH2AX and observed a marked increase in γH2AX levels in cerebellar protein lysates from SCA7 mice (Figure 2D). We also measured γH2AX in an inducible PC12 cell culture model of SCA7 and, similarly, observed an increase in γH2AX immunoreactivity in PC12 cells stably expressing ataxin-7 65Q in comparison to PC12 cells stably expressing ataxin-7 10Q (Figure S4).

To further examine the pathological relevance of DNA damage to SCA7 disease pathogenesis, we studied progenitor neurons derived from induced pluripotent stem cells generated from SCA7 patients and first-degree relative controls, using previously characterized independent clonal lines (Ward et al., 2019). When we immunostained neural progenitor cells (NPCs) derived from patients with SCA7 and related controls for γH2AX, we documented dramatically increased γH2AX foci formation in SCA7 NPCs (Figure 2E). We also analyzed the formation of 53BP1 foci because γH2AX recruits 53BP1 to facilitate the process of DNA double-strand break repair (Fernandez-Capetillo et al., 2003), and we observed a marked increase in 53BP1 foci formation in SCA7 NPCs in comparison with control NPCs (Figure 2F). These results, which were obtained for SCA7 NPCs under normal cell culture conditions without use of a stress treatment or insult, indicate that polyglutamine-expanded ataxin-7 protein is a potent inducer of DNA damage.
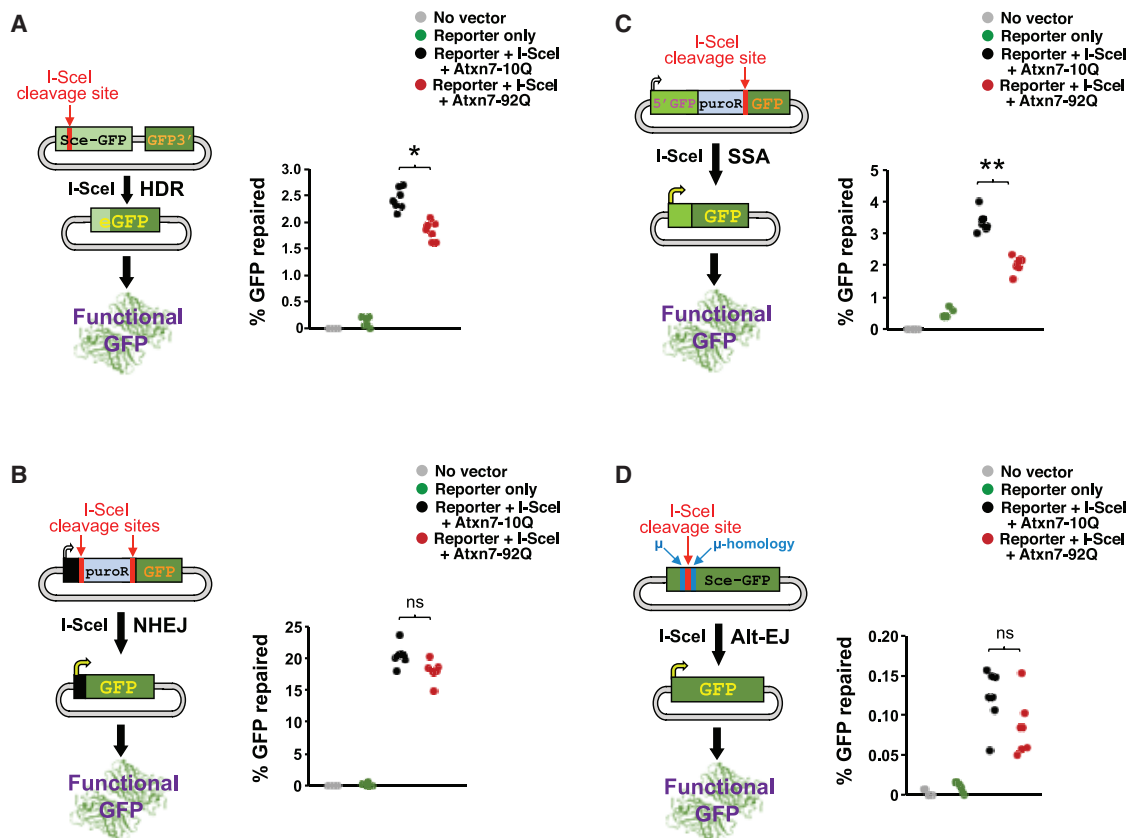
Figure 3. Polyglutamine-expanded ataxin-7 blocks specific DNA repair pathways
We transfected HEK293T cells with only the DNA repair reporter construct (Reporter only) or co-transfected with either an ataxin-7 10Q or ataxin-7 92Q expression construct in combination with the indicated DNA repair reporter (illustrated in each panel) and a vector encoding I-*Sce* I nuclease, which cleaves the reporter construct as demarcated to allow for quantitation of double-strand DNA break repair. Untransfected HEK293T cells (No vector) served as another negative control.
(A) Results for homology-directed repair reporter construct; n = 6 independent experiments. Two-tailed t test; *p < 0.05.
(B) Results for non-homologous end joining repair reporter construct; n = 6 independent experiments. Two-tailed t test.
(C) Results for single-strand annealing repair reporter construct; n = 6 independent experiments. Two-tailed t test; **p < 0.01.
(D) Results for alternative end joining repair reporter construct; n = 6 independent experiments. Two-tailed t test.

Specific DNA repair pathways are impaired in SCA7

Here, we have shown that polyglutamine-expanded ataxin-7 promotes DNA damage, and because we have previously reported that PARP1 hyperactivation contributes to SCA7 cerebellar degeneration (Stoyas et al., 2020), an important question is whether polyglutamine-expanded ataxin-7 is directly impairing the DNA repair process. To examine that issue, we performed DNA repair assays, using repair reporter plasmids with GFP genes susceptible to cleavage and inactivation by the meganuclease I-*Sce*I (Bennardo et al., 2008; Pierce et al., 1999; Stark et al., 2004). Upon I-*Sce*I cleavage, GFP fluorescence can only be recovered by the activity of a specific DNA repair pathway. For each repair reporter construct, we measured recovery of GFP fluorescence for HEK293T cells transfected with either ataxin-7 92Q or ataxin-7 10Q. We evaluated four different DNA repair pathways: HDR, NHEJ, SSA, and alternative end joining (Alt-EJ). We found that polyglutamine-expanded ataxin-7 significantly reduced HDR and SSA, but did not affect NHEJ or Alt-EJ (Figure 3). These results suggest that altered DNA repair pathway function may occur in SCA7.

Genome-wide translocation sequencing reveals that polyglutamine-expanded ataxin-7 alters DNA repair pathways used on endogenous DNA

To further investigate the disruption of functional DNA repair in a more physiological setting, we next investigated the effect of polyglutamine-expanded ataxin-7 on repair of endogenous DNA in native chromosomes by modifying a method for detection of double-strand breaks in mammalian genomes, known as linear amplification-mediated high-throughout genome-wide translocation sequencing (LAM-HTGTS) (Hu et al., 2016). To introduce a double-strand DNA break, we designed single-guide RNA (sgRNA) guides complementary to a specific sequence on mouse chromosome 14, and employed CRISPR-Cas9 to create a "bait" site. When a double-strand break is introduced at the bait site, there are three possible outcomes: (1) proper repair by NHEJ/HDR, (2) an insertion/deletion (indel) mutation introduced by erroneous NHEJ-mediated repair, or (3) annealing with a DNA double-strand break, resulting in a translocation. We then performed LAM-HTGTS, but
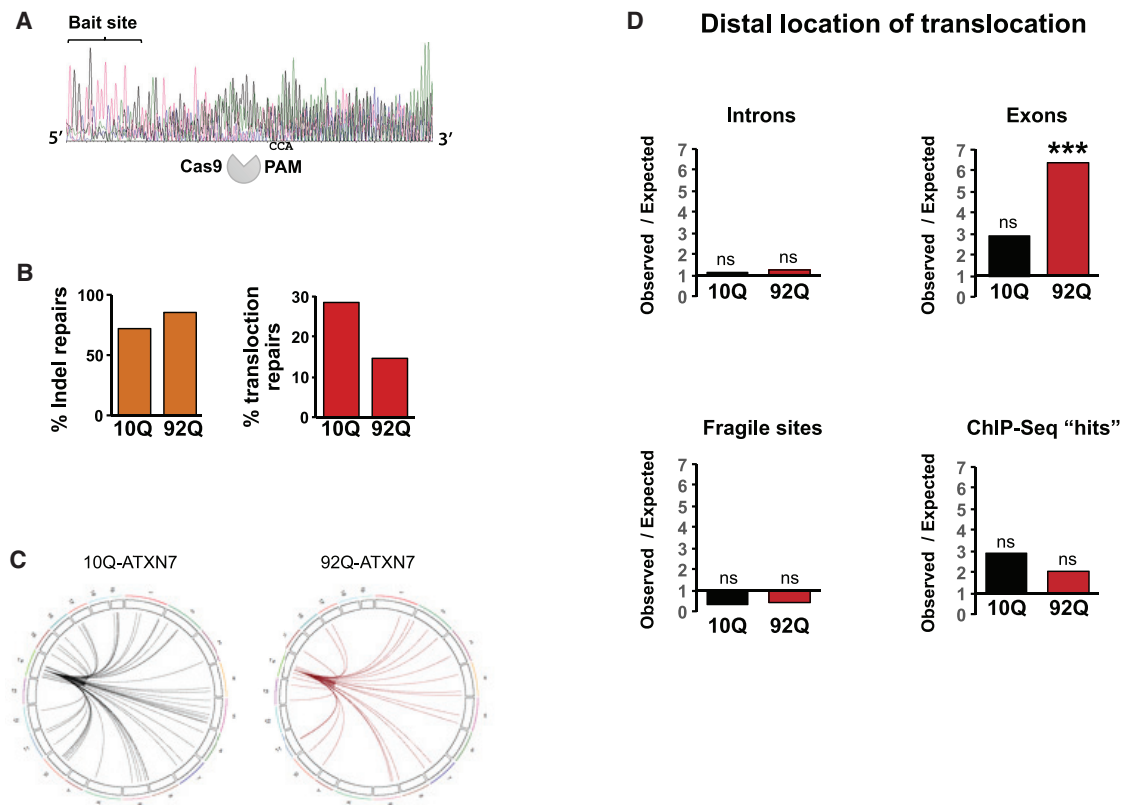
Figure 4. Polyglutamine-expanded ataxin-7 suppresses repair of double-strand DNA breaks by translocation and favors distal translocations with coding sequence regions

(A) We transfected Neuro2a cells with CRISPR-Cas9 and a plasmid encoding sgRNA targeting the bait site, co-transfected either an ataxin-7 10Q or ataxin-7 92Q expression construct and then performed the protocol shown in Figure S7A. Here, we see a representative sequence chromatogram for DNA libraries enriched for the bait CRISPR-Cas9 cut site, with *LpnPI* post-capture enrichment. Note the diversity of sequence 3′ to the bait site, consistent with common translocation events.

(B) Quantitation of repair outcomes for Neuro2a cells expressing ataxin-7 10Q or ataxin-7 92Q.

(C) We mapped 100 randomly selected local and distal translocations to the mm10 genome for Neuro2a cells expressing either ataxin-7 10Q or ataxin-7 92Q. Note that translocation events are distributed throughout the genome, as expected.

(D) After using permutation analysis to obtain an expected rate of distal translocation events for different DNA regions across the mm10 genome, we calculated the rate of distal translocation events at those different types of DNA sequence locations. Results are shown for next-generation sequencing from libraries pooled for genotype from three independent transfections. Fisher's exact test; ***p < 0.001.

included a 5-methyl cytosine modified primer that anneals 3′ to the site of the double-strand break to permit digestion of correctly repaired DNA by *LpnPI* to enrich for translocation events because *LpnPI* only cleaves its recognition site if it is at least hemi-methylated (Figure S5A). To validate this strategy, we transfected Neuro2a cells with CRISPR-Cas9 and the sgRNA and, then, performed LAM-HTGTS with the 5-methyl cytosine modified primer addition step but processed isolated DNA either with or without *LpnPI* treatment. As expected, bait reads containing translocated DNA were almost exclusively identified in the *LpnPI*-processed libraries. (Figure S5B). To confirm the specificity of the sgRNA targeting, we repeated our modified LAM-HTGTS approach, comparing Neuro2a cells transfected with sgRNA targeting the bait site or transfected with a non-targeting sgRNA. Sequence changes at the bait site (non-translocated reads) occur infrequently for the non-targeting sgRNA but account for most non-translocated reads

with sgRNA targeting the bait site (Figure S5C). When we mapped the translocation event sequencing reads obtained for LAM-HTGTS on target sgRNA versus non-targeting sgRNA, we found that translocation reads were enriched with sgRNA targeting the bait site, and distal translocations were only detected with sgRNA targeting the bait site (Figure S5D).

After validating our modified LAM-HTGTS protocol, we used this protocol to compare the outcome of DNA repair for the bait site introduced into Neuro2a cells expressing either ataxin-92Q or normal ataxin-7 10Q. When we analyzed the sequence reads obtained from libraries generated from Neuro2a cells expressing ataxin-7 92Q or ataxin-7 10Q (Figure S6), we confirmed that sequences 3′ to the bait site are much more divergent than sequences 5′ to the bait side, which is indicative of frequent translocations, as expected (Figure 4A, Data S3). We then compared repair outcomes for Neuro2a cells expressing ataxin-7 92Q to Neuro2a cells expressing ataxin-7 10Q, and

found that indel repair is more frequent for ataxin-7 92Q expressing cells, whereas translocation repair is much less frequent for ataxin-7 92Q expressing cells (Figure 4B). These results are consistent with the repair reporter construct studies, which indicated that NHEJ is not affected by polyglutamine-expanded ataxin-7, but HDR and SSA, which are established pathways for creating DNA translocations (Bhargava et al., 2016), are impaired by polyglutamine-expanded ataxin-7 (Figure 3). Translocation events were well distributed throughout the genome for Neuro2a cells expressing either ataxin-7 10Q or ataxin-7 92Q (Figure 4C), as expected from repairing random double-strand DNA breaks. However, when we evaluated the nature of DNA sequences involved in distal translocation formation, we detected a marked increase in translocations forming with exon-containing DNA in Neuro2a cells expressing ataxin-7 92Q (Figure 4D). This was not the case for other different types of DNA sequences, such as intron DNA, fragile sites, or genes identified as having altered H3K9ac from the ChIP-seq experiment (Figure 4D). When we analyzed translocation events using a second independent bait site, we observed this same pattern of preferential exon-containing DNA translocations in ataxin-7 92Q expressing Neuro2a cells (Figure S7, Data S3).

DISCUSSION

Decades of neurogenetics research have revealed vulnerability to DNA damage as a recurrent theme in inherited cerebellar ataxia and suggest that cerebellar neurons are exquisitely susceptible to defects in DNA repair. Indeed, ataxia telangiectasia (AT), AT-like disease, ataxia with oculomotor apraxia types 1 and 2, Cockayne syndrome, and spinocerebellar ataxia with axonal neuropathy are all inherited disorders caused by mutations in DNA repair genes (Paulson and Miller, 2005). Although patients affected with some of these diseases also display an increased predisposition to cancer, for all these diseases, cerebellar neurodegeneration is a predominant feature. We recently documented increased DNA damage in neurons cultured *ex vivo* from SCA7 model mice in combination with PARP1 hyperactivation and proposed that SCA7 cerebellar degeneration results from impaired function of sirtuin-1 because of depletion of nuclear nicotinamide adenine dinucleotide (Stoyas et al., 2020). As a thorough understanding of selective vulnerability in SCA7 requires that we examine how the disease mutation alters normal ataxin-7 protein function as a core component of the STAGA co-activator complex (Helmlinger et al., 2004; Martinez et al., 2001; Palhan et al., 2005), here, we pursued unbiased ChIP-seq analysis of H3K9ac in the degenerating cerebellum of SCA7 266Q knockin mice and matched littermate controls and identified more than 700 genes with increased H3K9 promoter acetylation in SCA7 cerebellum. One of the most significant pathways implicated in STAGA-complex-mediated epigenetic dysregulation in the cerebellum of SCA7 mice was DNA repair, suggesting that altered transcription regulation of DNA repair gene expression may underlie a potential propensity to DNA damage in SCA7 or may represent an attempt to compensate for increased DNA damage in SCA7 cerebellar degeneration.

To determine whether a predisposition to DNA damage exists in SCA7 and how extensive the predilection to DNA damage

might be in SCA7, we surveyed a number of different SCA7 model systems, including HEK293T cells transiently expressing mutant ataxin-7, inducible PC12 cells stably expressing mutant ataxin-7, primary cerebellar granule neurons from SCA7 knockin mice, and progenitor neurons derived from the stem cells of patients with SCA7. In all cases, we observed a marked increase in DNA damage in SCA7 cells, in agreement with a recent study of SCA7, using the inducible PC12 cell culture model (Niss et al., 2021). To define specific DNA repair processes that are impaired in SCA7, we performed direct functional assays of DNA repair and observed ataxin-7 polyglutamine length-dependent inhibition of two pathways: HDR and SSA. Because these experiments used reporter plasmids that are extrachromosomal, we sought a method for detection of break repair at endogenous DNA in native chromosome context and, thus, selected LAM-HTGTS, a powerful high-throughput next-generation sequencing technique developed for monitoring of DNA double-strand break formation (Hu et al., 2016). However, we modified the LAM-HTGTS protocol by using CRISPR-Cas9 to create the double-strand DNA break at a specific site and also added a step with a 5′-methyl-cytosine-modified primer to promote *LpnPI* endonuclease cleavage of sealed breaks to enrich for translocation events. Nonetheless, our results clearly illustrate that canonical NHEJ repair is highly efficient at ligating broken double-strand breaks because we observed a 2- to 5-fold greater rate of indel repair reads than translocation reads upon LAM-HTGTS, even though we had modified the protocol to enrich for detection of translocation events. Our unbiased, native-chromosome DNA-repair experimentation revealed that expression of polyglutamine-expanded ataxin-7 yielded greatly reduced translocations in comparison with normal ataxin-7, which is consistent with retained canonical NHEJ repair, decreased HDR activity, and decreased SSA repair in SCA7 cells because the classical NHEJ pathway is known to prevent translocation by ligating broken double-strand breaks (Roukos and Misteli, 2014). However, both HDR and SSA can resolve unsealed ends by promoting distal translocations, when the lack of NHEJ repair permits end resection to occur at exposed double-strand breaks (Bhargava et al., 2016); hence, markedly decreased translocation events in Neuro2a cells expressing polyglutamine-expanded ataxin-7 likely reflects a combination of impaired HDR and SSA activity with retained NHEJ repair, as observed in the direct DNA repair reporter assays. Taken together, these findings indicate that alterations in DNA repair pathway function in SCA7 neurons in the cerebellum may predispose the subject to excessive DNA damage, initiating a cascade of events culminating in neuron demise and, thus, highlighting DNA repair as an appealing target for therapy development for SCA7 and related ataxias.

One intriguing finding from the modified LAM-HTGTS study of DNA break repair of endogenous double-strand breaks was the marked tendency for translocations involving exons in Neuro2a cells expressing polyglutamine-expanded ataxin-7. What might account for this increased likelihood of translocations into exons in cells expressing polyglutamine-expanded ataxin-7? One possibility would be faulty transcription-coupled repair. PNKP has been shown to promote transcription-coupled double-strand break repair (Chakraborty et al., 2016) and often performs transcription-coupled repair when in complex with huntingtin

CellPress
OPEN ACCESS

protein, RNA polymerase II subunit A, ataxin-3, and cyclic AMP-response element binding protein (CBP) (Gao et al., 2019). We detected increased H3K9ac of the PNKP gene promoter in the cerebellum of SCA7 266Q knockin mice and documented markedly increased expression of the *Pnkp* gene; hence, polyglutamine-expanded ataxin-7 incorporation into the STAGA complex could result in transcriptional dysregulation of PNKP, affecting transcription-coupled repair. Because the polyglutamine-expanded huntingtin protein can impair PNKP function, disrupting transcription-coupled repair (Gao et al., 2019), it is also possible that polyglutamine-expanded ataxin-7, which accumulates in the nucleus, could inappropriately interact with members of the nuclear transcription-coupled repair complex and similarly interfere with its function, akin to aggregated polyglutamine-expanded ataxin-7 alteration of FUS localization and function (Niss et al., 2021).

Limitations of the study

Defects in DNA repair are emerging as likely contributors to polyglutamine neurodegeneration in HD and various SCAs because a number of single-nucleotide polymorphisms (SNPs) in DNA repair genes have been associated with age of disease onset in patients with HD and SCA (Bettencourt et al., 2016), and various reports have documented that polyglutamine-expanded disease proteins can impair DNA repair pathways (Fujita et al., 2013; Gao et al., 2019; Maiuri et al., 2019; Niss et al., 2021). In the case of SCA7, here, we found that altered normal function of ataxin-7 in the STAGA co-activator complex might be responsible for a widespread dysregulation of H3K9 acetylation across >35 genes involved in DNA repair, resulting in a significantly altered canonical DNA repair pathway interactome in the cerebellum. Our findings highlight a central role for altered DNA repair in SCA7 disease pathogenesis but cannot distinguish between two different models. In one scenario, polyglutamine-expanded ataxin-7 incorporation into the STAGA complex is the direct cause of altered GCN5-dependent histone acetyltransferase activity, as we and others have reported previously (Helmlinger et al., 2006; McMahon et al., 2005; Palhan et al., 2005), yielding markedly increased H3K9 promoter acetylation in DNA repair genes, resulting in increased expression of key DNA repair genes. Another possible explanation is that impaired DNA damage repair in SCA7 cerebellar disease promotes altered epigenetic regulation and increased transcription of DNA repair genes is, consequently, a compensatory response. Differentiating between these two models will be the focus of future work. Because protein levels often correlate with increases in mRNA dosage (Ishikawa et al., 2017; Zhang et al., 2016), we hypothesize that the stoichiometry of protein complexes involved in DNA repair are disrupted in SCA7, and resulting inappropriate repair protein component ratios inhibit HDR function (Krejci et al., 2012). However, even if that hypothesis is correct, we cannot explain why the HDR and SSA pathways function aberrantly in SCA7, underscoring the need for continued research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- ● KEY RESOURCES TABLE
- ● RESOURCE AVAILABILITY
  - ○ Lead contact
  - ○ Materials availability
  - ○ Data and code availability
- ● EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - ○ Cell lines
  - ○ Animals
- ● METHOD DETAILS
  - ○ Immunoblotting
  - ○ Chromatin immunoprecipitation and sequencing
  - ○ Pathway analysis
  - ○ Validation by qPCR
  - ○ Flow cytometry
  - ○ Fluorescence microscopy
  - ○ Single molecule DNA repair capture assay
  - ○ Cell and genomic DNA preparation
  - ○ DNA processing and library preparation
  - ○ Data analysis
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

AUTHOR CONTRIBUTIONS

A.R.L.S., P.M.S., and J.R.D. provided the conceptual framework for the study. P.M.S., J.R.D., L.C.B., N.S., B.L.S., and A.R.L.S. designed the experiments. P.M.S., J.R.D., L.C.B., C.N., M.R.-Z., A.A., E.N.G., J.B., B.L.S., and A.R.L.S. performed the experiments. A.R.L.S., P.M.S., and J.R.D. wrote the manuscript.

DECLARATIONS OF INTERESTS

The authors declare no competing interests.

REFERENCES

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46 (W1)*, W537–W544.

Bártová, E., Krejcí, J., Harnicarová, A., Galiová, G., and Kozubek, S. (2008). Histone modifications and nuclear architecture: a review. J. Histochem. Cytochem. *56*, 711–721.

Bennardo, N., Cheng, A., Huang, N., and Stark, J.M. (2008). Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. PLoS Genet. *4*, e1000110.

Bettencourt, C., Hensman-Moss, D., Flower, M., Wiethoff, S., Brice, A., Goizet, C., Stevanin, G., Koutsis, G., Karadima, G., Panas, M., et al.; SPATAX Network (2016). DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. Ann. Neurol. *79*, 983–990.

Bhargava, R., Onyango, D.O., and Stark, J.M. (2016). Regulation of single-strand annealing and its role in genome maintenance. Trends Genet. *32*, 566–575.

Chakraborty, A., Tapryal, N., Venkova, T., Horikoshi, N., Pandita, R.K., Sarker, A.H., Sarkar, P.S., Pandita, T.K., and Hazra, T.K. (2016). Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes. Nat. Commun. *7*, 13049.

Chatr-Aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N.K., O'Donnell, L., Oster, S., Theesfeld, C., Sellam, A., et al. (2017). The Bio-GRID interaction database: 2017 update. Nucleic Acids Res. *45* (*D1*), D369–D379.

Chen, S., Peng, G.H., Wang, X., Smith, A.C., Grote, S.K., Sopher, B.L., and La Spada, A.R. (2004). Interference of Crx-dependent transcription by ataxin-7 involves interaction between the glutamine regions and requires the ataxin-7 carboxy-terminal region for nuclear localization. Hum. Mol. Genet. *13*, 53–67.

d'Adda di Fagagna, F., Reaper, P.M., Clay-Farrace, L., Fiegler, H., Carr, P., Von Zglinicki, T., Saretzki, G., Carter, N.P., and Jackson, S.P. (2003). A DNA damage checkpoint response in telomere-initiated senescence. Nature *426*, 194–198.

David, G., Abbas, N., Stevanin, G., Dürr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., et al. (1997). Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. Nat. Genet. *17*, 65–70.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat. Protoc. *7*, 1728–1740.

Fernandez-Capetillo, O., Celeste, A., and Nussenzweig, A. (2003). Focusing on foci: H2AX and the recruitment of DNA-damage response factors. Cell Cycle *2*, 426–427.

Fujita, K., Nakamura, Y., Oka, T., Ito, H., Tamura, T., Tagawa, K., Sasabe, T., Katsuta, A., Motoki, K., Shiwaku, H., et al. (2013). A functional deficiency of TERA/VCP/p97 contributes to impaired DNA repair in multiple polyglutamine diseases. Nat. Commun. *4*, 1816.

Gao, R., Chakraborty, A., Geater, C., Pradhan, S., Gordon, K.L., Snowden, J., Yuan, S., Dickey, A.S., Choudhary, S., Ashizawa, T., et al. (2019). Mutant huntingtin impairs PNKP and ATXN3, disrupting DNA repair and transcription. eLife *8*, e42988.

Garden, G.A., and La Spada, A.R. (2008). Molecular pathogenesis and cellular pathology of spinocerebellar ataxia type 7 neurodegeneration. Cerebellum *7*, 138–149.

Gatchel, J.R., Watase, K., Thaller, C., Carson, J.P., Jafar-Nejad, P., Shaw, C., Zu, T., Orr, H.T., and Zoghbi, H.Y. (2008). The insulin-like growth factor pathway is altered in spinocerebellar ataxia type 1 and type 7. Proc. Natl. Acad. Sci. USA *105*, 1291–1296.

Ge, S.X., Jung, D., and Yao, R. (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. Bioinformatics *36*, 2628–2629.

Helmlinger, D., Hardy, S., Sasorith, S., Klein, F., Robert, F., Weber, C., Miguet, L., Potier, N., Van-Dorsselaer, A., Wurtz, J.M., et al. (2004). Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. Hum. Mol. Genet. *13*, 1257–1265.

Helmlinger, D., Hardy, S., Abou-Sleymane, G., Eberlin, A., Bowman, A.B., Gansmüller, A., Picaud, S., Zoghbi, H.Y., Trottier, Y., Tora, L., and Devys, D. (2006). Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. PLoS Biol. *4*, e67.

Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W., and Frock, R.L. (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. Nat. Protoc. *11*, 853–871.

Ishikawa, K., Makanae, K., Iwasaki, S., Ingolia, N.T., and Moriya, H. (2017). Post-translational dosage compensation buffers genetic perturbations to stoichiometry of protein complexes. PLoS Genet. *13*, e1006554.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Krejci, L., Altmannova, V., Spirek, M., and Zhao, X. (2012). Homologous recombination and its regulation. Nucleic Acids Res. *40*, 5795–5818.

Kuo, L.J., and Yang, L.X. (2008). Gamma-H2AX—A novel biomarker for DNA double-strand breaks. In Vivo *22*, 305–309.

La Spada, A.R., Fu, Y.H., Sopher, B.L., Libby, R.T., Wang, X., Li, L.Y., Einum, D.D., Huang, J., Possin, D.E., Smith, A.C., et al. (2001). Polyglutamine-expanded ataxin-7 antagonizes CRX function and induces cone-rod dystrophy in a mouse model of SCA7. Neuron *31*, 913–927.

Maiuri, T., Suart, C.E., Hung, C.L.K., Graham, K.J., Barba Bazan, C.A., and Truant, R. (2019). DNA damage repair in huntington's disease and other neurodegenerative diseases. Neurotherapeutics *16*, 948–956.

Martinez, E., Kundu, T.K., Fu, J., and Roeder, R.G. (1998). A human SPT3-TA-FII31-GCN5-L acetylase complex distinct from transcription factor IID. J. Biol. Chem. *273*, 23781–23785.

Martinez, E., Palhan, V.B., Tjernberg, A., Lymar, E.S., Gamper, A.M., Kundu, T.K., Chait, B.T., and Roeder, R.G. (2001). Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. Mol. Cell. Biol. *21*, 6782–6795.

McMahon, S.J., Pray-Grant, M.G., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. Proc. Natl. Acad. Sci. USA *102*, 8478–8482.

Niss, F., Zaidi, W., Hallberg, E., and Ström, A.L. (2021). Polyglutamine expanded Ataxin-7 induces DNA damage and alters FUS localization and function. Mol. Cell. Neurosci. *110*, 103584.

O'Geen, H., Echipare, L., and Farnham, P.J. (2011). Using ChIP-seq technology to generate high-resolution profiles of histone modifications. Methods Mol. Biol. *791*, 265–286.

Palhan, V.B., Chen, S., Peng, G.H., Tjernberg, A., Gamper, A.M., Fan, Y., Chait, B.T., La Spada, A.R., and Roeder, R.G. (2005). Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. Proc. Natl. Acad. Sci. USA *102*, 8472–8477.

Paulson, H.L., and Miller, V.M. (2005). Breaks in coordination: DNA repair in inherited ataxia. Neuron *46*, 845–848.

Paulson, H.L., Bonini, N.M., and Roth, K.A. (2000). Polyglutamine disease and neuronal cell death. Proc. Natl. Acad. Sci. USA *97*, 12957–12958.

Pierce, A.J., Johnson, R.D., Thompson, L.H., and Jasin, M. (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev. *13*, 2633–2638.

Richardson, C., Moynahan, M.E., and Jasin, M. (1998). Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. Genes Dev. *12*, 3831–3842.

Ross, C.A. (1997). Intranuclear neuronal inclusions: a common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? Neuron *19*, 1147–1150.

Roukos, V., and Misteli, T. (2014). The biogenesis of chromosome translocations. Nat. Cell Biol. *16*, 293–300.

Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M.B. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. Nat. Biotechnol. *27*, 66–75.

Sanders, S.L., Jennings, J., Canutescu, A., Link, A.J., and Weil, P.A. (2002). Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. Mol. Cell. Biol. *22*, 4723–4738.

Schindelin, Johannes, Arganda-Carreras, Ignacio, Frise, Erwin, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A., and de Lange, T. (2002). DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. Curr. Biol. *12*, 1635–1644.

Stark, J.M., Pierce, A.J., Oh, J., Pastink, A., and Jasin, M. (2004). Genetic steps of mammalian homologous repair with distinct mutagenic consequences. Mol. Cell. Biol. *24*, 9305–9316.

Stevanin, G., Dürr, A., and Brice, A. (2000). Clinical and molecular advances in autosomal dominant cerebellar ataxias: from genotype to phenotype and physiopathology. Eur. J. Hum. Genet. *8*, 4–18.

Stimpson, K.M., Song, I.Y., Jauch, A., Holtgreve-Grez, H., Hayden, K.E., Bridger, J.M., and Sullivan, B.A. (2010). Telomere disruption results in non-random formation of de novo dicentric chromosomes involving acrocentric human chromosomes. PLoS Genet. *6*, e1001061.

Stoyas, C.A., Bushart, D.D., Switonski, P.M., Ward, J.M., Alaghatta, A., Tang, M.B., Niu, C., Wadhwa, M., Huang, H., Savchenko, A., et al. (2020). Nicotinamide pathway-dependent sirt1 activation restores calcium homeostasis to achieve neuroprotection in spinocerebellar ataxia type 7. Neuron *105*, 630–644.e9.

Terasawa, M., Shinohara, A., and Shinohara, M. (2014). Canonical non-homologous end joining in mitosis induces genome instability and is suppressed by M-phase-specific phosphorylation of XRCC4. PLoS Genet. *10*, e1004563.

Ward, J.M., Stoyas, C.A., Switonski, P.M., Ichou, F., Fan, W., Collins, B., Wall, C.E., Adanyeguh, I., Niu, C., Sopher, B.L., et al. (2019). Metabolic and organelle morphology defects in mice and human patients define spinocerebellar ataxia type 7 as a mitochondrial disease. Cell Rep. *26*, 1189–1202.e6.

Wei, P.C., Lee, C.S., Du, Z., Schwer, B., Zhang, Y., Kao, J., Zurita, J., and Alt, F.W. (2018). Three classes of recurrent DNA break clusters in brain progenitors identified by 3D proximity-based break joining assay. Proc. Natl. Acad. Sci. USA *115*, 1919–1924.

Yoo, S.Y., Pennesi, M.E., Weeber, E.J., Xu, B., Atkinson, R., Chen, S., Armstrong, D.L., Wu, S.M., Sweatt, J.D., and Zoghbi, H.Y. (2003). SCA7 knockin mice model human SCA7 and reveal gradual accumulation of mutant ataxin-7 in neurons and abnormalities in short-term plasticity. Neuron *37*, 383–401.

Yu, X., Ajayi, A., Boga, N.R., and Ström, A.L. (2012). Differential degradation of full-length and cleaved ataxin-7 fragments in a novel stable inducible SCA7 model. J. Mol. Neurosci. *47*, 219–233.

Zhang, Yong, Liu, Tao, Meyer, Clifford. A ., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*. https://doi.org/10.1186/gb-2008-9-9-r137.

Zhang, H., Liu, T., Zhang, Z., Payne, S.H., Zhang, B., McDermott, J.E., Zhou, J.Y., Petyuk, V.A., Chen, L., Ray, D., et al.; CPTAC Investigators (2016). Integrated proteogenomic characterization of human high-grade serous ovarian cancer. Cell *166*, 755–765.

Zoghbi, H.Y., and Orr, H.T. (2000). Glutamine repeats and neurodegeneration. Annu. Rev. Neurosci. *23*, 217–247.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-H2A.X Phospho (Ser139) Monoclonal Antibody	BioLegend	Cat#613402; RRID: AB_315795
Rabbit Phospho-Histone H2A.X (Ser139) Monoclonal Antibody	Cell Signaling	Cat#2577; RRID: AB_2118010
Mouse anti- GAPDH Monoclonal Antibody (6C5)	ThermoFisher	Cat#AM4300; RRID: AB_2536381
Mouse anti-beta Actin Monoclonal Antibody	Abcam	Cat#ab8226; RRID: AB_306371
Mouse anti-MAP2 Monoclonal Antibody (AP20)	Merck Millipore	Cat#MAB3418; RRID: AB_94856
Rabbit anti-ATXN7 "L" Antibody	La Spada et al., 2001	N/A
Goat Anti-Rabbit IgG(H+L), Mouse/Human ads-HRP	SouthernBiotech	Cat#4050-05; RRID: AB_2795955
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, HRP	ThermoFisher	Cat#A16078; RRID:AB_2534751
Goat-anti mouse IgG (H+L) Superclonal SecondaryAntibody, Alexa Fluor 488	ThermoFisher	Cat#A28175; RRID:AB_2536161
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed SecondaryAntibody, Alexa Fluor 594	ThermoFisher	Cat#A-11012; RRID: AB_2534079
Rabbit anti-acetyl-Histone H3 (Lys9) Antibody	Merck Millipore	Cat#07-352; RRID: AB_310544
Drosophila Spike-in Antibody	Active Motif	Cat#61686; RRID: AB_2737370
Chemicals, peptides, and recombinant proteins		
HALT protease inhibitor cocktail	ThermoFisher	Cat#78443
Drosophila spike-in chromatin control	Active Motif	Cat#53083
Ribonuclease A from bovine pancreas	Merck Millipore	Cat#R6513
Proteinase K	NEB	Cat#P8107S
TURBO DNase	ThermoFisher	Cat#AM2238
Hoescht 33342	ThermoFisher	Cat#H3570
VectaShield Antifade Mounting Medium	Vector Laboratories	Cat#H-1000
Lipofectamine2000	ThermoFisher	Cat#11668500
T4 DNA ligase	NEB	Cat#M0202S
LpnPI enzyme	NEB	Cat#R0663S
Phusion DNA polymerase	ThermoFisher	Cat#F530L
Hexamine cobalt III chloride	Sigma	Cat#481521
Horse serum	ThermoFisher	Cat#16050-122
Tet System Approved FBS	Takara	Cat#631106
STEMdiff Neural Progenitor Medium	StemCell Technlogies	Cat#05833
Critical commercial assays		
Micro BCA Protein Assay Kit	ThermoFisher	Cat#23235
Dynabeads Protein G for Immunoprecipitation	ThermoFisher	Cat#10004D
Zymo ChIP DNA clean and concentrator kit	Zymo	#D5201
Direct-zol RNA Miniprep kit	Zymo	Cat#R2050

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SuperScript IV VILO Master Mix	ThermoFisher	Cat#11756050
SuperScript III kit	ThermoFisher	Cat#18080051
Power SYBR Green PCR master mix	Life Technologies	Cat#4367659
C1 Streptavidin Dynabeads	ThermoFisher	Cat#11205D
PureLink PCR Purification Kit	ThermoFisher	Cat#K310002
Blood and Tissue DNA Kit	QIAGEN	Cat#69506
KAPA HyperPrep DNA Library Preparation Kit	Roche	Cat#KK8502
Deposited data		
Raw and processed ChIP-seq data	This paper	GEO: GSE166118
Raw and processed LAM-HTGTS data	This paper	GEO: GSE166119
Experimental models: Cell lines		
Human IPSC-derived neuron progenitor cells	Ward et al., 2019	N/A
Human HEK293T cell line	ATCC	Cat#CRL-3216; RRID: CVCL_0063
Mouse Neuro2a cell line	ATCC	Cat#CCL-131; RRID: CVCL_0470
Rat PC12 Tet-off Atxn7 cells	Yu et al., 2012	N/A
Experimental models: Organisms/strains		
Mouse: SCA7 266Q mice	Yoo et al., 2003	N/A
Oligonucleotides		
Biotinylated bait primer site #1 5′ /5Biosg/ ACTAAAAAGAGCGCCCCTCTTTC 3′	IDT	N/A
Biotinylated bait primer site #2 5′ /5Biosg/ GGTTTCCCCTCCGCTCTTTA 3′	IDT	N/A
5-Cytosine methylated primer site #1 5′ TTTGGGAAGAGC/iMe-dC/ TGAGCATTCTGAG 3′	IDT	N/A
5-Cytosine methylated primer site #2 5′ TCAACCCCGTCCC/iMe-dC/ AGTGGAAAACACCC 3′	IDT	N/A
qPCR primers	Table S2	N/A
Recombinant DNA		
Plasmid expressing sgRNA to the target site 1 on Chromosome 14 (5′ GGTAGACATCAGCGCTGTTG 3′)	This paper	N/A
Plasmid expressing sgRNA to the target site 2 on Chromosome 14 (5′ CTCGCAGCCCTCCACCGTAA 3′)	This paper	N/A
Software and algorithms		
GraphPad Prism 6.0	GraphPad	https://www.graphpad.com/scientific-software/prism/ RRID: SCR_002798
Fiji distribution of ImageJ	(Schindelin et al., 2012)	https://fiji.sc/#download, RRID: SCR_002285
FlowJo	FlowJo, LLC	https://www.flowjo.com/ RRID: SCR_008520
Galaxy	Afgan et al., 2018	https://usegalaxy.org/ RRID: SCR_006281
MACS V2.1.1	(Zhang et al., 2008)	RRID:SCR_013291
GOrilla	Eden et al., 2009	http://cbl-gorilla.cs.technion.ac.il/ RRID: SCR_006848

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ShinyGO	Ge et al., 2020	http://bioinformatics.sdstate.edu/go/ RRID: SCR_019213
R	R Project for Statistical Computing	https://www.r-project.org/, RRID: SCR_001905
RStudio	RStudio	https://www.rstudio.com/, RRID: SCR_000432
Custom mm10_MACS_Interval_to_TSS script	Github; supplemental information	https://github.com/jrdelaney/ Published_Code/blob/main/mm10_ MACS_Interval_to_TSS_ CellReports2021.R

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Albert La Spada (alaspada@hs.uci.edu).

Materials availability
This study did not generate any new viral constructs or mouse lines, but it did produce a new set of plasmids expressing sgRNA to a designated target site on Chromosome 14, which are available upon request after completion of a material transfer agreement.

Data and code availability

- The ChIP-seq dataset obtained from SCA7 mice and WT control littermates has been deposited at GEO and are publicly available as of the date of publication (NCBI GEO: GSE166123).
- The LAM-HTGTS dataset obtained from murine Neuro2a cells transfected with 10Q-ATXN7, 92Q-ATXN7 or control plasmids has been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- All original 'R' code is available as Supplemental information (Data S2), and at the following website: https://github.com/jrdelaney/Published_Code/blob/main/mm10_MACS_Interval_to_TSS_CellReports2021.R
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines
Established HEK293T and Neuro2A cell lines were purchased from the American Type Culture Collection (ATCC) and validated by short tandem repeat (STR) profiling. Routine microscopic morphology tests were performed prior to each experiment. All cells were grown in antibiotics (penicillin - streptomycin, ThermoFisher, cat#15140163) and 10% FBS (HyClone, cat# SH30071.03HI). Cells were cultured at 37°C with 5% CO_2. Transfections in HEK293T cells were performed for 48 hours prior to molecular analysis.

Rat PC12 Tet-off cells express either N-terminal FLAG and C-terminal GFP-tagged full-length ataxin-7 with 10Qs or N-terminal FLAG and C-terminal GFP-tagged full-length ataxin-7 with 65Qs (Yu et al., 2012). Cells were cultured in DMEM (ThermoFisher, cat# 11965092) with 10% horse serum (ThermoFisher, cat# 16050-122) and 5% Tet System Approved FBS (Takara, cat# 631106). Hygromycin (ThermoFisher, cat# 10687010) and G418 (ThermoFisher, cat# 10131035) were used to maintain incorporation of ataxin-7 and tetracycline-regulated transactivator constructs respectively. 1 μg/ml Doxycycline (Sigma, cat# D9891) was used to keep the ataxin-7 expression off. Doxycycline was removed from the media to induce the ataxin-7 expression when desired.

Demographics and generation of NPCs cell line has been extensively documented by Ward et al. (2019). NPCs were originally generated from patient-derived iPSCs. Two families, one consisting of affected son and unaffected mother, and the other consisting of affected daughter and unaffected father donated fibroblast for IPSCs generation. NPCs were cultured in STEMdiff Neural Progenitor Medium (StemCell Technologies, cat# 05833) for 5-15 passages in 5% CO2 at 37°C with medium changes every 2-3 days per manufacturer's instructions.

Mouse primary cultures of granule neurons from SCA7 and wild-type littermates were isolated on postnatal day 6. Individual cerebella were isolated and trypsinized for 10 min at 37°C in HBSS + 3% BSA (ThermoFisher, cat# BP1600100) with 3.75 ug/ml of trypsin (Sigma, cat# T9935-50MG) followed by 10 min incubation with 3.75 ug/ml of trypsin inhibitor (Sigma, cat# T6522-25MG). Granule neurons were dissociated in the dissociating medium containing DMEM + glutamine (ThermoFisher, cat# 11965092), 10% FBS

(HyClone, cat# SH30071.03HI) and 1 mM sodium pyruvate (ThermoFisher, cat# 11360070) and plated on the poly-L-ornithine (Sigma, cat# P3655-100mg)-coated coverslips in Neurobasal A (ThermoFisher, cat# 10888022) medium containing B27 supplement (ThermoFisher, cat# 17504001), 0.5 mM L-glutamine and 20 mM KCl and maintained at 5% CO2 at 37°C with a half volume medium changes every 3 days.

Animals

All animal experimentation adhered to NIH guidelines and was approved by the University of California, San Diego, Duke University and University of California, Irvine Institutional Animal Care and Use Committees. The development of the SCA7 mice has been previously described (Yoo et al., 2003). SCA7 hemizygous mice carrying either 266 or 210 CAG repeats and age-matched control wild-type littermates were used in this study. All SCA7 knock-in mice have been maintained as congenic on the C57BL/6J strain background for > 20 generations were used in the experiments. Numbers and age used in all experiments are reported in the accompanying figure legends, with sex ratios equivalent to or not significantly different from 1:1. In experiments involving both males and females, we did not observe a significant difference in outcome measures based upon sex. Mice were housed in groups of 2-5 per cage under 12 h dark-light cycle with food and water *ad libitum*. All measurements were performed by a blinded examiner.

METHOD DETAILS

Immunoblotting

Cells were grown to 80% confluency on 10 cm plates at 37°C. Media was aspirated, cells washed in PBS, and the cellular monolayer was immersed in iced RIPA buffer supplemented with HALT protease inhibitor cocktail (ThermoFisher, cat#78443). Following 10 minutes on ice, lysates were collected using a cell lifter (Fisher Scientific, cat#08-100-240). Lysates were spun at 10,000 g for 10 minutes at 4°C and supernatant saved and quantified by BCA assay (ThermoFisher, cat#23235). For mouse tissue, tissue was first resuspended in RIPA buffer containing protease inhibitors and disrupted by pipetting through a P1000 pipette tip 20X. Tissue was then sonicated on a Pico Biorupter (Diagenode, USA) for three 5x cycles of 30 s on 30 s off at 4°C. Tissue lysates were similarly centrifuged and quantified by BCA. 30μg protein was loaded per well of a 4%–12% SDS-PAGE gel (Invitrogen, cat# NW04125BOX) and transferred onto PVDF membrane (Abcam, cat# ab133411-101MEMB). The membrane was blocked in 5% dry milk (Genesee, cat#20-241). Antibodies used were as follows: γH2AX (BioLegend, cat# 613402), γH2AX (Cell Signaling, cat# 2577), GAPDH (ThermoFisher, cat# AM4300), ATXN7 "L" (La Spada et al., 2001), and β-actin (Abcam, cat# ab8226). All primary antibodies were used at 1:1,000 dilution. Secondary HRP antibodies were anti-rabbit (Southern Biotech, cat# 4050-05) or anti-mouse (ThermoFisher, cat# A16078), and were used at 1:5,000 dilution. Quantitation of band intensity was performed in ImageJ and all normalizations were to the loading control displayed in the corresponding figure.

Chromatin immunoprecipitation and sequencing
Sample preparation

Cerebella were dissected from 9 week-old mice. While n = 2 is sufficient for ChIP-Seq determination of most peaks (Rozowsky et al., 2009), we doubled this to n = 4 per group. Tissue was rinsed quickly with iced PBS and snap frozen in liquid nitrogen. Tissue was fixed in 1 mL of 1% formaldehyde in PBS for 10 min rotating at room temperature, using a P1000 tip to initially disperse tissue. Fixation was quenched with 111 μL 1.25M glycine for 2 min at room temperature (RT). Cells were washed twice in iced PBS, centrifugation was performed at 3,000 g for 1 min at 4°C. All the buffers in the following steps contained 1X HALT protease inhibitor cocktail (ThermoFisher, cat#78443). Cells were spun at 3,000 g for 1 min at 4°C with 1 mL nuclei isolation buffer (50mM Tris-HCl pH8, 60mM KCl, 0.5% NP-40) added. Tissue was homogenized by pipetting up-and-down with a P1000 pipette tip, 20 times. Tissue was then rotated at 4°C for 10 min. Nuclei were centrifuged at 2,000 g for 3 min at 4°C. Supernatant was discarded and nuclei resuspended in 200 μL lysis buffer (50mM Tris-HCl pH8, 10mM EDTA pH8, 0.5mM EGTA pH8, 0.5% SDS). Suspension was sonicated on a Pico Biorupter for three 8x cycles of 30 s on, 90 s off at 4°C, with a 15 min cooling break between 8x cycles. Chromatin was then diluted in 800 μL dilution buffer (20mM Tris-HCl pH8, 2mM EDTA pH8, 150mM NaCl, 1% Triton X-100).

Chromatin immunoprecipitation (ChIP)

60 μL protein G dynabeads (ThermoFisher, cat#10004D) were used per ChIP reaction. Bead buffer was removed and beads were washed in bead wash buffer (0.1% bovine serum albumin (BSA) in PBS) twice. Wash buffer was removed and beads resuspended in bead blocking buffer (0.5% BSA, 1 μg/μl glycogen, 0.01% sodium azide, in PBS). Beads were rotated at RT for 1 hr. Beads were washed twice in bead wash buffer and then resuspended in 1:1.5 lysis:dilution buffers, prior to obtaining blocking bead aliquots (40 μL / ChIP reaction). In the remaining beads, 3 μg primary antibody targeting H3K9ac (Millipore, cat# 07-352) was added and 0.5 μg spike-in control *Drosophila* antibody (Active Motif, cat# 61686) was added. Antibodies were incubated with beads overnight at 4°C. Blocking bead supernatant was removed and diluted chromatin added. 5 ng of *Drosophila* spike-in chromatin control (Active Motif, cat# 53083) was added per ChIP reaction. Chromatin was cleared by rotating at 4°C for 1 hr. 10 μL input per ChIP was aliquoted for sequencing control and validation controls. Input was then housed at identical temperature conditions as the ChIP chromatin, except no beads or wash steps were used. The remaining chromatin was added to 20 μL antibody-bound beads. Beads were rotated at 4°C overnight. Beads were then washed twice with 500 μL wash buffer 1 (20mM Tris-HCl, 150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), twice with wash buffer 2 (20mM Tris-HCl, 500mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), twice with wash

buffer 3 (20mM Tris-HCl, 250mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1mM EDTA), twice with TE+0.1% Triton X-100, twice with TE+50mM NaCl, and twice with TE (all done at RT), and then eluted with 200 μL elution buffer (0.1M sodium bicarbonate, 1% SDS). 90 μL elution buffer was added to input samples. Beads were eluted by rotation at RT for 20 min, and then finally, eluted, immunoprecipitated chromatin was removed. 10 μL 5M NaCl was added to remove DNA crosslinks and incubated for 16 hr at 65°C. Samples were cooled to RT. 7 μL of 10mg/ml RNase (Millipore, cat# R6513) was added and incubated at 37°C for 1 hr. 7 μL proteinase K (20mg/ml, NEB, cat# P8107S) was added and incubated at 45°C for 1 hr, then inactivated by a 3 min incubation at 95°C. Samples were purified using the Zymo ChIP DNA clean and concentrator kit (Zymo, cat# D5201) following the manufacturer's protocol.

Library preparation and sequencing
Library was prepared using standard UCSD Genomics core protocols to generate single ended TruSeq adapters using KAPA Hyper-Prep DNA Library Preparation Kit (Roche, cat# KK8502). Library PCR amplification was performed at 15 cycles and agarose gel size selection performed at the 300-600 bp range. Samples were analyzed by Agilent Tapestation to determine concentrations and loaded into the IGM Core's Illumina HiSeq 4000.

Determination of peaks and post-processing
Sequencing FASTQ files were uploaded into a public Galaxy instance at https://usegalaxy.org (Afgan et al., 2018). A new workflow was created, incorporating the published Model-based Analysis of ChIP-Seq (MACS) algorithm to call peaks (Feng et al., 2012). Specifically, for each ChIP, the paired Input and ChIP FASTQ files were input to FASTQ Groomer, FASTQ Quality Trimmer (quality score > 30), mapped to mm10 using Bowtie2, and aligned reads were filtered for duplicates by RmDup. MACS was then run with these paired files using an MFOLD high confidence threshold of 16, an estimate band size of 300 bp, and a peak cutoff threshold of $p < 1e-05$. Output BED peak tracks were uploaded as custom tracks into the UCSC genome browser and displayed. To determine differentially marked genes, a custom R script (Data S2) was written to map BED peaks as overlapping or within a genomic region corresponding to up to 0.25 kb upstream of transcription start sites (TSSs) and up to 1 kb downstream of TSSs. TSS location data for mm10 was downloaded from the UCSC genome browser. Peaks that were called on a gene were considered differential peaks if at least two animals from the same genotype were observed to have the peak and zero animals from the opposite genotype were observed to have the peak.

Pathway analysis
Gene sets from differentially marked genes were input into the GOrilla tool (Eden et al., 2009) and into ShinyGO v0.66 (Ge et al., 2020), and analyzed for GO term enrichment relative to the background of TSS marked genes, amenable to the peak-calling script.

Validation by qPCR
RNA was isolated from the whole cerebellum of animals using Direct-zol RNA kit (Zymo, cat# R2050), and treated with DNase I in the form of TURBO-DNase (ThermoFisher, cat# AM2238) to remove traces of genomic DNA. Reverse transcription was performed with SuperScript IV VILO Reverse Transcriptase (ThermoFisher, cat# 11756050). Primers corresponding to significant peaks were designed using the NCBI primer blast tool using peak segment regions as the query DNA sequence. Corresponding cDNA from Ensembl was input into the NCBI primer blast tool for reverse transcriptase quantitative real time PCR primer design. ChIP'd DNA and control input DNA was used from all mice on which ChIP-seq was performed, and quantitative real time PCR was performed using input DNA as a control, with an additional correction for ChIP conditions, using the *Drosophila* chromatin ChIP normalization kit as per manufacturer's instructions (Active Motif, cat# 53083). For reverse transcriptase quantitative real time PCR, the loading control used for ddCt calculations was *Tubg1*. Superscript III kits (ThermoFisher, cat# 18080051) were used as per manufacturer's protocol to produce cDNA from an independent cohort of 9 week-old SCA7 mice and wild-type littermate control mice (n = 3 / group). Power SYBR Green PCR master mix (Life Technologies, cat# 4367659) was used to perform real-time PCR on an Applied Biosystems 7500 real time PCR machine. Specific primer sequences are available in Table S2.

Flow cytometry
HEK293T cells containing constructs for DNA repair templates were first generated by stable selection in puromycin with individual repair templates containing an I-*SceI* meganuclease target site (HR [Pierce et al., 1999], NHEJ [Bennardo et al., 2008], SSA [Stark et al., 2004], Alt-EJ [Bennardo et al., 2008]) to reduce variability between experiments. HEK293T cells were seeded to 100K cells / 24-well tissue culture well. After 24 hr to allow to adhere, media was removed and cells were transfected with 100ng CMV control, 10Q-ATXN7-Myc, or 92Q-ATXN7-Myc with 2 μL Lipofectamine2000 (ThermoFisher, cat#11668500) in 500 μL OptiMEM. In experimental wells, 100 ng I-*SceI* expressing plasmids (Richardson et al., 1998) were also co-transfected, whereas in control wells to measure background rates of spontaneous repair without I-*SceI* induced dsDNA breaks, no additional plasmid was transfected. 16 hours post transfection, media was replaced with antibiotic containing DMEM. 40 hr post transfection, cells were trypsinized in 500 μl for 5 min and then pooled with 400 μl iced media. Cells were centrifuged 3,000 g for 1min at 4°C, and supernatant removed. Cells were resuspended in 500 μl iced PBS, transferred to flow cytometry tubes on ice, and read on the flow cytometer within 1-2 hours post-trypsinization. The red / 594 nm channel on the flow cytometer was used to control for autofluorescence. GFP+ / Red- cells were quantified and then background repair was subtracted from non-I-*SceI* transfected controls. Ratios of GFP+ cells of 92Q- and 10Q- ATXN7 were compared in six independently transfected experiments.

Fluorescence microscopy

Cells were seeded at 25,000 cells per well in a Nunc Lab-Tek II Chambered Coverglass, 8-well chamber slide (Thermo Scientific, cat# 155409). For anaphase bridge assays, cells were grown for two days on the chamber slide, media was aspirated, 500 μL of PBS slowly added and then aspirated, and 4% paraformaldehyde added for 15 min at RT. Cells were then washed once with PBS and stained for 30 min with Hoescht 33342 (ThermoFisher, cat# H3570) (1μg/mL in PBS). Cells were washed once in PBS and mounted in 20 μL VectaShield Antifade Mounting Medium (Vector Laboratories, cat#H-1000). Slides were imaged on a Nikon Eclipse 80i at 40X and scored for anaphase bridges by a blinded observer. For antibody staining, cells were grown and processed similarly, but additionally permeabilized for 2 min with 0.1% Triton X-100 in PBS, blocked with 5% BSA / 5% goat serum for 30 min, stained with primary antibodies for 1 hour at RT and 16 hours at 4°C, washed 3x 5 min in PBS, stained with secondary Alexa Fluor antibodies (ThermoFisher, cat# A-11012 and A28175) and Hoescht 33342 (ThermoFisher, cat# Cat#H3570), and washed 3x 5 min and 1x 30 min in PBS prior to mounting with VectaShield.

Single molecule DNA repair capture assay

To capture DNA associated with a designated bait dsDNA break site, our protocol was based upon a previously published translocation capture assay (Hu et al., 2016). However, the referenced assay requires an engineered endonuclease site downstream of the primary dsDNA break site, so the protocol was modified to include endonuclease cleavage by *LpnPI*, which can target many sites as long as a methylated primer can be provided to create a heterodimer.

Cell and genomic DNA preparation

For ataxin-7 transfection experiments, 5 million murine Neuro2a cells were transfected with 10Q-ATXN7 or 92Q-ATXN7 and co-transfected with a plasmid expressing sgRNA to a designated target site on Chromosome 14 (site 1: 5′ GGTAGACATCAGCGCTGTTG 3′, site 2: 5′ CTCGCAGCCCTCCACCGTAA 3′), along with CRISPR-Cas9 and a constitutive GFP reporter. For the quality control experiment, only the CRISPR-Cas9 plasmid co-expressing site 2 targeted sgRNA or non-targeted sgRNA were used. 48 hours after transfection, cells were trypsinized for 5 min, rinsed in PBS, and three replicates were pooled. 1 million GFP positive cells per pooled condition were sorted by flow cytometry. Cells were then resuspended in 200 μL PBS and 20 μL NEB Proteinase K (800U/mL) and incubated for 30 min at 56°C. Proteinase K was inactivated by heating at 95°C for 5 min and then 5 μl RNaseA (8 mg/ml) was added and incubated for 30 min at 37°C. Treated lysate was purified using the Blood and Tissue DNA Kit (QIAGEN, cat# 69506) using the manufacturer's protocol, with the final elution performed twice with 100 μl AE buffer.

DNA processing and library preparation

A Diagenode Bioruptor Pico was used to shear DNA to 750 bp fragments. Each sample was then bait-amplified using 8x 50μl reactions using Phusion DNA polymerase (ThermoFisher, cat# F530L) standard reactions containing 1.5μl 2.5mM dNTPs, 0.5μl of 1 μM biotinylated bait primers manufactured by IDT (site 1: 5′ /5Biosg/ACTAAAAAGAGCGCCCCTCTTTC 3′, site 2: 5′ /5Biosg/GGTTTCCCCTC CGCTCTTTA 3′), 0.5 μl Phusion polymerase, 25 μl sonicated DNA, 10 μl 5X Phusion buffer, and 12.5 μl water. A thermal cycler was run at 98°C 2 min, steps 2-81 as 95°C 15 s, 58°C 15 s, 72°C 60 s, and step 82 as 72°C for 2 min, with freezing at −20°C immediately after program completion. The Invitrogen PureLink PCR Cleanup Kit (ThermoFisher, cat#K310002) was used to remove excess RNA and biotinylated primer, with buffer 2 conditions and elution in 2x 100μl E1 buffer. 50 μl 5M NaCl and 2.5 μl 0.5M EDTA pH 8 were added to each sample. Biotinylated bait DNA was then further purified using magnetic C1 Streptavidin Dynabeads (ThermoFisher, cat#11205D). 40 μl/ sample beads were first washed in 600 μl 1X B&W buffer (2M NaCl, 10mM Tris-HCl pH 7.4, 1mM EDTA pH 8 in water) twice, and resuspended with the biotinylated-DNA samples. Bead-DNA complexes were rotated for 2 hr at RT and then 16 hr at 4°C. Beads were washed 3x from unbound DNA using 600 μl 1X B&W buffer and once with 1 mL water. Bead-DNA complexes were resuspended in 45 μl water and then the following was added for an on-bead ligation reaction: 10 μl 10X T4 ligase buffer, 5 μl 20mM hexamine cobalt III chloride (Sigma, cat# 481521), 5 μl bridge adapters as per the reference Hu et al. (2016), 5 μl 3U/μl T4 DNA ligase (NEB, cat# M0202S), and 30 μl 50% PEG8000 (Promega, cat#V3011) in water. A thermal cycler was used for the ligation reaction using settings of 1 hr at 25°C, 1 hr at 22°C, a manual pipette resuspension step, 1 hr at 22°C, and 16 hr at 16°C. Ligation product was then washed by first adding 100 μl 2X B&W buffer per sample, captured on magnets, and then washing with 200 μl 1X B&W buffer twice followed by a 200 μl water wash. *LpnPI* digestion was performed on indicated samples (all except non-*LpnPI* controls) using beads resuspended in 23.5 μl water, 3 μl 10X NEB Cutsmart buffer, 0.5 μl of 100 μM 5-Cytosine methylated primer manufactured by IDT (site 1: 5′ TTTGGGAAGAGC/iMe-dC/ TGAGCATTCTGAG 3′, site 2: 5′ TCAACCCCGTCCC/iMe-dC/AGTGGAAAACACCC 3′), 0.5μl *LpnPI* enzyme (NEB, cat#R0663S), 1 μl 30X enzyme activator, at running conditions of 37°C for 4 hr, then 65°C heat inactivation for 20 min. *LpnPI* digested samples were washed with 600 μl B&W buffer 5x, then once with 1 mL water, and resuspended in 200 μl water. For Illumina adaptor inclusion and barcoding, resuspended samples were split into 8 PCR reactions consisting of 10 μl 5X Phusion buffer, 4 μl 2.5mM dNTPs, 2 μl 10μM I5 nested primer, 2 μl 10μM I7 primer, 0.5 μl Phusion polymerase, 25 μl bead-DNA complex, and 6.5 μl water. Reaction was run using 95°C 30 s, 18 cycles of 95°C 15 s, 60°C 30 s, 72°C 1 min, then 72°C 6 min, and ending with 4°C until removed from thermal cycler. PCR products were pooled and removed of beads by extraction of supernatant on magnet followed by centrifugation at 15,000 g for 5 min at RT. Supernatant was transferred to a new tube and purified using the Invitrogen PureLink PCR Cleanup Kit with Buffer 2 conditions and eluted with 60 μl water. Final Illumina binding adapters were added by 4 PCR reactions of P5-I5 and P7-I7 adaptor primers, using the following conditions: 15 μl DNA products, 16.5 μl water, 10 μl 5X Phusion buffer, 4 μl 2.5mM dNTPs, 2 μl of each 10 μM P5-I5 and P7-I7 primer, 0.5 μl Phusion polymerase,

with running conditions of 95°C 3 min, 16 cycles of 95°C 15 s, 62°C 30 s, 72°C 1 min, then 72°C 6 min, and 4°C until removal from thermal cycler. Final purification of library prior to NGS steps was performed by running on a 1X TBE gel containing GelRed stain (VWR, cat# 89139-134), excised in the 350-1,250bp range (directly above primer-dimer band), and extracted using a QIAGEN gel purification kit. Additional purification was performed from the gel purification by a second purification through the Invitrogen PureLink PCR Cleanup Kit with Buffer 2 conditions and eluted with 30 μl water. Library concentration was estimated using a NanoDrop, and library pools were created using equimolar DNA per sample. Finally, libraries were run on an Illumina MiSeq at 300 bp depth (QC reads, P5-I5-end only) or 250 bp depth (experimental reads, P5-I5 end only).

Data analysis

For QC analysis, FASTQ files for each barcode were processed in R using vmatchPattern function in Biostrings. The pattern aligned to was: "AACCGCTTCAGCTTGGTGTAGACGGTGTGCAAGCCCCCCA," which contains the nested primer, 20 bp downstream genomic sequence, and is 63 bp upstream of the CRISPR-Cas9 cleavage site, allowing up to 3 mismatches. To determine indels, we only counted reads that contained the sequence "TGGCCCCTCTCAGATCCACCATTTTGCACTCATTATTCTG" and were not isogenic to the reference sequence in length. This sequence allows for up to 15 bp deletion indels distal to the cut site (up to 63+15bp total indel), since the sequence begins 15 bp distal to the cut site. Reads which were not indels were separately saved, trimmed by 110 bp (the region containing the nested primer up to the CRISPR-Cas9 cut site), and then end-end aligned to mm10 using Bowtie2 in Galaxy (Afgan et al., 2018). Aligned BAM files were converted to BED files using BAM to BED in Galaxy. All reads within 50 kb of the bait Cas9 cleavage site were designated "local" translocations, and all reads on other chromosomes or > 50 kb from the bait site were considered "distal" translocations.

For the 10Q versus 92Q ATXN7 analysis, FASTQ files were similarly processed. For site 1 analysis, the initial vmatchPattern alignment was "GTTGAAGACCTCAGAGGTGATGCTGGCTGCTTTCTTCATC," which contains the nested primer, 20 bp downstream genomic sequence, and is 40 bp upstream of the CRISPR-Cas9 cut site. To determine indels, we counted reads that contained the sequence "TGTCTACCTGTAGCAAGGTGATGCATTTCTACTACACACA" and were not isogenic to the reference sequence in length. As for site 2, this sequence allows for up to 15 bp deletion indels distal to the cut site (up to 40+15bp total indel), since the sequence begins 15 bp distal to the cut site. The remainder of the analysis was identical to the QC analysis with translocations defined as "distal," i.e., > 50 kb away from cut site translocations.

To determine overlap with mm10 annotations BED files for each, corresponding annotation was downloaded from the UCSC genome browser (exons, intron, ChIP-Seq hit genes) or supplemental data files (brain progenitor cell fragile sites (Wei et al., 2018)). Any discrepancy from mm9 and mm10 genomes was first converted using the LiftOver function of the UCSC genome browser. To create a randomized control "expect" rate, BED locations were permuted 1000x over the mm10 chromosomal genome.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was done using Microsoft Excel or Prism 6.0 (GraphPad). Statistical significance was defined at $p < 0.05$. All t tests were two-tailed Student's t tests, and level of significance (alpha) was always set to 0.05. For one-way and two-way analysis of variance (ANOVA), if statistical significance ($p < 0.05$) was achieved, then we performed post hoc analysis, as specified, to account for multiple comparisons. Details of each experiment, including the number of biological replicates and technical replicates, are provided in the legend for each figure.