Published in final edited form as:

Cell. 2023 October 26; 186(22): 4898–4919.e25. doi:10.1016/j.cell.2023.09.008.

Antagonistic roles of canonical and alternative-RPA in disease-associated tandem CAG repeat instability

Terence Gall-Duncan[1,2], **Jennifer Luo**[1,2], **Carla-Marie Jurkovic**[3], **Laura A. Fischer**[4], **Kyota Fujita**[5], **Amit L. Deshmukh**[1], **Rachel J. Harding**[6,7], **Stephanie Tran**[1,2], **Mustafa Mehkary**[1,2], **Vanessa Li**[1,2], **David E. Leib**[8], **Ran Chen**[9], **Hikari Tanaka**[5], **Amanda G. Mason**[10], **Dominique Lévesque**[3], **Mahreen Khan**[1,2], **Mortezaali Razzaghi**[11], **Tanya Prasolava**[1], **Stella Lanni**[1], **Nozomu Sato**[1], **Marie-Christine Caron**[12], **Gagan B. Panigrahi**[1], **Peixiang Wang**[1], **Rachel Lau**[1], **Arturo López Castel**[13], **Jean-Yves Masson**[12], **Lynette Tippett**[14,15], **Clinton Turner**[16], **Maria Spies**[11], **Albert R. La Spada**[17,18,19], **Eric I. Campos**[1,2], **Maurice A. Curtis**[15,20], **François-Michel Boisvert**[3], **Richard L.M. Faull**[15,20], **Beverly L. Davidson**[8], **Masayuki Nakamori**[21], **Hitoshi Okazawa**[5], **Marc S. Wold**[11], **Christopher E. Pearson**[1,6,22,23,*]

[1]Genetics & Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada.

[2]Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.

[3]Immunology and Cell Biology, Université de Sherbrooke, Sherbrooke, Québec, Canada.

[4]Developmental Biology and Center of Regenerative Medicine, Washington University School of Medicine, St. Louis, USA.

[5]Neuropathology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan.

[6]Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada.

[7]Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada.

[8]Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19146.

[9]Pediatrics, Division of Hematology and Oncology, Washington University School of Medicine, St. Louis, MO 63110, USA.

[10]Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.

[*]**Correspondence:** Christopher E. Pearson (cepearson.sickkids@gmail.com).

[22]Senior author

[23]Lead contact

Author Contributions

TGD and VL conducted ddPCR experiments. AGM conducted qRT-PCR. TGD conducted western blots and densitometry quantifications. DEL and BLD provided dissected zQ-175 HD mouse brain. TGD, RL, and PW synthesized repair substrates, TGD and TP generated RPA-deficient functional cell extract, and TGD, JL, and GBP performed *in vitro* repair and binding assays. RJH performed AlphaFold analyses. MR, LAF, RC, MS, and MSW purified RPA/Alt-RPA and conducted oligonucleotide binding assays and FRET assays. CMJ, ST, EIC, and F-MB conducted BioID and statistical analysis, DL conducted mass spectrometry analysis, TGD, MM, and ST conducted analysis of BioID dataset. TGD, VL, and NS conducted fragment length analysis and SL conducted instability index calculations for SCA1 mouse brains. KF, HT, and HO prepared Rpa1-overexpressing SCA1 mouse brain tissues and conducted mouse immunohistochemistry. MN conducted repeat instability analysis in HT1080 cell model. LT, CT, MAC, ARLS, and RLMF provided post-mortem patient brain tissues, anonymized patient information, and sized inherited repeat lengths via PCR. M-CC and J-YM purified FAN1 proteins, ALD performed nuclease assays. TGD, MSW, and CEP conceived of, designed, and coordinated experiments. TGD, JL, KF, LAF, C-MJ, ALD, and GBP prepared figures. TGD, MSW, and CEP wrote the manuscript, ALC, J-YM, DEL, BLD, F-MB, EIC, RLMF, MAC, ARL, and HO provided edits and comments. All authors approved of the final text, data and figures.

Declaration of interests: The authors declare no competing interests.

[11]Biochemistry and Molecular Biology, Carver College of Medicine, University of Iowa, Iowa City, IA, USA.

[12]CHU de Québec-Université Laval, Oncology Division, Laval University Cancer Research Center, Molecular Biology, Medical Biochemistry, and Pathology, Québec, QC, Canada.

[13]BIOTECMED, University of Valencia, Valencia, Spain.

[14]School of Psychology, University of Auckland, Auckland, New Zealand

[15]University Research Centre for Brain Research, University of Auckland, Auckland, New Zealand

[16]Anatomical Pathology, LabPlus, Auckland City Hospital, Auckland, New Zealand

[17]Pathology & Laboratory Medicine, Neurology, and Biological Chemistry, University of California Irvine School of Medicine, Irvine, CA, USA

[18]Neurobiology & Behavior, University of California, Irvine, CA, USA

[19]Center for Neurotherapeutics, University of California, Irvine, CA 92697, USA

[20]Anatomy and Medical Imaging, University of Auckland, Auckland, New Zealand

[21]Neurology, Osaka University Graduate School of Medicine, Osaka, Japan.

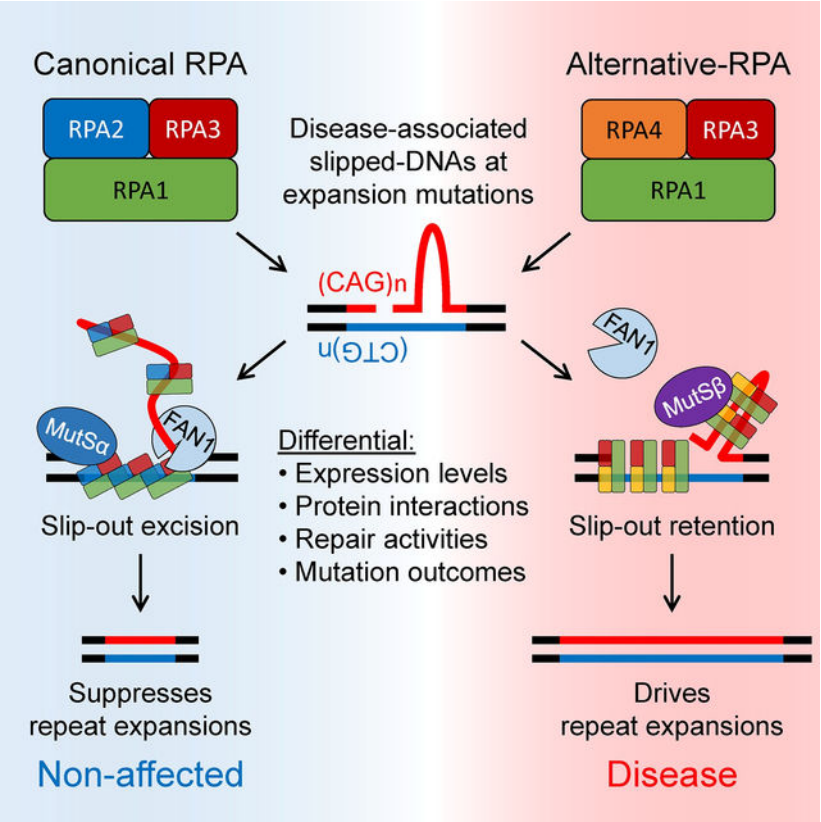
SUMMARY

Expansions of repeat DNA tracts cause >70 diseases, where ongoing expansions in brains exacerbate disease. During expansion mutations, ssDNAs form slipped-DNAs. We find ssDNA-binding complexes, canonical RPA (RPA1, RPA2, and RPA3) and Alternative-RPA (RPA1, RPA3, and primate-specific RPA4), are upregulated in Huntington disease and SCA1 patient brains. Protein interactomes of RPA and Alt-RPA reveal unique and shared partners, including modifiers of CAG instability and disease presentation. RPA enhances *in vitro* melting, FAN1 excision, and repair of slipped-CAGs, and protects against CAG expansions in human cells. RPA-overexpression in SCA1 mouse brains ablates expansions, coincident with decreased mATXN1-aggregation, reduced brain DNA damage, improved neuron morphology, and rescued motor phenotypes. In contrast, Alt-RPA inhibits melting, FAN1 excision, and repair of slipped-CAGs, and promotes CAG expansions. These findings suggest a functional interplay between the two RPAs, where Alt-RPA may antagonistically offset RPA's suppression of disease-associated repeat expansions, which may extend to other DNA processes.

IN BRIEF

Balance between expression, function, and protein-protein interactions of single-strand DNA-binding proteins canonical RPA and primate-specific Alternative-RPA, regulates somatic trinucleotide CAG repeat expansions associated with neurodegeneration in Huntington Disease and Spinocerebellar Ataxia Type 1.

Graphical Abstract



INTRODUCTION

Seventy neurodegenerative diseases are caused by expansions of gene-specific tandem repeat sequences[1], several being associated with (CAG)•(CTG) expansions including Huntington disease (HD) and multiple spinocerebellar ataxias (SCAs). Inherited expansions continue to somatically expand with age in affected tissues, suggesting that ongoing expansions drive disease age-of-onset (AOO), progression, and severity[2–4].

Screens for repeat disease modifiers validated three AOO modifiers, strengthening a connection of somatic expansions to disease[5–10]. The first two modifiers, repeat tract length and interruptions, modulate expansions by altering the formation of slipped-CAG DNAs —expansion mutation intermediates. The third modifier were DNA repair proteins (FAN1, MSH3, MLH1, PMS2, PMS1, LIG1, RRM2B, and POLD1)[10–16], that modulate somatic expansions by processing slipped-CAG structures. Naturally-occurring *Msh3* variants modulate CAG expansions in mice and humans[17], strenghting a causal link between instability and disease. Modifier screens identified many such proteins involved in expansions – although others crucial to expansions, such as MSH2 and MLH3, were not identified[11,18–20], suggesting the existence of additional modifiers.

Slipped-CAG DNA structures form from misaligned base-pairing as ssDNA re-anneals following transcription, replication, repair, or recombination[8,21,23–25]. Slipped-DNAs were identified at the expanded (CAG)•(CTG) repeat of the mutant *DMPK* gene in DM1 patient

tissues, with tissue-specific levels correlating with somatic expansion levels[26]. Additionally, a small-molecule targeting slipped-CAGs of >5 excess repeats induces contractions of the mutant CAG tract in HD mouse brains, further supporting long slipped-DNAs as mutagenic intermediates[27]. Thus, understanding proteins involved in slipped-DNA processing may reveal important modifiers of repeat instability. Single-strand DNA-binding (SSB) proteins are candidates, as slipped-DNAs form from single-stranded repeats[21,22].

The main SSB in humans is canonical RPA, a complex of RPA1, RPA2, and RPA3 (Figure 1A). RPA is essential for life, ubiquitously expressed and mediates virtually every process involving ssDNA[28–32] including DNA duplex destabilization, melting of unusual DNA structures, reannealing DNAs, protecting single-strand DNA, and protein recruitment to DNA[33–37]. Primates express an alternative form of RPA, Alt-RPA, which differs from canonical RPA via swapping of RPA2 for a paralog, RPA4 (Figure 1A)[38,39], present only in primates and some mammals (excluding mice)[38–40]. RPA4 has 47% and 63% amino acid sequence identity and similarity to RPA2, with similar domain organization (Figure 1B)[39]. RPA4 is less abundant than RPA2, is present in non-proliferating tissues, with low expression in brain[38,41]. RPA4 is reduced in cancers and cannot support cell cycle progression, consistent with a role in genome maintenance in non-proliferating cells[38,40,42]. While canonical RPA has been characterized in >3000 studies, Alt-RPA's characterization is limited to a small number of studies[38–40,42–44].

We assessed the role of RPA and Alt-RPA in somatic CAG repeat instability. We observe RPA and Alt-RPA are upregulated 2- and 10-fold, respectively, in HD and SCA1 patient brains relative to age- and sex-matched controls. Our data show RPA enhances *in vitro* slipped-CAG repair, while high levels of Alt-RPA block slipped-DNA repair. Both SSBs bind slipped-DNAs, but only RPA efficiently melts them. FAN1 nuclease mediated slipped-CAG cleavage is enhanced by RPA but inhibited by Alt-RPA. BioID of each RPA subunit revealed unique and shared associations with proteins important for somatic repeat instability. Expression of RPA inhibits, while Alt-RPA expression enhances, repeat expansions in human cells. Over-expressing murine Rpa1 ablates spontaneous somatic CAG expansions in SCA1 mouse brains, reduces disease biomarkers and improves motor functions. These results indicate that RPA/Alt-RPA modulate CAG repeat instability.

RESULTS

RPA and Alt-RPA are upregulated in HD and SCA1 patient brains

We quantified RNA and protein levels of RPA subunits in post-mortem brains from HD and SCA1 patients relative to unaffected individuals (Table S1). In HD, the striatum is highly degeneration prone[45], the cerebellum is moderately prone [46,47], and frontal pole does not degenerate[48–50], whereas in SCA1, the cerebellum is the most highly degenerated, with evidence of striatal degeneration in late stage disease[51–53]. In both HD and SCA1 somatic expansions are high in the striatum, limited in the cerebellum[2–4], and extremely low in the fontal pole (not shown).

Within the striatum, all RPA transcripts are significantly upregulated in HD patients relative to unaffected individuals (Figure 1C). *RPA1* is increased ~1.5-fold (p=0.0254), *RPA2*

~2-fold (p<0.0001), *RPA3* ~1.2-fold (p=0.0044), and *RPA4* exhibits striking ~5–6-fold upregulation (p<0.0001). Digital droplet PCR (ddPCR) upregulation was corroborated by real-time quantitative reverse transcription PCR (qRT-PCR) in a separate cohort of HD patients for *RPA2* (~1.5-fold, p=6.6e^{-6}) and *RPA4* (~8-fold, p=9.5e^{-5}) (Figure S1A). Protein upregulation was observed for RPA2 (~2-fold, p=0.0112) and RPA4 (~3–4-fold, p=0.0026) (Figures 1E–F, S1B). Upregulations occurred in 2 HD patient fibroblasts with ~2–3-fold upregulation of RPA2 and ~5–10-fold upregulation of RPA4 relative to 5 control lines (Figures 1H, S1I).

The cerebellum also shows RPA transcript upregulation (Figure 1D) with *RPA1* upregulated in HD (~1.5-fold, p<0.0091) and SCA1 (~2-fold, p<0.0001). *RPA2* was upregulated ~2-fold in HD (p<0.0001) and mild in SCA1 (~1.2-fold, p<0.0001) (Figure 1D). *RPA3* was upregulated ~2-fold in HD (p<0.0001), but not in SCA1 (p=0.583) (Figure 1D). *RPA4* was moderately upregulated by ~2.5-fold in HD (p=0.0091) and SCA1 (p<0.0001) (Figure 1D). qRT-PCR of a separate cohort confirms cerebellar upregulation of *RPA2* (~2-fold, p=8.6e^{-7}) and *RPA4* (~6-fold, p=2.8e^{-7}) (Figure S1A). Protein upregulation is evident in HD patient cerebellum for RPA2 (~2.5-fold, p=0.0099) and RPA4 (~2.5-fold, p=0.0065) (Figures 1E&G, S1B). In contrast, in SCA1 patient cerebellum, while RPA4 protein is upregulated ~4.5-fold (p=0.0003), RPA2 protein is not upregulated (p=0.1624) despite *RPA2* transcript upregulation (Figures 1E&G). Similarly, RPA2 is not upregulated in SCA1 (p=0.8524) and SCA3 (p=0.4642) patient fibroblasts, despite upregulations of RPA4 (~5–6-fold, p<0.0001) in both lines (Figures 1H, S1I).

In contrast to the other brain regions, the frontal pole of HD patients shows no upregulation of *RPA1* (p=0.2333), and mild upregulation for *RPA2* (~1.5-fold, p=0.0006), *RPA3* (~1.2-fold, p=0.0015), and *RPA4* (~2-fold, p=0.039) (Figure S1C). Neither RPA2 or RPA4 proteins are upregulated in the frontal pole (Figure S1D&E).

We demonstrate Alt-RPA is highly upregulated in HD and SCA1 patient brains, while canonical RPA is mildly upregulated. Expression patterns corelate with degenerative vulnerability in HD: RPA4 shows striking upregulation and RPA2 shows mild upregulation in the striatum (the most degenerated), mild upregulations of RPA2 and RPA4 in the cerebellum (less degenerated), and no upregulation in RPA2 and RPA4 in the frontal pole (devoid of neurodegeneration). A caveat of human postmortem tissues is the inability to assess DNA, RNA, or proteins in cells lost by degeneration. Therefore, our quantifications may underestimate RPA/Alt-RPA upregulation levels due to neurodegeneration incurred during disease progression. Brain region specificity could not be assessed in the SCA1 patients due to limited samples. RPA4 was highly upregulated while RPA2 was unaltered in SCA1 cerebellum (most degenerated in SCAs). This suggest upregulation of Alt-RPA expression (or increased ratios of Alt-RPA:RPA) may coincide with disease-specific repeat expansions and/or neurodegeneration[2–4].

RPA and Alt-RPA levels correlate with striatal degeneration in patients and fluctuate with age in HD mice

We assessed RPA and Alt-RPA level correlations with patient clinical information. We found no linkage of *RPA2* or *RPA4* levels with age-at-onset, age-at-death, post-mortem delay, or

inherited repeat length (Figures S2A,B,D&E, Data S1A). Lower striatal *RPA4* levels are mildly correlated with longer disease duration (R^2=0.65; Figure S2C), although *RPA2* levels are not linked with disease duration (patient lifespan after symptom onset). We observed negative correlations for RPA and Alt-RPA expression with increasing striatal grades (Figure S2F, Data S1A).

We assessed if Rpa2 protein levels are temporally regulated relative to neurodegeneration and CAG instability in HD mice and found Rpa2 fluctuated significantly with age in the striatum of zQ175 HD mice, mirroring the timing of striatal neurodegeneration and somatic instability (Figure S1I&F, Data S1A). No changes in Rpa2 levels were observed in the zQ175 cerebellum (Figure S1G–H, Data S1A).

Structural differences could underlie differential RPA and Alt-RPA functions

We hypothesized that regional sequence differences between RPA2 and RPA4[43] result in structural distinctions between RPA and Alt-RPA complexes. The AlphaFold predicted RPA4 structure model (Figure S3A–D)[54–56] has a DNA-binding domain (DBD) and a winged-helix domain, like RPA2 (67% sequence similarity). The Alt-RPA hetero-trimerization core was modelled by superposition of RPA4 DBD-G onto RPA2 DBD-D (Figure S3E), revealing changes in surface electrostatics compared to RPA, which may alter nucleic acid interactions (Figure S3F). Other structural differences could alter conformations or subunit-protein interactions to affect function. For example, although most complex interface residues in RPA2 (inter-protein distance <3.5Å) are conserved in RPA4, the RPA1-interacting FKIM (Phe-Lys-Ile-Met) motif of RPA2 (Figure S3G) is not conserved in RPA4.

Nick-in-repeat slipped-DNAs undergo structural transitions during in vitro repair that are resolved by RPA and, to a lesser degree, by Alt-RPA

We assessed functions of RPA and Alt-RPA in slipped-DNA repair using an *in vitro* assay[57–61] and circular (CAG)50•(CTG)30 slipped-DNA substrates with nicks in the repeat (nick-in-repeat) or in flanking sequences (nick-in-flank). These substrate's size and structure mirror the *mutagenic slipped-DNA expansion intermediates* that occur in patient tissues[26]. Nicks in the repeat generate a heterogenous mixture of *slipped-DNA structural conformations*; whereas nicks in the flank generate an anchored slip-out (Figure 2A bottom, see grey, white and black triangles)[59,62]. Southern-based detection permits molar assessments of reaction products, eliminating concerns of inter-lane loading variations or variable exposures.
We validate conformational distinctions between starting nick-in-repeat and nick-in-flank slipped-DNAs as distinct electrophoretic species and their repair by HeLa cell extract to fully-paired duplex products (Figure 2B–D). We also confirm that a slower-migrating species in nick-in-repeat, but not nick-in-flank, reactions are altered slipped-DNA structural conformations, which did not arise from DNA synthesis (green versus white and black triangles in lane 1 and 2 of Figure 2C–D, Figure S4A–C, Data S1B). This assay permits interrogation of RPA and Alt-RPA in the formation of DNA structural conformations and repair of slipped-CAGs expansion intermediates.

To test RPA and Alt-RPA functions in altering DNA structural conformations of nick-in-repeat substrates, and in efficiency of correct slip-out repair, we prepared HeLa extracts

depleted of RPA2 via siRNA (Figure S4D). RPA2 depletion also removes RPA1, thereby depletes RPA and Alt-RPA (Figure S4D), as previously observed[40]. We purifiedhuman RPA, Alt-RPA, and yeast RPA (scRPA)[39,42]. Preparations are pure, functional, heterotrimeric complexes, and nuclease-free (Figure S4E–G). Depleted extracts do not support SV40 *in vitro* DNA replication unless supplemented with human RPA, but not Alt-RPA, scRPA, or bacterial SSB (bSSB) (Figure S4H)[39,63]. The absence of RPA and Alt-RPA mildly increases levels of the DNA structural conformations in all reactions (green versus white or black triangles, Figure 2B–2J, lanes 2 versus 3), with the highest increase being ~2.8-fold (Figure 2E) and the lowest being ~1.1-fold (Figure 2I). Thus, neither RPA nor Alt-RPA are required to form the slow-migrating DNA structural conformations.

RPA supplementation diminished levels of the DNA structural conformations (8–24% reduction); whereas Alt-RPA supplementation caused less reduction (1–12%) depleted repair reactions (Figure 2C–F, lane 3 versus 4, %SI quantifications, green triangles). The ability to diminish DNA structural conformations is specific for human RPA and Alt-RPA, as neither scRPA, nor bSSB, could alter their levels to the same degree (Figure 2I–J). Altered slipped-DNA structural conformations are likely interconverting structural isomers of structurally dynamic slipped-DNAs. Our results suggest formation of altered slipped-DNA conformations in nick-in-repeat reactions does not require RPA or Alt-RPA, but their elimination is enhanced by RPA and to a lesser degree by Alt-RPA.

Quantification of repair: Low levels of RPA and Alt-RPA enhance correct slip-out repair, while high levels of Alt-RPA, but not RPA, inhibit slip-out repair

Repair of all Δ(CAG)20 slip-outs by non-depleted HeLa cell extract yield the expected nick-directed correct DNA repair products; the nicked strand being repaired using the continuous strand as template, yielding the correct repair products (CTG)50•(CAG)50 or (CTG)30•(CAG)30 (Figure 2B–D, arrowheads indicate correct repair products, compare lane 1 and 2). Correct repair was confirmed by radionucleotide-incorporated repair (Figures S4A–B).

Quantifying non-depleted HeLa repair, nick-in-flank slip-outs are only partially repaired (~24%; Figure 2B lane 2, repaired products (arrowheads) versus unrepaired substrate (grey triangle), whereas equivalent nick-in-repeat slip-outs are efficiently repaired (up to 93%, >7-fold greater) (Figure 2 compare lane 2 between B&C and E&G, arrowheads versus white/black triangles). A high efficiency (~73%, ~3-fold greater) of correct repair is also evident for the nick-in-repeat substrate (Figures 2 lane 2 in B & D and F & H, arrowheads versus white/black triangles). Repair is also greater when the nick resides in the non-slipped strand (Figure 2, lane 2 in C, E, G, and I versus D, F, H, and J), consistent with published findings[59,60,64]. We conclude that correct repair efficiency is highly sensitive to nick location, with nick-in-repeat slip-outs repaired more efficiently than nick-in-flank slip-outs.

RPA-/Alt-RPA-depleted HeLa extract correctly repairs slip-outs at similar levels as normal extracts (Figure 2B–J, lane 2 versus lane 3). Addition of RPA or Alt-RPA had mild stimulating effects on slip-out repair (Figure 2B–D, lane 3 versus 4 (RPA) and lane 5 (Alt-RPA), and 2D-J, lane 3 versus 4). Repair of a single excess repeat [(CAG)48•(CTG)47]

is also mildly reduced by depleted extract, being partially rescued by RPA and, to a lesser degree, by Alt-RPA (Figure S4I). Similarly, G-T mismatch was repaired by RPA-/Alt-RPA-depleted extract (Figure 2K, lane 2 versus lane 3), consistent with reports that HMGB1 compensates for loss of RPA to act in G-T and slip-out repair[86–94]. Despite this, G-T mismatch repair by depleted extracts is enhanced by low levels of RPA or Alt-RPA (Figure 2K, lane 3 versus lanes 4&7). Increasing RPA levels further enhance G-T repair, consistent RPA reported enhancement of mismatch and slip-out repair[65–75]. Conversely, increasing amounts of Alt-RPA diminishes G-T repair efficiency to levels of the depleted extract alone (Figure 2K, lane 3 versus 9). This suggests that low levels of Alt-RPA stimulate, but high levels diminish, G-T mismatch repair.

We tested how altering RPA:Alt-RPA levels, mimicking upregulations in HD and SCA1, influenced efficiency of correct slip-out repair. A 2:1 ratio of RPA:Alt-RPA caused a ~1.6-fold stimulation of correct repair (62% to 98%) with increasing ratios of RPA:Alt-RPA (6:1, 10:1, and 10:0) showing a similar stimulation for nick-in-repeat substrates nicked in the non-slipped strand (Figure 2G, lane 3 versus 4–8). A milder repair stimulating effect (81% to 85%) was observed with nick-in-repeat substrate nicked in the slipped strand (Figure 2H, lane 3 versus 4–8). Thus, low RPA levels competitively enhance slip-out repair in the presences of Alt-RPA, with no further enhancement occurring with increasing RPA levels.

In contrast, increasing levels of Alt-RPA over RPA causes inhibition of correct nick-in-repeat slip-out repair (Figure 2E&F, lanes 5–8 versus 4). Repair was reduced ~6-fold (85% to 15%) when shifting Alt-RPA:RPA molar ratios from 0:1 to 10:1 for nick-in-repeat substrates nicked in the non-slipped strand (Figure 2E, lane 8 versus 4). Similarly, reductions in repair occurred (~3-fold, 74% to 25%) by shifting Alt-RPA:RPA ratios from 0:1 to 10:1 for nick-in-repeat substrates nicked in the slipped strand (Figure 2F, lanes 8 versus 4). Repair inhibition was evident at the lowest increases of Alt-RPA:RPA; shifting from 0:1 to 2:1 Alt-RPA:RPA decreases repair for nick-in-repeat substrates ~1.2-fold (85% to 69%) nicked in the non-slipped strand and ~1.5-fold (74% to 50%) nicked in the slipped strand (Figure 2E&F, lanes 4–5 versus 3). A threshold level of Alt-RPA:RPA may reside between the 1:1 and 2:1, as enhanced repair by RPA at 0:1 is unaltered at a 1:1 ratio (Figure S4J). Thus, low levels of Alt-RPA could competitively overcome the enhancing effect of RPA upon slip-out repair, and further increases of Alt-RPA progressively inhibits slip-out repair to very low levels.

Together, these results suggest neither RPA nor Alt-RPA are essential for the formation of altered slipped-DNA conformations or repair of slipped-DNAs. However, in a concentration-sensitive manner, RPA competitively diminishes altered conformations and enhances slip-out repair, while high levels of Alt-RPA competitively and progressively inhibit slip-out repair. Slipped-DNA repair modulation is specific for human RPA and Alt-RPA as neither scRPA nor bSSB, could enhance slip-out repair (Figure 2I) or diminish the altered slipped-DNA conformations (Figure 2I&J). The DNA-binding mode of bSSB is different from RPA and both scRPA and bSSB differ in their protein-protein interactions from human RPA[42,63,76]. These data support specific interactions of the human SSB proteins with DNA or with other proteins (or both) as critical for their differential effects in slipped-DNA repair.

Alt-RPA has altered binding, poorly melts slipped-DNAs, and inhibits FAN1 cleavage

To understand differences in repair efficiency, we investigated DNA-binding of RPA and Alt-RPA using radiolabeled linearized nick-in-repeat slipped-CAG substrates by electrophoretic mobility shift assays (EMSA). Both RPA and Alt-RPA bound slipped-DNAs, with RPA yielding two distinct slower-migrating protein-DNA complexes with increasing RPA concentration (Figures 3A, S5E, lane 4 in each panel). Conversely, Alt-RPA yields only one protein-DNA complex at all concentrations (Figures 3A, S5E, lane 7). As RPA and Alt-RPA have ~20–30 nucleotide binding sites[31,39], these differences suggest RPA, but not Alt-RPA, is melting at least ~50 nucleotides of ssDNA.

Further EMSA analysis confirmed Alt-RPA and RPA both have high affinity for ssDNA (Figure S5A)[39]. Alt-RPA has reduced interactions with fully-paired duplexes and various bubble DNA substrates, with initial shifts requiring ~2.5 to 3.5 higher levels, and saturation binding requiring ~2.5 to 5 higher levels, of Alt-RPA relative to RPA (Figures S5A, S5B). We also probed DNA-binding by RPA and Alt-RPA, using a Förster resonance energy transfer (FRET) assay (Figure S5C) and observed, consistent with EMSA data, that more Alt-RPA than RPA is needed to reach 50% binding to duplex and slip-out substrates (Figure S5D).

Next, we investigated RPA- and Alt-RPA-mediated DNA melting/remodeling rates of bubble, non-repeat slip-out, and CAG-slip-out substrates via kinetic FRET assays with RPA or Alt-RPA under conditions of optimal melting (30 mM KCl and 0 mM MgCl2). RPA and Alt-RPA melt a bubble substrate at equal rates (Figure S5F) as published[77]. In contrast, Alt-RPA is less efficient than RPA at melting CAG and non-repeat slip-outs (Figure S5G–H). To further assess the melting kinetics of CAG slip-outs, we conducted melting in the presence of 5 mM MgCl2, a metal required for slip-out repair[57,59,78–80] which stabilizes DNA and slows melting[77]. Under these conditions, Alt-RPA shows ~64-fold slower melting than RPA (50% substrate melted in ~190 seconds for Alt-RPA versus ~3 seconds for RPA, $p<0.0001$, Figure 3B). Competition melting assays revealed a competitive ability of Alt-RPA to slow melting even in the presence of RPA as increasing molar ratios of Alt-RPA:RPA from (1:3 to 3:1) significantly slows melting of slip-outs by 5- to 17-fold ($p=0.0062$ to <0.0001). This suggests the kinetics of Alt-RPA binding and melting partially duplex DNA structures, including slip-outs, are slower than canonical RPA. Therefore, we predict slip-outs will have longer half-lives in the presence of Alt-RPA.

RPA's ability to melt unusual DNA structures and modulate nuclease activities[33,81–83] coupled with the differential ability of RPA and Alt-RPA to melt slip-outs, suggests that these complexes differentially affect slip-out processing. To test this hypothesis, we utilized FAN1 nuclease, which suppresses somatic CAG repeat expansions in HD mice brains by cleaving slip-outs with unique specificity[57] via cycles of DNA-binding-cleavage-dissociation, pausing between each cycle. SSB protein melting of DNA structures may be critical for FAN1 excision of excess repeats[57].

We assessed RPA's and Alt-RPA's influence upon FAN1 *endo*- and *exo*-nucleolytic activities on CAG-slip-outs (Figures 3C, S5I–J). In a dose-sensitive manner, RPA mildly stimulates FAN1 *endo*- and *exo*-nucleoytic digestion of slip-out DNA (Figures 3C, S5I). In

contrast, Alt-RPA inhibits FAN1's *endo*- and *exo*-nucleolytic digestion (Figures 3C, S5I). In competition assays with constant FAN1 and Alt-RPA levels, increasing RPA significantly enhances FAN1 slip-out digestion ($p<0.001$ to $p<0.0001$; Figure 3C and S5I), which would result in increased repeat stability. In contrast, competition assays with constant FAN1 and RPA levels, but increasing Alt-RPA, preferentially and significantly blocks cleavage ($p<0.0001$; Figures 3C, S5I), which would lead to expansions. RPA and Alt-RPA has similar effects on FAN1 digestion non-repeat flapped DNA (Figure S5J). Thus, Alt-RPA's altered binding to slipped-DNAs may be contributing to reduced slip-out melting and FAN1 nuclease activity while RPA promotes these processes.

Unique, shared, and novel protein-protein associations for RPA/Alt-RPA

Functional overlaps and distinctions between RPA and Alt-RPA may be revealed by the proteins they associate with. To identify protein-protein associations of RPA and Alt-RPA we conducted BioID proximity labelling assays[84–87] for RPA1, RPA2, RPA3, or RPA4 (Figures 4, S6, S7, Table S3, and Data S1C)[88] in HEK293T cells. We identify >2000 proximal protein associations for RPA1, RPA2, and RPA3, and ~700 associated proteins for RPA4 (>5 log2-fold enrichment and p-value <0.01 versus untransfected controls (Figures 4A, S6, and Table S3)[88]. The interactome is extensive including: i) CAG instability modifiers (MSH2, MSH3, MLH1, PMS1, XPG, etc.); ii) HD/SCA disease modifiers (HTT, RM2B/p53R2, UBR5, TCERG1/CA150, SETD2/HYPB/KMT3A, etc.); iii) CAG/CTG disease pathogenesis; iv) other repeat expansion diseases (RFC1, ATXN2, ATXN10, EIF4A3, DIP2B, etc.); v) DNA damage responses (P53, MRE11, NBN, LIG3, BRCA1, BRCA2, XPC, etc.); vi) chromatin biology; vii) DNA/RNA metabolism; viii) altered protein associations upon DNA damage (hydroxyurea); and ix) the largest GO subset is RNA metabolizing proteins, consistent with a growing appreciation of RPA (and possibly Alt-RPA) in RNA metabolism[89–93] (Figures 4, S6, S7, Table S3, and Data S1C). Associations with CAG instability modifiers are discussed here; other associations are discussed in Data S1C.

BioID captured known RPA interactors, including p53, RAD51, and DNA polymerases[28,94,95], and Alt-RPA interactors RAD51 and RFC[42,43] (Figures 4A–B, S7). Gene Ontology (GO) analysis showed proportional distributions of each subunit for molecular function, biological processes, and cellular components, with minor differences for RPA4 (Figure S6A and Data S1)[96,97]. Most associated proteins (1674) are shared between RPA1, RPA2, and RPA3 with many (581) shared between all four subunits (Figure 4A). Unique associations were also identified for each subunit, revealing 60 to 150 unique associations (4–8% of all associations, Figure 4A). Shared proximal associations are consistent with RPA/Alt-RPA being heterotrimeric complexes.

RPA/Alt-RPA differentially associate with proteins that regulate CAG/CTG repeat instability

Unique associations with RPA1 or RPA3 should reflect shared interactions with RPA and Alt-RPA, while those with RPA2 or RPA4 might reflect unique associations between RPA and Alt-RPA. BioID revealed differential protein associations between DNA repair proteins that regulate somatic repeat instability (Figure 4B). Strikingly, RPA4 but not RPA1–3, shows enriched association with MSH3 (Figure 4B), a modifier of age-of-onset and disease

progression[11–15], which with MSH2 forms MutSβ (required for CAG expansions)[17,98]. RPA1–3, but not RPA4, show enriched associations with MSH2 and MSH6 (Figure 4B), which form MutSα (required for mismatch repair but not expansions)[99,100]. Similarly, RPA1–3, but not RPA4, show enriched associations with MLH1 and PMS1, both modifiers of repeat diseases, which form MutLα (required for expansions) (Figure 4B)[11–15,19,101,102]. RPA4 also shows preferential enriched association with XPG which modifies CAG instability (Figure 4B)[58,71,103–105].

BioID findings were validated using cellular co-immunoprecipitation (IP) experiments in HD patient-derived and control fibroblasts (Figure 4C). p53, a known RPA interactor[106–110], co-IP'd with both RPA2 (unique to RPA) and RPA4 (unique to Alt-RPA). HTT (the Huntingtin protein) and MLH1, both identified by BioID as preferentially enriched with RPA1, were only observed upon RPA2 IP, but not RPA4 IP, suggesting their interaction with RPA but not Alt-RPA (Figure 4C). RPA2 co-IP confirmed associations with both MSH2, MSH6 (unique to MutSα), and MSH3 (unique to MutSβ) with the MSH6 interaction substantially higher than with MSH3, consistent with BioID. In contrast, Alt-RPA co-IP shows interaction with MSH2 and MSH3 (MutSβ) but not MSH6 (MutSα), consistent with BioID.

Direct interactions of RPA and Alt-RPA complexes with the MutSα (MSH2-MSH6) and MutSβ (MSH2-MSH3) complexes were assessed using purified protein complexes. Alt-RPA binds MutSα and MutSβ, while RPA only binds MutSα (Figure 4D). The detection of an RPA-MutSβ interaction in cells, but not *in vitro*, suggests this interaction is transient or indirect in cells. Similarly, that Alt-RPA interacts with MutSα *in vitro*, but not in cells, suggests that this interaction may be too weak to detect in cells. Altogether these results support the preferential interactions of Alt-RPA with MutSβ, and RPA with MutSα, MutL, and HTT.

RPA upregulation inhibits, and Alt-RPA upregulation promotes, somatic CAG expansions in human cells

To assess the role of RPA and Alt-RPA on somatic repeat instability, we overexpressed human RPA2 or RPA4 in the HT1080 cell model of CAG instability[27,40,57,98,111]. Cells lacking the overexpression constructs demonstrate an average somatic expansion gain of 4.1 CAG units in 10-days relative to the starting length (Figure 5A). Overexpression of RPA2 promotes significant contraction-biased instability ($p<0.0001$, average loss of ~60 CAGs) resulting from both inhibition of somatic expansions ($p<0.0001$) and promotion of somatic contractions ($p=0.004$) (Figure 5A). In contrast, overexpression of RPA4 promotes significant expansion-biased instability ($p=0.01$, average gain of ~37 CAGs) resulting exclusively from promotion of somatic expansions ($p=0.004$) (Figure 5A). Since cell proliferation was inhibited via serum starvation, these effects likely occur through replication-independent DNA repair.

RPA upregulation inhibits somatic CAG repeat expansions and rescues molecular, cellular, and motor phenotypes in SCA1 mice

To assess the *in vivo* effects of RPA modulation, we overexpressed *Rpa1* in the brains of mutant Ataxin-1 Q135 knock-in SCA1 mice[112], which mirror disease features of SCA1 patients including neuronal somatic CAG expansions in striatum and cerebellum[3,113–117], ataxia, cerebellar and brainstem atrophy[115], neuronal genome-wide DNA damage, and ubiquitin-positive expanded protein aggregates[115]. Murine Rpa1 was overexpressed by injecting AAV-EGFP-Rpa1 into the subarachnoid space for broad brain delivery in 5-week-old mice. Mice lack a functional *RPA4* and hence lack Alt-RPA, so RPA1 overexpression leads solely to canonical RPA upregulation. Previously, we demonstrated AAV-Rpa1 overexpression in these mice rescued motor phenotypes (gait and rotarod), Purkinje neuron morphology, elevated Purkinje DNA damage levels, and partially rescued impaired transcription, splicing, and abnormal cell cycle (Figure S8A)[112]. Following confirmation of striatal AAV delivery in the SCA1 mice via eGFP expression (Figure S8B), we demonstrate *Rpa1* RNA upregulation in the striatum (~2-fold, p=0.0057) relative to control mice (Figure S8C). *Rpa2* and *Rpa3* RNA levels are also upregulated in the same mice (Figure S8C), suggesting that Rpa1 upregulation is sufficient to upregulate the RPA complex, consistent with *in vivo* and experimental RPA1 overexpression studies[118–120].

Somatic CAG expansions were assessed from inherited length-matched mice, (CAG)~135, via fragment length analysis and repeat instability indices[121]. In control AAV-EGFP SCA1 mice we observe modest but consistent cerebellar expansions (Figures 5B, S8D) and extremely high levels of striatal expansions (>60 CAG gains, Figures 5B, S8E), as in SCA1/HD mice and humans[3,116,117,122]. AAV-mediated Rpa1 upregulation ablates expansions in the cerebellum (Figures 5B, S8D), although expansion levels did not permit statistical significance (p=0.76). Strikingly, Rpa1 upregulation within the striatum completely inhibits expansions ($p=8.63e^{-10}$, Figures 5B–C, S8D–E). These results, consistent with human cellular data (Figure 5A), suggest RPA protects against somatic CAG expansions.

RPA upregulation reduces genome-wide DNA damage and mutant ATXN1 aggregation

SCA1 and HD patients and mice present molecular disease markers, including spontaneous brain DNA damage and activated DNA damage response (DDR)[112,123–131]. As RPA acts in DDR[106,110,132,133], we assessed AAV-Rpa1 upregulation upon markers of double-strand DNA breaks (γ-H2AX and 53BP1). We observe significant reductions in γ-H2AX (p=0.0005) and 53BP1 (p=0.0007) in striatal medium spiny neurons (MSNs) and cerebellar Purkinje cells of AAV-Rpa1, relative to AAV-EGFP SCA1 mice (Figures 6A–B, S9). This suggests RPA upregulation suppresses spontaneous DDR, coincident with suppression of neuronal CAG expansions.

Aggregation of expanded polyglutamine proteins is a disease biomarker linked to polyglutamine and CAG repeat size[27,134–140]. Mutant mATXN1 aggregates are distinguished from non-aggregated ATXN1 as ubiquitin-positive nuclear inclusions that are resistant to proteasomal degradation[141]. In AAV-EGFP SCA1 mice, ~52% of MSNs exhibit >1 ubiquitin-positive ATXN1 aggregates, dropping to ~13% (p=0.003) in AAV-Rpa1

SCA1 mice where ATXN1 was not aggregated (Figure 6E). This suggests RPA-mediated suppression of neuronal CAG expansions causes strong reductions in mATXN1 aggregates.

DISCUSSION

Repeat expansion mutations involve single-stranded DNA intermediates and unusual DNA structures that require SSBs to stabilize, protect, melt, anneal and recruit DNA repair proteins to. Here we investigated the well-studied RPA and understudied Alt-RPA[38–40,42–44], and provide data supporting a model of antagonistic roles for Alt-RPA↔RPA in disease-associated somatic instability,with Alt-RPA opposing some RPA functions (Figure 7). RPA and, to a greater degree Alt-RPA, are upregulated in HD and SCA1 patient brain regions vulnerable to somatic expansions and degeneration. RPA enhances correct repair of slipped-DNA avoiding expansion mutations, whereas high levels of Alt-RPA inhibit slip-out repair, where retention of the excess repeats promotes expansion. Both RPA/Alt-RPA bind ssDNA with high affinity, yet RPA more rapidly melts slip-outs than Alt-RPA, suggesting a mechanism for differential slip-out repair outcomes mediated by these SSBs. Supporting this, RPA enhances, and Alt-RPA inhibits, slip-out excision by FAN1 nuclease, which diminishes expansions in HD mouse brains (Figure 3). These differences correlate with RPA-enhanced and Alt-RPA-inhibited *in vitro* repair of slipped-DNAs (Figure 2) and with the RPA-stabilization and Alt-RPA-expansions of CAG tracts (Figure 5). These data cumulatively support a model where RPA guards against somatic CAG expansions, thereby diminishing disease phenotypes, while Alt-RPA diminishes RPA's activity, promoting expansions and worsening phenotypes (Figure 7).

Supporting our model, overexpression of RPA in SCA1 mice prevents somatic expansions in the brain, diminishes disease markers (brain DNA damage and polyglutamine-aggregates) (Figures 6, S9), rescues neuron morphology, motor phenotypes, and partially rescued impaired transcription, splicing, and abnormal cell cycle[112]. These findings reveal RPA and Alt-RPA as active players in CAG repeat stability and instability, respectively. In bacterial and yeast models of instability, where repeat contractions predominate, an absence of SSBs enhanced contractions[142,143], consistent with our findings that metazoan RPA is required to protect against the predominating repeat expansions. Moreover, our finding that RPA and Alt-RPA modulate repeat instability in non-dividing human and murine neurons is consistent with overexpression of SSBs stabilizing repeats in replicating systems[142–144].

Our BioID protein-interactome, representing unbiased interactomes for RPA and Alt-RPA subunits, provides a resource to delineate the pathways involved in DNA metabolism in disease and non-diseased states. Differential associations of SSBs with known protein modulators of somatic CAG expansions may contribute to how RPA and Alt-RPA promote stability or expansions, respectively. For example, preferential interaction of RPA with MutSα is consistent with a protective role of MSH6 against repeat instability[98,99,145–147]. Enriched associations of RPA4 versus RPA2, with MutSβ (MSH2-MSH3) and XPG, is consistent with a requirement of MSH2, MSH3, and XPG for CAG instability in various systems[17,18,58,103–105]. Alt-RPA may work with MutSβ and FAN1 to mediate CAG repeat expansions, via formation and/or retention of excess repeats in poorly repaired slip-outs (Figure 7). Association of HTT with RPA1, part of RPA and Alt-RPA (Figure 4A–C),

could be consistent with a role of HTT in DNA repair[148–151]. While RPA1–4, like MSH2 and MLH3, were not identified in repeat disease modifier screens[11–15,152], our data reveals RPA1–4 as key players in CAG instability and possibly in disease. Notably, a screen for modifiers of transgenerational CAG instability in SCA3 families, identified RPA3 as a modifier[153].

Dysregulation of RPA1–4 can be clinically impactful given that increases or decreases in RPA expression has distinct effects. In humans, chromosomal microduplications or microdeletions of *RPA1* that result in modest increases (~1.5-fold) or decreases (~0.8-fold) in canonical RPA are associated with disease[118,154,155], whereas *RPA1* haploinsufficiency causes defective ATR-dependent DDR and G2/M checkpoint arrest[154]. Duplications of *RPA1* alter DDR and cell cycle changes distinctly from RPA haploinsufficiency[118] and a gain-of-function point mutation in RPA1, which increases DNA-binding activity, was associated with telomere-shortening disease[187]. Overexpression of RPA1–3 in cell culture leads to endoreduplication, attenuated DSBs, chromosomal instability, and cell death[118,156]. This data supports a direct link for RPA overexpression suppressing elevated γ-H2AX and 53BP1 levels in SCA1 mouse brains (Figure 6 and S9)[112,123–131,157]. RPA4 overexpression in the absence of exogenous stress also leads to DSB accumulation (γ-H2AX foci), and subsequent cell death[40], similar to the spontaneously activated DDR in SCA and HD patient cells and brains. That both depletion of canonical RPA and overexpression of Alt-RPA lead to DNA damage and apoptosis[118,156], further supports Alt-RPA↔RPA antagonistic interactions. Humans inheriting Xq21.33 duplications encompassing only *RPA4* and *DIAPH2* genes, are cognitively normal (males and females), suggesting Alt-RPA dosage variations may not be pathogenic[158]. Whether the levels of HD-related RPA and Alt-RPA expression change are causally or compensatory related to instability and/or neurodegeneration cannot be definitively linked currently.

Ablation of expansions by RPA overexpression in the SCA1 mouse striatum coincides with reduced disease phenotypes, further strengthens the therapeutic potential of targeting somatic expansions. CAG expansion levels in some HD brain regions correlate with disease age-of-onset, supporting a causal relationship of expansions to HD disease[4]. In SCA1 this relationship is understudied and not readily obvious in post-mortem analyses. Correlating instability and disease solely due to neurodegeneration is misleading, as HD and SCA1 patients can be clinically affected prior to any observed neurodegeneration[159–164]. Cell-type sorted HD patient brains show large CAG expansions and genome-wide transcriptional dysregulation in degenerating MSNs, but also in non-degenerating striatal interneurons, cerebellar Purkinje cells and cortical cell types, supporting cellular dysfunction in the absence of neurodegeneration[165,166]. In SCA1, cerebellar neurodegeneration is critical to disease yet does not account for all phenotypes in SCA1 patients or mice[51,167,168]. SCA1 pathology extends beyond the cerebellum and brainstem to involve the striatum and temporal lobe[51,167,168]. Over-expression of RPA can diminish expansions in brains of SCA1 mice and some disease phenotypes, supporting the concept that somatic CAG expansion could drive SCA1 disease. Therefore, somatic CAG expansion, regardless of a direct correlation to neurodegeneration levels, may drive the rate of various aspects of disease pathogenesis in different cell types in HD and SCA1. Furthermore, RPA overexpression

in SCA1 mice correlates with reduced disease biomarkers and phenotypes, supporting a potential link between somatic expansions and pathogenesis.

RPA4 related sequences are only found in placental mammals and active RPA4 genes exist in primates and certain mammals. *RPA4* retention in primates suggests it does/did confer some advantage. Though Alt-RPA inhibits replication, it can support some DNA repair in non-proliferating cells – raising questions about the natural role of Alt-RPA. In revealing previously unknown functions of canonical RPA and Alt-RPA, we highlight an antagonistic RPA↔Alt-RPA interaction that raises several interesting questions: Do Alt-RPA↔RPA antagonistic interactions impact RPA function across other DNA repair processes? Does regulated Alt-RPA↔RPA interactions sustain cells in a post-mitotic state or ensure high-fidelity DNA repair processes? Can Alt-RPA, like RPA, MMR and other DNA repair proteins, participate in health-requiring mutations, such as in immunoglobulin maturation[169,170]? Might Alt-RPA (being highly expressed in the testes), like RPA, modulate meiotic recombination to increase genetic diversity[171,172]? Does perturbation of the relative concentrations of RPA4 and RPA2 occur in and/or contribute to cancer? Our results suggest that most forms of DNA metabolism in primates could be affected by the antagonistic RPA↔Alt-RPA interaction.

LIMITATIONS OF THE STUDY

While our animal results support a protective role for RPA against somatic expansion, mice have a non-functional un-expressed *Rpa4* pseudogene, limiting our ability to equivalently assess Alt-RPA overexpression *in vivo.* Engineering a human-murine hybrid Alt-RPA in mice or “correction” and over-expression of the murine *Rpa4* pseudogene (distinct from human *RPA4*) is possible but would require extensive characterization of the engineered complex formation and function relative to the human Alt-RPA, whose functions are poorly understood. This analysis would be further hampered by the inability to silence or delete the essential RPA2 gene which in turn limits the ability to interpret Alt-RPA’s role in isolation upon somatic expansions. However, we provide evidence in human cells supporting a role for Alt-RPA promoting somatic CAG expansions, in opposition to the suppression of expansions by canonical RPA.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact: Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Christopher E. Pearson (cepearson.sickkids@gmail.com).

Materials Availability: Homemade antibodies for RPA4 and plasmid constructs used for BioID are available from the lead contact upon request. No other unique reagents were generated for this study.

Data and code availability:

- The mass spectrometry proteomics data collected for BioID have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are publicly available as of the date of publication with the dataset identifier PXD044158, as indicated in the key resource table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All patients, mice, and cell line (summarised here and in the key resources table) are described in detail in Table S1. The sex and age of HD, SCA1, and unaffected post-mortem individuals, where possible, was matched as closely as possible with near-equal males to females. Gender was unreported. Our study centered on adult post-manifest, hence age ranged from 32–80 years. The sex of HD and control mice was near-equal males to females. All SCA1 and associated control mice were male. Gender does not apply. Our study centered on post-natal animals hence ages ranged from 11–82 weeks.

Human participants:

Patient tissue sample collection, preparation, and patient descriptions.: Post-mortem patient tissues were provided by the Neurological Foundation Human Brain Bank with institutional ethics approval #011654 (7 HD patients and 7 unaffected individuals; striatum, cerebellum, frontal pole) directed by RLMF and MAC. ARLS provided 3 HD patients and 3 unaffected individuals; striatum and cerebellum, and the National Ataxia Foundation Biobank (3 SCA1 patients; cerebellum). Tissues were collected from patients using previously characterized protocols[174]. Briefly, unfixed brain is sectioned into discrete blocks and frozen with powdered dry ice, double wrapped in aluminum, and then stored at −80°C until processing. Known clinical information is outlined in Table S1.

Animals:

zQ175 HD mouse model description, handling, and tissue collection.: Animal protocols were approved by the Children's Hospital of Philadelphia Institutional Animal Care and Use Committee. Heterozygous zQ175 mice and littermate controls were housed on a 12-hour light/dark cycle in a temperature- and humidity-controlled environment with *ad libitum* access to food and water. Mice were anesthetized with ketamine/xylazine and perfused with ice-cold saline. Striatum and cerebellum samples were immediately collected on ice, flash frozen in liquid nitrogen, and stored at −80°C.

SCA1 KI mouse model description, handling, and tissue collection.: All mouse experiments, handling, and sacrifice were performed in strict accordance with the Guidelines for Proper Conduct of Animal Experiments by the Science Council of Japan. Handling, ethics, and tissue harvesting was previously described[112]. Briefly, mice were euthanized with ethyl ether, and tissues were collected within 5 minutes of death. Tissues were

immediately frozen in liquid nitrogen and then kept in −80°C until processing. Mutant *Atxn1* knock-in mice were crossed with background mice (C57BL/6J) during breeding. After multiple crosses, heterozygous knock-in mice with 125–140 repeats were used for all subsequent experiments, with non-transgenic siblings being used as controls.

Cell lines:

Patient derived cell line culturing.: Q43, Q40, and control 1 cell lines were a gift from Dr. Ray Truant (McMaster University) and were previously characterised[175]. Some lines were purchased from the Coriell Biorepository; Q43 HD line (code: GM02191), Q45 SCA1 line (code: GM06927), Q53 SCA3 line (code: GM06153). Control cell line 2 and 3 were a gift from Dr. Guy Rouleau (McGill University) and control cell line 4 and 5 were a gift from Dr. Elise Heon (University of Toronto; C4 and C5). For long-term storage, cells were immersed in Cellbanker 1 (Amsbio, catalogue #11888) and frozen in liquid nitrogen. All cells were cultured in DMEM (10% FBS, 1% supplemented L-glutamine, 1% Penicillin-Streptomycin) at 37°C with 5% CO2. Cells were plated at ~50% confluency, and split at 85–95% confluency by Trypsin. Viability was checked using Trypan Blue exclusion tests during all splits and prior to experimentation; a viability of 90% or greater was maintained for cells prior to experimentation.

METHOD DETAILS

RNA preparation from patient and mouse brain tissues.—Tissues (stored and −80°C and kept immersed liquid nitrogen during handling) were crushed with a frozen metal mortar and pestle partway buried in dry ice, and frozen crushed tissues were immediately transferred to a 1.4 mm acid washed tube pre-filled with zirconium Beads and 300–1000 μL of TRIzol reagent. Smaller tissues were directly inserted into tubes without crushing. Tubes were inverted to ensure immersion of the whole tissue, and were placed at room temperature for 10 minutes to allow the TRIzol reagent to denature and remove proteins bound to RNA. Tubes were placed on ice after ten minutes and then placed in a MagNA Lyser Instrument (Roche; item #03358968001). Tubes were oscillated at 7000 OSC 3 times for 20 seconds each oscillation, with a 3-minute incubation on ice between each 20 second oscillation. TRIzol was transferred to a different tube, RNA precipitated by an equal volume of 100% EtOH and then purified using the Direct-zol RNA purification kit using the manufacturers protocol, which includes in-column DNase treatment (Zymo research; catalog # R2071). Whole RNA was reverse transcribed using the SuperScript IV First-Strand Synthesis System kit using the manufacturers protocol (ThermoFisher Scientific; catalogue #18091050).

Droplet digital PCR (ddPCR) for RNA transcript expression quantification.— FAM and HEX fluorophore labelled probes specific for each RNA target of interest was ordered from Bio-Rad (pre-made probe designs), and manufacturer's "PrimePCR ddPCR gene expression probe assays" protocol was used for ddPCR reactions. In brief: 10–50 ng of total cDNA (depending on target abundance, empirically derived from preliminary runs) was mixed with: 1) 2x ddPCR Supermix for Probes (no dUTP), 2) 20x target primers/probe mix (FAM), and 3) 20x reference primers/probe (HEX), topped up to a final reaction volume of 20 μL with DNase-/RNase-free water. Plate was sealed with aluminum, mixed well and centrifuged briefly to collect the reaction, and then kept at room temperature for 3 minutes

to equilibrate the reaction temperature to room temperature. Droplets were generated using the Automated Droplet Generator (Bio-Rad; catalogue number: 10043138). Plate containing generated droplets was re-sealed with aluminum, and then subjected to PCR in a C1000 Thermal Cycler (Catalog #185–1197) to the following cycles: 1) 1× 95°C, 10 minutes, 2) 40x (94°C, 30 seconds followed by 55°C, 1 minute), 3) 1x (98°C, 10 minutes), 4) held at 4°C until further processing. A ramp rate of 2°C/second was used for all the cycles. After cycling, the plate was transferred to a QX200 Droplet Reader (Bio-Rad; catalogue #186–4101) for fluorescent detection, and was analyzed using the Bio-Rad QuantaSoft Software. All experiments were conducted using at least 3 technical replicates.

Quantitative reverse transcription PCR (qRT-PCR).—RNA isolation from patient brain tissues, and cDNA generation performed as described in the ddPCR methods above. mRNA quantification was performed using TaqMan probes (Life Technologies) and TaqMan Universal PCR Mix (ThermoFisher Scientific; catalogue # 4304437) on a 7500 Real Time PCR System (Applied Biosystems). Gene expression was normalized to 18S rRNA. Delta CT values were calculated as Cttarget - Ct18S. Note: all experiments were conducted using at least 3 technical replicates.

Protein lysate preparation (cultured cells).—RIPA Lysis and Extraction Buffer (ThermoFisher Scientific; catalogue #89901) was mixed with and appropriate volume of 100x Halt Protease and Phosphatase Inhibitor Cocktail (ThermoFisher Scientific; catalogue #78430) and kept on ice. Fibroblast cells were kept adhered to the culture flask, media removed, and washed twice with 1x sterile PBS. After the last wash, as much excess PBS was removed as possible and 150 μL - 700 μL of RIPA with protease inhibitor (depending on cell numbers) was added directly to the culture flask. Flask was tilted to allow RIPA to cover the whole surface that cells were grown on, and then placed flat (to ensure whole surface was covered with RIPA) on ice for 1 hour. Cells were then scraped on ice using a rubber scraper, and RIPA was collected into an epi-tube. Cells were then sonicated on ice using a microtip (amplitude 20, 15 cycles, each consisting of 1 second on and 1 second off), with the microtip being cleaned twice with water and then 70% ethanol (wiped dry with a Kim-Wipe) between each sample. Cell debris was pelleted by centrifuging at 21,000x*g* for 15 minutes at 4°C, and the supernatant collected for subsequent experiments. Supernatant was aliquoted into 100 μL aliquots to avoid repeat freeze-thaws and then stored at −80°C between uses.

Protein lysate preparation (tissues).—Tissues (stored in −80°C and kept immersed in liquid nitrogen during handling) were crushed with a frozen metal mortar and pestle placed on dry ice, and frozen crushed tissues were immediately transferred to a 1.4 mm Acid Washed tube pre-filled with Zirconium Beads and 300–1000 μL of RIPA solution without detergent (50 mM Tris HCl pH 7.4, 150 mM NaCl, 1 mM EDTA) with an appropriate volume of 100x Halt Protease and Phosphatase Inhibitor Cocktail (ThermoFisher Scientific; catalogue #78430), on ice. Smaller tissues were directly inserted into tubes without crushing. Tubes were inverted to ensure immersion of the whole tissue, and then placed back on ice before processing in a MagNA Lyser Instrument (Roche; item #03358968001). Tubes were oscillated at 7000 OSC 3 times for 20 seconds each oscillation, with a 3-minute incubation on ice between each 20 second oscillation. To ensure complete tissue homogenization,

samples were briefly spun to assess the level of unhomogenized tissue left (if any), and additional oscillation cycles were performed as needed. An equal volume of RIPA double detergent (2%DOC, 2% Igepal, 2% Triton X-100) with an appropriate volume of 100x Halt Protease and Phosphatase Inhibitor Cocktail (ThermoFisher Scientific; catalogue #78430) was added to each tube. Parafilm and clips were attached to the lid of each tube, and then incubated on a sample roller overnight at 4°C. The next day, tubes were centrifuged at 14000x*g* for 15 minutes at 4°C, and the supernatant collected for subsequent experiments. Supernatant was aliquoted into 100 μL aliquots to avoid repeat freeze-thaws and then stored at −80°C between uses.

SDS-PAGE and western blotting.—10 μg - 100 μg of protein lysate was used per sample (depending on target protein abundance, and kept consistent between samples on the same gel). Samples were prepared using 4x NuPAGE LDS Sample Buffer (ThermoFisher Scientific; catalogue #NP0007) and 10x NuPAGE Sample Reducing Agent (ThermoFisher Scientific; catalogue #NP0004), and denaturing the samples at 70°C for 10 minutes. The denatured samples were electrophoresed at 100–120 volts for 1.5–2.5 hours on NuPAGE 4–12% Bis-Tris Proteins Gels (ThermoFisher Scientific; catalogue # NP0321BOX) in NuPAGE MES SDS Running Buffer (ThermoFisher Scientific; catalogue #NP0002). Samples were run in parallel with Full range rainbow MW marker (ThermoFisher Scientific, catalogue #RPN800E) and/or HiMark Pre-stained Protein Standard (ThermoFisher Scientific, catalogue #LC5699). Gels were wet-tank transferred to PVDF Western Blotting Membranes (Sigma-Aldrich, Cat #3010040001; activated in 100% methanol for 1–2 minutes prior to use) in tris-glycine (with 10–20% methanol) overnight (16–24 hours typically) at 4°C using a constant voltage of 20–30V. The next day, membranes were blocked in 5–10% w/v milk dissolved in 1xTBS + 0.1% Tween-20 (TBST) for 1 hour at room temperature. Blots are then incubated with primary antibody at room temperature for 2 hours using the same solution used for, washed 3 times in 1xTBST at room temperature (10 minutes/wash), incubated with secondary antibody at room temperature for 1 hour in the same solution used for blocking, washed 3 times in TBST at room temperature (10 minutes/wash), and then detected with ECL according to the manufacturer's instructions (GE Healthcare Amersham ECL™ Prime Western Blotting Detection Reagent, Cat #RPN2232) by autoradiograph. Densitometric quantification of bands was performed using Image Studio Lite Version 5.2 (LI-COR Biosciences).

Antibodies used for western blotting.—*Primary antibodies:* Anti-RPA2 clone 9H8 (1:1000, monoclonal mouse, Abcam catalogue #ab2175), Anti-RPA4 (1:4000–1:8000, sheep serum, homemade), Anti-Actin Protein Antibody (1:30,000, monoclonal mouse, BD Transduction Laboratory, catalogue #612657). *Secondary antibodies:* Peroxidase-AffiniPure Sheep Anti-Mouse IgG H+L (1:2000, Cedarlane Labs, catalogue #515035062), Sheep IgG (H+L) Highly Cross-Adsorbed Donkey anti-Ovine HRP (1:2000–1:6000, ThermoFisher Scientific, catalogue #A16047)

Antibodies used for IF.—*Primary antibodies:* anti-phospho-H2AX (γ-H2AX) clone JBW301 (1:200, monoclonal mouse, Millipore Sigma catalogue #05–636), anti-53BP1 (1:5000, polyclonal rabbit, Novus bio catalogue #NB100–304SS), anti-Ataxin1 clone N76/8

(1:100, monoclonal mouse, EMD Millipore catalogue #MABN37), anti-ubiquitin clone P4D1 (1:1000, monoclonal mouse, Cell Signaling Technology catalogue #3936S), anti-DARPP32 clone 19A3 (1:200, monoclonal rabbit, Cell Signaling Technology catalogue # 2306), anti-calbindin clone EG-20 (1:2000, polyclonal rabbit, catalogue EMD Millipore #05–636), anti-calbindin clone CB-955 (1:2000, monoclonal mouse, Millipore Sigma catalogue #C9848). *Secondary antibodies:* Goat anti-Mouse IgG (H+L) Superclonal Recombinant Secondary Antibody Alexa Fluor 555 (1:200, ThermoFisher Scientific catalogue #A28180), Goat Anti-Rabbit IgG H&L Alexa Fluor 488 (1:200, Abcam catalogue #ab150077), Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody Alexa Fluor 568 (1:200, ThermoFisher Scientific catalogue #A-11011), Goat Anti-Mouse IgG H&L Alexa Fluor 488 preabsorbed (1:200, Abcam catalogue #ab150117).

Functional cell extract preparation.—Cells are grown in 20 cm plates to ~70–80% confluence. Media is removed, and cells are washed twice with ice cold hypotonic buffer (20 mM HEPES-KOH pH 7.8, 5 mM KCl, 0.15 mM MgCl2 and 0.1 mM DTT). Remove as much excess hypotonic buffer from washes as possible, add 300 μL of hypotonic solution with 3 μL of 100x Halt Protease and Phosphatase Inhibitor Cocktail (ThermoFisher Scientific; catalogue #78430), and then scrape the cells using a rubber scraper. Collect cells into a Dounce homogenizer. Dounce cells ~10–15 times using a tightly fitting pestle (B pestle). Homogenization of cells can be checked on a slide under a light microscope, and additional homogenization can be conducted as needed. Transfer extract to a larger volume conical centrifuge tube (15 mL or 50 mL) and let stand on ice for 30 minutes. Centrifuge at 1700x*g* for 10 minutes at 4°C to pellet large cell debris, then transfer to high-speed centrifuge tubes and centrifuge again at 12,000x*g* for 10 minutes at 4°C to clarify the extract further. Remove supernatant and freeze as beads by dripping into liquid nitrogen, and then store in −80°C.

***In vitro* repair reaction and Southern blotting.—**Substrate generation and repair reactions were performed as previously described[59,60]. In brief, each repair reaction consists of 1 μL slipped-DNA substrate, ATP, rNTP-ATP, calf phosphatase, calf kinase, dATP, dCTP, dGTP, dTTP, cell extract, with and without supplementation with purified protein. The reaction is incubated at 37°C for 1 hour, after which the reaction is stopped by 2% SDS, 2 mg/mL proteinase K, 0.05 M EDTA and incubating for another hour at 37°C. Phenol:chloroform extraction is performed and then DNA purified using MinElute Reaction Cleanup Kit according to the manufacturer's protocol (Qiagen, catalogue # 28206). The DNA is then digested with *Eco*RI and *Hin*dIII overnight at 37°C. The next day the DNA is electrophoresed at 200 volts on a 4% polyacrylamide gel in 1x TBE for 1 hour and 25 minutes. The electrophoresed DNA is then transferred from the gel to a nitrocellulose membrane using a Owl HEP Series Semidry Electroblotting System (ThermoFisher Scientific, catalogue #HEP-1). The transferred membrane is then immersed in denaturing solution (1.6% w/v NaOH pellets in ddH2O) for 20 minutes at room temperature with gentle agitation, renatured in Southern neutralising solution for 20 minutes at room temperature with gentle agitation, and then washed in 5x SSPE for 20 minutes at room temperature with gentle agitation. Following this, the membrane is rolled into a glass hybridization tube and blocked with salmon sperm DNA in Southern prehybridization solution for an hour at

42°C. After this, a 32P radioactively labelled probe complimentary to the DNA is added to the prehybridization solution and allowed to hybridize overnight at 42°C. The next day, the radioactive probe in prehybridization solution is removed and the blot is washed 3x with Southern wash solution (0.1% SDS in 0.1% SSPE v/v in water) - each wash being at least 30 minutes long at 65°C). Lastly, the membrane is exposed to an autoradiograph to visualize the DNA.

Quantifications of repair reaction products were done using previously published methods[19,24,32–35]. Since products were assessed by Southern blots, all analyses are molar, and can be compared within a reaction, eliminating any concerns of inter-lane loading variations, inter-experimental variation, or variable exposures. Densitometric intensities reflect molar amounts. Since the correct slip-out repair products electrophoretically resolve as distinct bands, isolated from unrepaired DNA species, quantifications are accurate, with some inter experiment variation, as we previously published[19,24,32–35]. The same is true for quantifying G-T mismatch repair reactions, as we have published[19,24,34,35]. In the case of the structural intermediates, that arise from only the nick-in-repeat slipped-DNA substrates, which are “repair” products that have not incurred DNA synthesis, we quantify the levels of the intermediates, as done for the correct repair products. We note that the structural intermediates electrophoretically resolve as a series of tightly-spaced ladder-bands or as a smear – typical of heterogenous random coil conformations[36–43].

siRNA administration to cultured cells.—The RPA2 siRNA was from Santa Cruz biotechnology (catalogue #sc-38230). siRNA was used according to the manufacturer’s specific protocols.

***In silico* analysis of canonical-RPA and Alt-RPA.—**Human RPA2 and RPA4 sequences were extracted from Uniprot and aligned using Clustal Omega. The RPA4 structure model and associated model prediction data were downloaded from AlphaFold (accession Q13156) and analyzed using Pymol and APBS.

Electrophoretic mobility shift assay (EMSA).—Radioactively or fluorescently labelled DNA and proteins/compound are incubated at room temperature for 15–30 minutes in a reaction containing purified proteins and DNA in a binding buffer (3mM HEPES pH 7.9, 16 mM NaCl, 0.04 mM EDTA, 0.2 mM DTT, 0.06 mg/ml BSA, and 2% glycerol). Following this incubation, 1–2 μL of 10x native sample binding dye (50% glycerol with Bromophenol Blue) are added to the reaction and loaded onto a 4–8% polyacrylamide gel or 1–1.5% agarose gel as quickly as possible. The reaction is electrophoresed for 1–3 hours at room temperature or overnight at 4°C. For fluorescently labelled DNA, gels are visualized using a fluorescent detection system (typically an Amersham Typhoon laser-scanning platform or Bio-Rad ChemiDoc MP imaging system). For radioactively labelled DNA, gels are dried to a Whatman paper and then exposed to autoradiograph to detect band shifts.

Kinetic FRET binding and melting assays.—Two complimentary oligonucleotides with a Cy3 or Cy5 fluorophore were annealed to one another such that the fluorophores were adjacent to one another. Cy3 was excited by an external 530 nm laser, while the Cy5

fluorophore was excited by the emission of the Cy3 fluorophore, allowing for observance of both fluorescent signals when the two strands were annealed to one another. Once melted, only the Cy3 emission will be observed. Binding and melting can be quantified as a function of the observance of one versus two signals (FRET calculation in Supplementary Figure S5C). FRET quantifications were normalized as a value from 1 to 0 so individual experiments could be compared to one another. Cy3 and Cy5 emission intensities were assessed via a Cary Eclipse Fluorescence Spectrophotometer. Equilibrium binding was quantified by plotting normalized FRET values to protein concentration. Increasing amounts of RPA or Alt-RPA were titrated into the solution containing 1 nM of each DNA substrate, and the midpoint of DNA substrate binding was used to infer relative DNA binding affinities of RPA and Alt-RPA Protein-mediated unwinding of DNA (melting) was observed by saturating the DNA with the purified protein of interest, and observing the FRET signal over time. The data were plotted in GraphPad Prism and fitted to exponential functions. The time needed to reach FRET = 0.475 (i.e. half of the DNA being melted) was used to quantify the rate of substrate melting by each complex.

RPA, Alt-RPA, and bSSB protein purification.—Recombinant RPA, Alt-RPA, and bSSB were expressed and purified from *E. Coli* as previously described[176,177]. Purity and heterotrimeric complex formation for RPA and Alt-RPA was accessed by mass-photometry. Mass photometry was performed on 100 nM sample of purified RPA or Alt-RPA using the Refeyn TwoMP mass photometry instrument (Refeyn Ltd. Oxford, UK) in buffer containing 20mM Tris (pH 7.4), 100mM KCl, and 1mM DTT. Molecular weight calibrations were performed using two protein oligomer solutions, β-amylase (56, 112 and 224 kDa) and Thyroglobulin (670 kDa). Individual molecular weights collected from 3000 frames (59.9 seconds) were binned in 3kDa bins and plotted as frequency histograms and fitted to single Gaussians using GraphPad Prism.

FAN1 protein purification.—Recombinant human FAN1 protein is expressed and purified from Sf9 insect cells as described previously[57,178].

FAN1 nuclease assay.—FAN1 nuclease assays were performed in nuclease assay buffer (50 mM Tris HCl pH 8.0, 25 mM NaCl, 1 mM MnCl2, 1 mM dithiothreitol, 200 mg/ml BSA) with 100 nM of fluorescently labeled DNA incubated with 50 nM of FAN1 protein. Reactions were initiated by the addition of protein, incubated at 37°C, for 20 minutes then stopped with formamide loading buffer (95% formamide, 10 mM EDTA). Products of were separated using 6% denaturing sequencing gel for 1 hour at 2000 V and detected at fluorescence filter in the Typhoon FLA (GE Healthcare).

RPA and Alt-RPA subunit BioID and data analysis.—BioID is a proximity-labeling technique where a biotin tag is covalently attached to proteins that directly interact with the bait protein (including capturing weak and transient protein-protein interactions), as well as proteins in close proximity but that may not be direct interactors of the bait protein (indirect interactions)[84–87,179–181].

Cloning and expression in HEK293FT cells.: RPA1, RPA2, RPA3 and RPA4 expression constructs were generated from specific PCR amplification from a cDNA library, and

inserted into either a pgLAP1–3MYC-BioID2 vector using Gateway cloning as per the manufacture's protocols (Gateway BP Clonase - ThermoFisher Scientific, catalogue #11789100 and Gateway LR Clonase - ThermoFisher Scientific, catalogue #11791020). This vector produces a myc-tagged construct conjugated to a functional BioID2 at the protein N-terminal. pgLAP1–3MYC-BioID2 subunit plasmids were then transformed and stably integrated into HEK293-Flp-In-T-REx cells using the manufacturer's protocol (ThermoFisher Scientific, catalogue #R78007).

Biotinylation, pull down, and sample preparation.: 293-FT stable cell lines were induced to express pgLAP1–3MYC-BioID2-RPA 1–4 by Doxycyclin 48 hours prior to pull-down at ~40–50% cell confluency. The next day, 50 μM Biotin was added 24 hours prior to the pulldown, with or without hydroxyurea treatment of 1 mM. After 24 hours, cells were washed 3x with cold sterile 1x PBS, trypsinised, pelleted at 1500x*g* for 5 minutes at 4 °C, and then washed again 2x with cold sterile 1x PBS. 600 μL of cold, freshly made lysis buffer (8M urea, 50 mM HEPES, pH 7.4, 1 mM PMSF, 1 mM DTT, 1% Triton X-100) was added to each pellet and allowed to lyse on ice for 1 hour. Pellets were then sonicated on ice (30 amplitude, 2 cycles of 10 second on and 10 seconds off, with 600 μL of fresh lysis buffer being added to the pellet between the two cycles). The lysate was centrifuged at 16,500x*g* for 10 minutes at 4 °C, and the supernatant was transferred to a fresh tube. Streptavidin Sepharose High Performance beads (Sigma-Aldrich, catalogue # GE17–5113-01) were washed 3x with 1 mL cold lysis buffer, and then added to each of the samples. Samples were rotated overnight with the beads at 4°C. The next day, the samples were centrifuged for 5 minutes at 1000x*g*, and the supernatant removed. Beads were washed 4x by rotated for 10 minutes at room temperature in 1 ml wash buffer (8M Urea, 50 mM HEPES, pH 7.4). Beads were pelleted by centrifuging for 2 minutes at 1000x*g* and transferred to a new tube. From this point on MS-grade water was used for the preparation of all buffers. Beads were washed 4x with 1 mL 20 mM ammonium bicarbonate water, and then 1x with the same buffer with 1 mM added biotin (to saturate unbound streptavidin). The bound proteins were then reduced using 50 μl of 20 mM ammonium bicarbonate buffer with 10 mM added DTT for 30 minutes, rotating at 60°C. Samples were cooled to room temperature for 5 minutes. Proteins were then alkylated in a light-tight container using 50 μL of 50 mM ammonium bicarbonate buffer with 15 mM chloroacetamide for 1 hour rotating at room temperature. The chloroacetamide was neutralized by adding DTT to a final concentration of 15 mM and then rotating for 3 minutes at room temperature. The proteins were then digested by adding 1 μg Pierce MS-grade trypsin (ThermoFisher Scientific, catalogue # 90058) and incubating overnight at 37°C while rotating. The next day, formic acid was added to a final concentration of 1%, and tubes rotated for 5 minutes at room temperature, to stop the reaction. Beads were centrifuged at 2000x*g* for 3 minutes at room temperature and the supernatant was collected into a fresh tube and then put to the side. Beads were then resuspended in 100 μL of 60% acetonitrile and 0.1% FA, and then rotated for 5 minutes at room temperature. The beads were then centrifuged again at 2000x*g* for 3 minutes at room temperature and the supernatant collected and added to the supernatant from two steps prior. The supernatant was dried by a centrifugal evaporator at 60°C until completely dried and then resuspended in 30 μL of 0.1% trifluoroacetic acid (TFA) buffer. Peptides were then purified with ZipTip 10-μl micropipette tips containing a C18 column as

per the manufacturer's protocol (EMD Millipore, catalogue # ZTC18M008). Peptides were eluted in new tubes, in a final volume of 30 µL comprised of 50% ACN and 1% FA buffer. The supernatant was dried by a centrifugal evaporator at 60°C until completely dried and then resuspended in 30 µL of 1% FA buffer. Peptides were then transferred to a glass vial and stored at −20°C until mass spectrometry analysis.

LC-MS/MS analysis.: 250 ng of each sample was injected into an HPLC nanoElute system (Bruker Daltonics), loaded onto a trap column with a constant flow of 4 µl/min (Acclaim PepMap100 C18 column, 0.3 mm id × 5 mm, Dionex Corporation, catalogue # 164567), and then eluted onto an analytical C18 Column (1.9 µm beads size, 75 µm × 25 cm, PepSep). Peptides were eluted over 2 hours in a gradient of acetonitrile (5–37%) in 0.1% FA at 500 nL/min while being injected into a TimsTOF Pro ion mobility mass spectrometer equipped with a Captive Spray nano electrospray source (Bruker Daltonics). Data was acquired using data-dependent auto-MS/MS with a 100–1700 m/z mass range, with PASEF enabled, number of PASEF scans set at 10 (1.27 seconds duty cycle), a dynamic exclusion of 0.4-minute, m/z dependent isolation window and collision energy of 42.0 eV. The target intensity was set to 20,000, with an intensity threshold of 2,500.

Protein identification by MaxQuant analysis.: Raw data files were analyzed using MaxQuant version 1.6.17.0 software392 and a Uniprot human proteome database (21/03/2020, 75,776 entries). The settings used for the MaxQuant analysis (with TIMS-DDA type in group-specific parameters) were: 2 miscleavages were allowed; fixed modification was carbamidomethylation on cysteine; enzymes were Trypsin (K/R not before P); variable modifications included in the analysis were methionine oxidation, protein N-terminal acetylation and protein carbamylation (K, N-terminal). A mass tolerance of 10 ppm was used for precursor ions and a tolerance of 20 ppm was used for fragment ions. Identification values "PSM FDR", "Protein FDR" and "Site decoy fraction" were set to 0.05. Minimum peptide count was set to 1. Label-Free-Quantification (LFQ) was also selected with a LFQ minimal ratio count of 2. Both the "Second peptides" and "Match between runs" options were also allowed.

Data analysis and statistics.: Following analysis, results were sorted by parameters set by Prostar software (Proteomics statistical analysis with R). Proteins positive for at least either one of the "Reverse", "Only.identified.by.site" or "Potential.contaminant" categories were eliminated, as well as proteins identified from a single peptide. An SLSA (Structured Least Square Adaptative) and DetQuantile imputation were performed for, respectively, POV (Partially Observed Value) and MEC (Missing in the Entire Condition) missing values. After a mean centering within each condition, results were sorted to retain proteins that were present in at least 2 of 3 biological replicates for each condition. For the mass spectrometry analysis, the specific protein-protein interaction networks involving each of the RPA or Alt-RPA proteins was quantified based on intensities to obtained enrichment ratios and MS/MS counts. Quantification of the identified proteins for each subunit measured the enrichment in comparison to the negative control HEK293-FT. Experiments were performed in biological triplicates. Enrichment ratios where significant when over the 90% percentile

of associated proteins. Enrichment ratios detected for each quantified protein were compared between subunits.

Co-immunoprecipitation from human fibroblast cell lines.—Cells were grown until 90% confluent in DMEM media (Thermo Fisher Scientific, Catalogue #11995073) containing 15% Fetal Bovine Serum (Thermo Fisher Scientific, Catalogue #12483020). Cells were fixed with 1% Paraformaldehyde (VWR, Catalogue #100503–914) for 20 minutes at room temperature and then quenched with 1.25M Glycine in PBS. Fixing media was removed and fixed cells washed twice with cold PBS. Modified Pierce RIPA Lysis buffer (Thermo Fisher Scientific, Catalogue #89900) supplemented with 0.5% of sodium deoxycholate and Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific, Catalogue #78446) was used to lyse cells. Plates were scraped with a plastic scraper to remove any adherent cells and to collect lysate. Lysate was then transferred to an epitube and incubated at 4°C for 30 minutes. The lysate was then sonicated 3 times for 5 seconds, with a 5 second break in between rounds (amplitude 20) and treated with 500 units of Benzonase for 1 hour to digest DNA. The lysate was centrifuged at 21,000 x*g* for 15 minutes, and supernatant containing proteins collected for immunoprecipitation. Supernatant was incubated overnight separately with either 5 μg Anti-RPA2 antibody (polyclonal rabbit, (Thermo Fisher Scientific, Catalogue #A300–244A), 5 μg Anti-RPA4 antibody (sheep serum, homemade), 5μg Normal Rabbit control IgG (New England Biolabs Catalogue #2729S), or 5 μg Normal Sheep control IgG (Thermo Fisher Scientific, Catalogue #31243). Following overnight incubation, the Protein G magnetic Dynabeads Immunoprecipitation kit (Thermo Fisher Scientific, Catalogue #10007D) was used for co-IP as per the manufacturer's instructions. Briefly, magnetic Protein G Dynabeads were washed 3 times in modified RIPA lysis buffer (2 minutes per wash with gentle rotation) and added to the supernatant-antibody mixture. This mixture was incubated for 1 to 3 hours with gentle rotation/nutation and then beads washed 3 times with the kit wash buffer to remove unspecific binding of protein to beads. After washes, proteins were eluted from beads by adding modified lysis buffer, 4x NuPAGE LDS Sample Loading dye (ThermoFisher Scientific, catalogue #NP0007) and 10x NuPAGE Sample Reducing Agent (ThermoFisher Scientific; catalogue #NP0004) and boiling at 70°C for 10 minutes. Proteins were assessed by western blot as described above.

Co-immunoprecipitation using purified protein complexes.—1 μg of purified RPA or Alt-RPA complexes were incubated with increasing amounts (1 μg, 2 μg, 5 μg) of purified MutSα or MutSβ complexes which both contained a FLAG-tagged MSH2 subunit. Purified proteins were mixed into hypotonic binding buffer (20 mM HEPES PH7.8, 5 mM KCl, 1.5 mM MgCl2) and incubated for 1 hour at 4°C. After allowing proteins to interact, IP was performed using Pierce Anti-FLAG magnetic beads (Millipore Sigma, Catalogue #M8823) using the manufacturer's instructions. In brief, magnetic beads were washed twice with hypotonic buffer (5 minutes per wash with gentle rotation) and the protein mixture was added to the beads. The protein mixture was incubated with beads for 2 hours at 4 ° with gentle rotation. Beads were then washed three times with PBS to remove non-specifically bound proteins (5 minutes per wash with gentle rotation). Proteins were eluted from the beads by adding hypotonic buffer, 4x NuPAGE LDS Sample Loading dye (ThermoFisher

Scientific, catalogue #NP0007) and 10x NuPAGE Sample Reducing Agent (ThermoFisher Scientific; catalogue #NP0004) and boiling the sample at 70°C for 10 minutes. Proteins were assessed by western blot as described above.

Overexpression or RPA2 or RPA4 in HT1080-(CAG)850 cells and small-pool PCR.—Cell were maintained in six-well tissue culture plates in DMEM supplemented with 10% FBS. At 60–70% confluency, cells were transfected with 250 ng of the appropriate plasmid DNA expressing RPA2 or RPA4 using Lipofectamine 2000 as per the manufacturer's instructions (ThermoFisher Scientific; catalogue #11668019). At 24 h post-transfection, the media was removed and fresh DMEM supplemented with 0.5% FBS was added to the cells. After ten days of culture under serum starvation, DNA was extracted using the Gentra Puregene Kit (Qiagen; catalogue #158043). The (CTG)•(CAG) repeats were sized by small pool PCR with the input of 2.0–2.7 genome equivalents, followed by Southern blot detection, as previously described[27,57]. Briefly, amplification of diluted genomic DNA was performed using an Expand Long Template PCR System (Sigma Aldrich; Catalogue # 11681834001) with primers 5′-ACCCTAGAACTGTCTTCGACTCC-3′ and 5′-TTCCCGAGTAAGCAGGCAGAG-3′ through a total of 24 cycles. Detection of amplicons was done through Southern blot using a digoxigenin-labeled (CAG)7 locked nucleic acid probe. At least 250 alleles were analyzed in each group. A $\chi 2$ test was performed to compare the frequencies of expanded, unchanged, and contracted alleles in each set of experiments as described previously[27,57].

AAV-overexpression in mouse brains.—AAV vector plasmids contained cDNA for either EGFP-Rpa1 (cloned by reverse transcription and PCR of RNA from mouse primary cortical neurons) or EGFP under control of a CMV promoter. AAV vectors generated after transient transfection of HEK293T cells, and the recombinant virus was isolated from two sequential continuous CsCl gradients. AAV were injected into 5-week-old mice (randomized method for injection of a particular virus vector) into the subarachnoid space above the cerebellar surface. Mice were anesthetized intraperitoneally with Nembutal and mounted on a stereotaxic apparatus (Narishige). Forehead was tilted down 20°, and a 1 mm diameter old was made at −9.2 mm from bregma, ± 0 mm lateral to the midline. A glass syringe was inserted into the hole (along the occipital bone, 3.5 mm from the hole) and 8 μL of AAV virus solution (~1000 particles) were injected in four orientations (60, 90, 270 and 330° clockwise rotation from the posterior to anterior line, with 2 μL injected at each orientation at a rate of 0.5 ul/min).

Fragment length analysis (capillary gel electrophoresis).—Genomic DNA was collected from mouse brain tissues following homogenisation with a MagNA Lyser Instrument (Roche; item #03358968001) (same method used in 2.3.1) and phenol:chloroform extraction with ethanol precipitation. Amplification was performed using the Expand Long Template PCR system (Roche Diagnostic, catalogue #11681834001) with 5% DMSO added with the following primers: Forward primer: 5'-CCGGAGCCCTGCTGAGGTG-3' and Reverse primer: 5'-CCAGACGCCGGGACACAAGGCTGAG-3'. The PCR cycles were as follows: 1) 1× 95°C for 5 minutes, 2) 35x (95°C for 30 seconds, then 64°C for 30 seconds, then 72°C for 5

minutes), 3) 72°C for 10 minutes, 4) infinite hold at 4°C. PCR products were denatured with HiDi formamide and boiling at 95°C for 5 minutes, and then processed by capillary gel electrophoresis with size markers on a 3130xl Genetic Analyzer (Applied Biosystems). Peak Scanner 2 software was used to visualise the repeat sizes and repeat lengths were calculated by subtracting the length of non-repeat sequence in the PCR product and dividing by 3 (representing one CAG unit).

Immunohistochemistry (IHC).—Whole mouse brains were fixed in 4% paraformaldehyde for 12–16h, embedded in paraffin, and 5 μm sagittal sections were obtained using a microtome. Xylenes was used to deparaffinize the sections, which were then rehydrated in serial dilutions of ethanol (100, 90, 80, 70%). Slides were then microwaved in 0.01M citrate buffer (pH 6.0) at 120°C for 15 min for antigen retrieval. Sections were blocked with 10% normal donkey serum in 1xPBS for 1 hour at room temperature. Sections were then incubated with primary antibodies in blocking solution for 1 hour a room temperature, washed 2x with 1xPBST, and then incubated with secondary antibodies in blocking solution for 1 hour at room temperature, washed 2x with 1xPBST, Nuclei were stained with 0.2 μg/mL DAPI in PBS (DOJINDO Laboratories, catalogue #D523), and then mounted. Images were acquired using a FV1200IX83 Olympus confocal microscope and BZ-X800 Keyence All-in-one fluorescence microscope. Fluorescent signal was quantified by ImageJ.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of RPA1–4 expression levels (protein and RNA) in Figure 1, Supplementary Figure S1, and Supplementary Figure S8: quantification and statistical details can be found in the Results and figure legends for Figure 1 and Supplementary Figure S1. In brief, RNA expression analysis was assessed via ddPCR on cohort 1 comprised of n = 7 HD patients (striatum, cerebellum, and frontal pole), n = 7 unaffected individuals (striatum, cerebellum, and frontal pole), and n = 3 SCA1 patients (cerebellum). Expression levels of RPA1–4 were normalised to levels of actin. Independent expression analysis qRT-PCR was assessed on a second cohort of n = 3 HD patients (striatum and cerebellum), and n = 3 unaffected individuals (striatum and cerebellum). In both cases, n = total number of patient tissues assessed, with each analysis being conducted in triplicate per patient per assay. ddPCR was also used to assess RPA1–3 RNA expression levels in SCA1 mice (+/− AAV treatment), where n = 3 AAV-GFP control SCA1 mice, and n = 5 AAV-GFP-RPA1 SCA1 mice. For ddPCR, data was plotted as box plots, with the black line representing the median and the whiskers corresponding to Tukey whisker extent. Each individual patient's or mouse's expression datapoints (average of n = 3 technical replicates) were also plotted to demonstrate dispersion. Statistical test: unpaired t-test comparing means with p-values reported within the plot and in-text. For qRT-PCR, data was plotted as bar graphs, with the bars corresponding to the mean expression data and the error bars corresponding to the standard deviation from the mean. Statistical test: unpaired t-test comparing means using GraphPad Prism Software with p-values reported within the plot and in-text.

For protein expression analysis, both HD patient cohorts were plotted and quantified as a single group as tissues was not available for protein expression analysis from all patients. For protein expression analysis: n = 6 HD patients (striatum), n = 6 unaffected individuals (striatum), n = 9 HD patients (cerebellum), n = 9 unaffected individuals (cerebellum), n = 2 SCA1 patients (cerebellum). For mouse protein expression analysis; n = 4 zQ175 HD mice (11-week striatum and cerebellum), n = 5 WT mice (11-week striatum and cerebellum), n = 4 zQ175 HD mice (48-week striatum and cerebellum), n = 3 WT mice (48-week striatum and cerebellum), n = 4 zQ175 HD (82-week striatum and cerebellum), n = 4 WT mice (82-week striatum and cerebellum). For cell line protein expression analysis; n = 5 control WT cell lines, n = 2 HD cell lines (Q40 and Q43), n = 1 SCA1 cell line (Q45), and n = 1 SCA3 cell line (Q53). Western blots were quantified via densitometry of RPA2 and RPA4 band intensity normalised to the actin loading control band intensity. Data was plotted as dot plots, with the horizontal band corresponding to the mean, and the error bars corresponding to the standard deviation. For cell lines, westerns were performed in duplicate or triplicate for each individual line, with all replicates being plotted for HD, SCA1, and SCA3 cell lines; replicates were averaged for control lines and each line was plotted together. Individual datapoints for each sample was also plotted to demonstrate dispersion from the mean. Statistical test: unpaired t-test comparing means using GraphPad Prism Software with p-values reported within the figures and in-text.

Correlation of RPA2 and RPA4 RNA expression with HD phenotypes in Supplementary Figure S2: quantification and statistical details, and patient information, can be found in the Results, figure legends for Supplementary Figure S2, Supplementary Data S1, and Supplementary Table S1. In brief, ddPCR RPA2 and RPA4 RNA expression data from cohort 1 HD patients and matched unaffected individuals (n = 7 HD patients striatum, cerebellum, and frontal pole, n = 7 unaffected individuals striatum, cerebellum, and frontal pole, n = 3 technical replicates for each ddPCR reaction), was correlated with clinical features as specified in Supplementary Figure S2. Correlation was carried out using R version 4.2.1. Linear models were fit to the data using the ‘ggpmisc’ R package (https://cran.r-project.org/web/packages/ggpmisc/index.html), using the stat_poly_line() and stat_poly_eq() functions. Coefficient of determination are represented as R^2 values reported within each plot. Boxplots were plotted for RPA2 and RPA4 RNA expression as a function of Vonsattel neuropathological grades, with the black line representing the median and the whiskers corresponding to Tukey whisker extent. Statistical test: unpaired t-test comparing means using GraphPad Prism Software, with p-values reported within the figures and in-text.

Quantification of *in vitro* repair reactions in Figure 2 and Supplementary Figure S4: quantification details can be found in the Results, Supplementary Data S1, and figure legends for Figure 2 and Supplementary Figure S4. In brief, Southern blots were quantified via densitometry of individual band intensity normalised to the whole lane intensity and reported as a percentage (i.e. the proportion of repaired/intermediate DNA relative to total DNA in the reaction). Each reaction was performed in triplicates (n = 3 technical replicates), the data was plotted as bar graphs with the bars corresponding to the mean and the error bars corresponding to the standard deviation from the mean. For Figure 2E–J, mean repaired and intermediate quantifications were reported as individual values rather than bar graphs.

Quantification of kinetic FRET assays in Figure 3 and Supplementary Figures S4 and S5: quantification and statistical details can be found in the Results, and figure legends of Figure 3 and Supplementary Figures S4 and S5. In brief, binding and melting curves were generated by detecting FRET intensity over time according to the equation outlined in Supplementary Figure S5C, which also normalises the specific intensity to a value between 1 and 0. Equilibrium bindings curves display the mean intensities as a function of protein concentration collected over triplicate experiments (n = 3 technical replicates), with each concentration point plotted as the legend-specified shape representing the mean and bars representing the standard deviation from the mean. Melting curves represent the mean intensities collected over triplicate experiments (n = 3 technical replicates), with lighter coloured dots/shading above and below each value representing the standard deviation from the mean. The 50% melting time was also plotted in a bar graph, with the bar representing the mean from triplicate experiments (n = 3 technical replicates) and the error bars representing the standard deviation from the mean. Statistical test: unpaired t-test comparing means using GraphPad Prism Software with p-values reported within the figures and in-text.

Quantification of FAN1 nuclease assays in Figure 3 and Supplementary Figure S5: quantification and statistical details can be found in the Results, and figure legends of Figure 3 and Supplementary Figure S5. In brief, FAN1 nuclease digestion of DNA was conducted using fluorescent oligonucleotides, run on a urea denaturing gel, and fluorescence was visualised. Gels were quantified via densitometry of individual band intensity normalised to the whole lane intensity and reported as a percentage (i.e. the proportion of cleaved DNA relative to total DNA in the reaction). Each reaction was performed in triplicates (n = 3 technical replicates) and the data was plotted as bar graphs with the bars corresponding to the mean and the error bars corresponding to the standard deviation from the mean. Statistical test: unpaired t-test comparing means using GraphPad Prism Software with p-values reported within the figures and in-text.

BioID datasets in Supplementary Figure 4 and Supplementary Figures S6 and S7: quantification and statistical details can be found in the Results, Supplementary Data S1, and figure legends of Figure 4 and Supplementary Figures S6 and S7. GO analysis of enriched genes was conducted using the The Gene Ontology Resource[96,97]. GO terms are provided with a p-value determined by a Fishers exact test. Data is represented as bar graphs with the GO terms as a function of their associated -log10(p-value). Raw mass spectrometry data was processed and significance of enriched genes was determined using parameters set by Prostar software (Proteomics statistical analysis with R) as specified in the STAR Methods section. Each individual gene's enrichment for a particular subunit within the BioID datasets were represented as dot plots, with values specified as per the figure legend. BioID data was derived from n = 3 biological replicates.

Quantification of small-pool PCR in Figure 5: quantification and statistical details can be found in the Results, and figure legend of Figure 5. In brief, following Southern blot detection of amplicons following small-pool PCR, individual band repeat sizes were calculated based on their migration relative to DNA size markers. At least n = 250 mutant

alleles were sized per group. Statistical test: $\chi 2$ (Chi-square) test using Microsoft Excel as per the following, with p-values reported within the figures and in-text:

$\chi 2$ tests for RPA2 overexpression versus Transfection Control: "Induced contractions, p=0.00358" = contracted alleles versus unchanged & expanded alleles ($\chi 2$ test by 2 rows [RPA2 and Transfection Control], 2 columns [contracted and unchanged + expanded])

"Inhibited expansions, p<0.0001" = expanded alleles versus unchanged & contracted alleles ($\chi 2$ test by 2 rows [RPA2 and Transfection Control], 2 columns [expanded and unchanged + contracted])

"Overall contraction bias, p<0.001" = contracted alleles versus unchanged alleles versus expanded alleles ($\chi 2$ test by 2 rows [RPA2 and Transfection Control], 3 columns [contracted, unchanged, and expanded])

$\chi 2$ tests for RPA4 overexpression versus Transfection Control: "No induced contractions, p=0.4" = contracted alleles versus unchanged & expanded alleles ($\chi 2$ test by 2 rows [RPA4 and Transfection Control], 2 columns [contracted and unchanged + expanded])

"Promoted expansion, p=0.00425" = expanded alleles versus unchanged & contracted alleles (chi-square test by 2 rows [RPA4 and Transfection Control], 2 columns [expanded and unchanged + contracted])

"Overall contraction bias, p=0.012" = contracted alleles versus unchanged alleles versus expanded alleles ($\chi 2$ test by 2 rows [RPA4 and Transfection Control], 3 columns [contracted, unchanged, and expanded])

Quantification of Fragment Length Analysis in Figure 5 and Supplementary Figure S8: quantification and statistical details can be found in the Results, and figure legends of Figure 5 and Supplementary Figure S8. Instability indices and repeat lengths were calculated as described in the STAR Methods. Indices were calculated on the striatum and cerebellum from n = 3 AAV-GFP control SCA1 mice, and n = 5 AAV-GFP-RPA1 SCA1 mice, with n = 3 technical replicates per mouse. Statistical test: Mann–Whitney *U* test using Microsoft Excel.

Fluorescent Intensity Analysis in Figure 7 and Supplementary Figure S9: quantification and statistical details can be found in the Results, and figure legends of Figure 7 and Supplementary Figure S9. Fluorescent intensity was calculated using ImageJ Software, with fluorescent intensity being normalised to the area being quantified. Fluorescent intensity was assessed from the striatum and cerebellum of n = 3 AAV-GFP control SCA1 mice, and n = 3 AAV-GFP-RPA1 SCA1 mice, with n = 3 technical replicates per mouse. Whole-brain analysis was conducted on n = 3 technical replicates per mouse, and individual cell analysis was conducted on at least n = 100 cells/nuclei were quantified per replicate (300 cells/nuclei assessed in total per mouse). The mean of each replicate per mouse was plotted in bar graphs, with the bars representing the mean between mice, and the bars representing the standard deviation from the mean. Each individual mouse was also plotted to demonstrate

dispersion. Statistical test: unpaired t-test comparing means using GraphPad Prism Software with p-values reported in the figure legends and in-text.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We acknowledge family members who made this work possible through tissue donations and support. We acknowledge the Neurological Foundation Human Brain Bank, Centre for Brain Research, the University of Auckland, the National Ataxia Foundation and the University of Florida Center for NeuroGenetics for the HD and SCA1 brain. We thank Drs. Guy Rouleau and Elise Heon for cells. We thank John D. Cleary for critically reading the manuscript. TGD and ALD are supported by Postdoctoral Researcher Fellowships from the Hereditary Disease Foundation and the Fox Family Foundation. DEL is supported by a Postdoctoral Researcher Fellowship from the Hereditary Disease Foundation. RJH is supported by the Hereditary Disease Foundation and a member of the Structural Genomics Consortium. The Structural Genomics Consortium is a registered charity (no: 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196], EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUbOPEN grant 875510], Janssen, Merck KGaA (aka EMD in Canada and US), Pfizer, and Takeda. MR is supported by a T32 Postdoctoral Fellowship through a National Institutes of Health/National Cancer Institute training grant in Free Radicals and Radiation Biology (CA078586). FMB is supported by the Canadian Institutes of Health Research (BMB-398925). EIC is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-05559). MS is supported by the National Institutes of Health/ National Institute of General Medical Sciences (R35 GM131704), ARLS is supported by the National Institutes of Health (R35 NS122140). J-YM is supported by the Canadian Institutes of Health Research (FRN-388879). HO is supported by Ministry of Education, Culture, Sports, Science, and Technology of Japan (22110001/ 22110002), Japanese Society for the Promotion of Science (16H02655, 19H01042, 22H00464) and Japan Agency for Medical Research and Development (Brain/MINDS). CEP is supported by the Canadian Institutes of Health Research (FRN-148910; FRN-173282), the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016–08355 RGPIN-2016–06355/498835), Petroff Family Foundation, Tribute Communities, Marigold Foundation, Kazman Family Foundation, and Fox Family Foundation. J-YM. holds a Tier 1 Canada Research Chair in DNA Repair and Cancer Therapeutics. CEP holds a Tier 1 Canada Research Chair in Disease-Associated Genome Instability.

Inclusion and diversity:

We support inclusive, diverse, and equitable conduct of research.

REFERENCES

1. Gall-Duncan T, Sato N, Yuen RKC, and Pearson CE (2022). Advancing genomic technologies and clinical awareness accelerates discovery of disease-associated tandem repeat sequences. Genome Res. 32, 1–27. 10.1101/gr.269530.120. [PubMed: 34965938]
2. Kacher R, Lejeune F-X, Noël S, Cazeneuve C, Brice A, Humbert S, and Durr A (2021). Propensity for somatic expansion increases over the course of life in Huntington disease. eLife 10, e64674. 10.7554/eLife.64674. [PubMed: 33983118]
3. Mouro Pinto R, Arning L, Giordano JV, Razghandi P, Andrew MA, Gillis T, Correia K, Mysore JS, Grote Urtubey D-M, Parwez CR, et al. (2020). Patterns of CAG repeat instability in the central nervous system and periphery in Huntington's disease and in spinocerebellar ataxia type 1. Hum. Mol. Genet. 29, 2551–2567. 10.1093/hmg/ddaa139. [PubMed: 32761094]
4. Swami M, Hendricks AE, Gillis T, Massood T, Mysore J, Myers RH, and Wheeler VC (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum. Mol. Genet. 18, 3039–3047. 10.1093/hmg/ddp242. [PubMed: 19465745]
5. Jones L, Wheeler VC, and Pearson CE Special Issue: DNA repair and somatic repeat expansion in Huntington's disease. J. Huntingt. Dis. 10, 3–5. 10.3233/JHD-219001.

6. Wright GEB, Black HF, Collins JA, Gall-Duncan T, Caron NS, Pearson CE, and Hayden MR (2020). Interrupting sequence variants and age of onset in Huntington's disease: clinical implications and emerging therapies. Lancet Neurol. 19, 930–939. 10.1016/S1474-4422(20)30343-4. [PubMed: 33098802]
7. Benn CL, Gibson KR, and Reynolds DS (2021). Drugging DNA damage repair pathways for trinucleotide repeat expansion diseases. J. Huntingt. Dis. 10, 203–220. 10.3233/JHD-200421.
8. López Castel A, Cleary JD, and Pearson CE (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. Nat. Rev. Mol. Cell Biol. 11, 165–170. 10.1038/nrm2854. [PubMed: 20177394]
9. Deshmukh AL, Porro A, Mohiuddin M, Lanni S, Panigrahi GB, Caron M-C, Masson J-Y, Sartori AA, and Pearson CE (2021). FAN1, a DNA repair nuclease, as a modifier of repeat expansion disorders. J. Huntingt. Dis. 10, 95–122. 10.3233/JHD-200448.
10. Trinh J, Lüth T, Schaake S, Laabs B-H, Schlüter K, Laβ J, Pozojevic J, Tse R, König I, Dominic Jamora R, et al. (2022). Mosaic divergent repeat interruptions in XDP influence repeat stability and disease onset. Brain J. Neurol, awac160. 10.1093/brain/awac160.
11. Flower M, Lomeikaite V, Ciosi M, Cumming S, Morales F, Lo K, Hensman Moss D, Jones L, Holmans P, Monckton DG, et al. (2019). MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. Brain 142, 1876–1886. 10.1093/brain/awz115. [PubMed: 31216018]
12. Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C, Stevanin G, Koutsis G, Karadima G, Panas M, et al. (2016). DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. Ann. Neurol. 79, 983–990. 10.1002/ana.24656. [PubMed: 27044000]
13. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium (2015). Identification of genetic factors that modify clinical onset of Huntington's disease. Cell 162, 516–526. 10.1016/j.cell.2015.07.003. [PubMed: 26232222]
14. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium (2019). CAG repeat not polyglutamine length determines timing of Huntington's disease onset. Cell 178, 887–900.e14. 10.1016/j.cell.2019.06.036. [PubMed: 31398342]
15. Laabs B-H, Klein C, Pozojevic J, Domingo A, Brüggemann N, Grütz K, Rosales RL, Jamora RD, Saranza G, Diesta CCE, et al. (2021). Identifying genetic modifiers of age-associated penetrance in X-linked dystonia-parkinsonism. Nat. Commun. 12, 3216. 10.1038/s41467-021-23491-4. [PubMed: 34050153]
16. Lee J-M, Huang Y, Orth M, Gillis T, Siciliano J, Hong E, Mysore JS, Lucente D, Wheeler VC, Seong IS, et al. (2022). Genetic modifiers of Huntington disease differentially influence motor and cognitive domains. Am. J. Hum. Genet. 109, 885–899. 10.1016/j.ajhg.2022.03.004. [PubMed: 35325614]
17. Tomé S, Manley K, Simard JP, Clark GW, Slean MM, Swami M, Shelbourne PF, Tillier ERM, Monckton DG, Messer A, et al. (2013). MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLOS Genet. 9, e1003280. 10.1371/journal.pgen.1003280. [PubMed: 23468640]
18. Manley K, Shirley TL, Flaherty L, and Messer A (1999). Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat. Genet. 23, 471–473. 10.1038/70598. [PubMed: 10581038]
19. Pinto RM, Dragileva E, Kirby A, Lloret A, Lopez E, St Claire J, Panigrahi GB, Hou C, Holloway K, Gillis T, et al. (2013). Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet. 9, e1003930. 10.1371/journal.pgen.1003930. [PubMed: 24204323]
20. Roy JCL, Vitalo A, Andrew MA, Mota-Silva E, Kovalenko M, Burch Z, Nhu AM, Cohen PE, Grabczyk E, Wheeler VC, et al. (2021). Somatic CAG expansion in Huntington's disease is dependent on the MLH3 endonuclease domain, which can be excluded via splice redirection. Nucleic Acids Res. 49, 3907–3918. 10.1093/nar/gkab152. [PubMed: 33751106]
21. Mirkin SM (2007). Expandable DNA repeats and human disease. Nature 447, 932–940. 10.1038/nature05977. [PubMed: 17581576]

22. Brown RE, and Freudenreich CH (2021). Structure-forming repeats and their impact on genome stability. Curr. Opin. Genet. Dev. 67, 41–51. 10.1016/j.gde.2020.10.006. [PubMed: 33279816]

23. Pearson CE, Nichol Edamura K, and Cleary JD (2005). Repeat instability: mechanisms of dynamic mutations. Nat. Rev. Genet. 6, 729–742. 10.1038/nrg1689. [PubMed: 16205713]

24. Pearson CE, Zorbas H, Price GB, and Zannis-Hadjopoulos M (1996). Inverted repeats, stem-loops, and cruciforms: Significance for initiation of DNA replication. J. Cell. Biochem. 63, 1–22. 10.1002/(SICI)1097-4644(199610)63:1<1::AID-JCB1>3.0.CO;2-3. [PubMed: 8891900]

25. Petruska J, Arnheim N, and Goodman MF (1996). Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. Nucleic Acids Res. 24, 1992–1998. 10.1093/nar/24.11.1992. [PubMed: 8668527]

26. Axford MM, Wang Y-H, Nakamori M, Zannis-Hadjopoulos M, Thornton CA, and Pearson CE (2013). Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues. PLOS Genet. 9, e1003866. 10.1371/journal.pgen.1003866. [PubMed: 24367268]

27. Nakamori M, Panigrahi GB, Lanni S, Gall-Duncan T, Hayakawa H, Tanaka H, Luo J, Otabe T, Li J, Sakata A, et al. (2020). A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo. Nat. Genet. 52, 146–159. 10.1038/s41588-019-0575-8. [PubMed: 32060489]

28. Caldwell CC, and Spies M (2020). Dynamic elements of replication protein A at the crossroads of DNA replication, recombination, and repair. Crit. Rev. Biochem. Mol. Biol. 55, 482–507. 10.1080/10409238.2020.1813070. [PubMed: 32856505]

29. Li X, and Heyer W-D (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 18, 99–113. 10.1038/cr.2008.1. [PubMed: 18166982]

30. Oakley GG, and Patrick SM (2010). Replication protein A: directing traffic at the intersection of replication and repair. Front. Biosci. Landmark Ed. 15, 883–900. 10.2741/3652. [PubMed: 20515732]

31. Chen R, and Wold MS (2014). Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. BioEssays News Rev. Mol. Cell. Dev. Biol. 36, 1156–1161. 10.1002/bies.201400107.

32. Wold MS (1997). REPLICATION PROTEIN A: A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaryotic DNA Metabolism. Annu. Rev. Biochem. 66, 61–92. 10.1146/annurev.biochem.66.1.61. [PubMed: 9242902]

33. Chen H, Lisby M, and Symington LS (2013). RPA coordinates DNA end resection and prevents formation of DNA hairpins. Mol. Cell 50, 589–600. 10.1016/j.molcel.2013.04.032. [PubMed: 23706822]

34. Safa L, Delagoutte E, Petruseva I, Alberti P, Lavrik O, Riou J-F, and Saintomé C (2014). Binding polarity of RPA to telomeric sequences and influence of G-quadruplex stability. Biochimie 103, 80–88. 10.1016/j.biochi.2014.04.006. [PubMed: 24747047]

35. Ray S, Bandaria JN, Qureshi MH, Yildiz A, and Balci H (2014). G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. Proc. Natl. Acad. Sci. 111, 2990–2995. 10.1073/pnas.1321436111. [PubMed: 24516170]

36. Ray S, Qureshi MH, Malcolm DW, Budhathoki JB, Çelik U, and Balci H (2013). RPA-Mediated Unfolding of Systematically Varying G-Quadruplex Structures. Biophys. J. 104, 2235–2245. 10.1016/j.bpj.2013.04.004. [PubMed: 23708363]

37. Qureshi MH, Ray S, Sewell AL, Basu S, and Balci H (2012). Replication Protein A Unfolds G-Quadruplex Structures with Varying Degrees of Efficiency. J. Phys. Chem. B 116, 5588–5594. 10.1021/jp300546u. [PubMed: 22500657]

38. Keshav KF, Chen C, and Dutta A (1995). Rpa4, a homolog of the 34-kilodalton subunit of the replication protein A complex. Mol. Cell. Biol. 15, 3119–3128. 10.1128/MCB.15.6.3119. [PubMed: 7760808]

39. Mason AC, Haring SJ, Pryor JM, Staloch CA, Gan TF, and Wold MS (2009). An alternative form of replication protein a prevents viral replication in vitro. J. Biol. Chem. 284, 5324–5331. 10.1074/jbc.M808963200. [PubMed: 19116208]

40. Haring SJ, Humphreys TD, and Wold MS (2010). A naturally occurring human RPA subunit homolog does not support DNA replication or cell-cycle progression. Nucleic Acids Res. 38, 846–858. 10.1093/nar/gkp1062. [PubMed: 19942684]

41. Menezo YJ, Russo G, Tosti E, Mouatassim SE, and Benkhalifa M (2007). Expression profile of genes coding for DNA repair in human oocytes using pangenomic microarrays, with a special focus on ROS linked decays. J. Assist. Reprod. Genet. 24, 513–520. 10.1007/s10815-007-9167-0. [PubMed: 17899356]

42. Kemp MG, Mason AC, Carreira A, Reardon JT, Haring SJ, Borgstahl GEO, Kowalczykowski SC, Sancar A, and Wold MS (2010). An alternative form of replication protein a expressed in normal human tissues supports DNA repair. J. Biol. Chem. 285, 4788–4797. 10.1074/jbc.M109.079418. [PubMed: 19996105]

43. Mason AC, Roy R, Simmons DT, and Wold MS (2010). Functions of alternative replication protein A in initiation and elongation. Biochemistry 49, 5919–5928. 10.1021/bi100380n. [PubMed: 20545304]

44. Choi J-H, Lindsey-Boltz LA, Kemp M, Mason AC, Wold MS, and Sancar A (2010). Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. Proc. Natl. Acad. Sci. U. S. A. 107, 13660–13665. 10.1073/pnas.1007856107. [PubMed: 20616048]

45. Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, and Richardson EP (1985). Neuropathological classification of Huntington's disease. J. Neuropathol. Exp. Neurol. 44, 559–577. 10.1097/00005072-198511000-00003. [PubMed: 2932539]

46. Singh-Bains MK, Mehrabi NF, Sehji T, Austria MDR, Tan AYS, Tippett LJ, Dragunow M, Waldvogel HJ, and Faull RLM (2019). Cerebellar degeneration correlates with motor symptoms in Huntington disease. Ann. Neurol. 85, 396–405. 10.1002/ana.25413. [PubMed: 30635944]

47. Rüb U, Hoche F, Brunt ER, Heinsen H, Seidel K, Del Turco D, Paulson HL, Bohl J, von Gall C, Vonsattel J-P, et al. (2013). Degeneration of the cerebellum in Huntington's disease (HD): possible relevance for the clinical picture and potential gateway to pathological mechanisms of the disease process. Brain Pathol. Zurich Switz. 23, 165–177. 10.1111/j.1750-3639.2012.00629.x.

48. Johnson EB, Ziegler G, Penny W, Rees G, Tabrizi SJ, Scahill RI, and Gregory S (2021). Dynamics of Cortical Degeneration Over a Decade in Huntington's Disease. Biol. Psychiatry 89, 807–816. 10.1016/j.biopsych.2020.11.009. [PubMed: 33500176]

49. Padron-Rivera G, Diaz R, Vaca-Palomares I, Ochoa A, Hernandez-Castillo CR, and Fernandez-Ruiz J (2021). Cerebellar Degeneration Signature in Huntington's Disease. Cerebellum Lond. Engl. 20, 942–945. 10.1007/s12311-021-01256-5.

50. Rees EM, Farmer R, Cole JH, Haider S, Durr A, Landwehrmeyer B, Scahill RI, Tabrizi SJ, and Hobbs NZ (2014). Cerebellar abnormalities in Huntington's disease: a role in motor and psychiatric impairment? Mov. Disord. Off. J. Mov. Disord. Soc. 29, 1648–1654. 10.1002/mds.25984.

51. Koscik TR, Sloat L, van der Plas E, Joers JM, Deelchand DK, Lenglet C, Öz G, and Nopoulos PC (2020). Brainstem and striatal volume changes are detectable in under 1 year and predict motor decline in spinocerebellar ataxia type 1. Brain Commun. 2, fcaa184. 10.1093/braincomms/fcaa184. [PubMed: 33409488]

52. Reetz K, Costa AS, Mirzazade S, Lehmann A, Juzek A, Rakowicz M, Boguslawska R, Schöls L, Linnemann C, Mariotti C, et al. (2013). Genotype-specific patterns of atrophy progression are more sensitive than clinical decline in SCA1, SCA3 and SCA6. Brain J. Neurol. 136, 905–917. 10.1093/brain/aws369.

53. Klaes A, Reckziegel E, Franca MC, Rezende TJR, Vedolin LM, Jardim LB, and Saute JA (2016). MR Imaging in Spinocerebellar Ataxias: A Systematic Review. AJNR Am. J. Neuroradiol. 37, 1405–1412. 10.3174/ajnr.A4760. [PubMed: 27173364]

54. Madeira F, Park Y. mi, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641. 10.1093/nar/gkz268. [PubMed: 30976793]

55. The PyMOL Molecular Graphics System (2015).

56. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, and Baker NA (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for

molecular simulations. Nucleic Acids Res. 35, W522–W525. 10.1093/nar/gkm276. [PubMed: 17488841]

57. Deshmukh AL, Caron M-C, Mohiuddin M, Lanni S, Panigrahi GB, Khan M, Engchuan W, Shum N, Faruqui A, Wang P, et al. (2021). FAN1 exo- not endo-nuclease pausing on disease-associated slipped-DNA repeats: A mechanism of repeat instability. Cell Rep. 37, 110078. 10.1016/j.celrep.2021.110078. [PubMed: 34879276]
58. Hou C, Zhang T, Tian L, Huang J, Gu L, and Li G-M (2011). The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins. Cell Biosci. 1, 11. 10.1186/2045-3701-1-11. [PubMed: 21711735]
59. Panigrahi GB, Lau R, Montgomery SE, Leonard MR, and Pearson CE (2005). Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. Nat. Struct. Mol. Biol. 12, 654–662. 10.1038/nsmb959. [PubMed: 16025129]
60. Panigrahi GB, Slean MM, Simard JP, Gileadi O, and Pearson CE (2010). Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. Proc. Natl. Acad. Sci. U. S. A. 107, 12593–12598. 10.1073/pnas.0909087107. [PubMed: 20571119]
61. Tian L, Hou C, Tian K, Holcomb NC, Gu L, and Li G-M (2009). Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J. Biol. Chem. 284, 20452–20456. 10.1074/jbc.C109.014977. [PubMed: 19525234]
62. Pearson CE, Tam M, Wang Y-H, Montgomery SE, Dar AC, Cleary JD, and Nichol K (2002). Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. Nucleic Acids Res. 30, 4534–4547. 10.1093/nar/gkf572. [PubMed: 12384601]
63. Sibenaller ZA, Sorensen BR, and Wold MS (1998). The 32- and 14-kilodalton subunits of replication protein A are responsible for species-specific interactions with single-stranded DNA. Biochemistry 37, 12496–12506. 10.1021/bi981110+. [PubMed: 9730822]
64. Panigrahi GB, Slean MM, Simard JP, and Pearson CE (2012). Human mismatch repair protein hMutLα is required to repair short slipped-DNAs of trinucleotide repeats. J. Biol. Chem. 287, 41844–41850. 10.1074/jbc.M112.420398. [PubMed: 23086927]
65. Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L, and Li G-M (2005). Reconstitution of 5'-directed human mismatch repair in a purified system. Cell 122, 693–705. 10.1016/j.cell.2005.06.027. [PubMed: 16143102]
66. Genschel J, and Modrich P (2009). Functions of MutLalpha, replication protein A (RPA), and HMGB1 in 5'-directed mismatch repair. J. Biol. Chem. 284, 21536–21544. 10.1074/jbc.M109.021287. [PubMed: 19515846]
67. Guo S, Zhang Y, Yuan F, Gao Y, Gu L, Wong I, and Li G-M (2006). Regulation of replication protein A functions in DNA mismatch repair by phosphorylation. J. Biol. Chem. 281, 21607–21616. 10.1074/jbc.M603504200. [PubMed: 16731533]
68. Yuan F, Gu L, Guo S, Wang C, and Li G-M (2004). Evidence for involvement of HMGB1 protein in human DNA mismatch repair. J. Biol. Chem. 279, 20935–20940. 10.1074/jbc.M401931200. [PubMed: 15014079]
69. Genschel J, and Modrich P (2003). Mechanism of 5'-directed excision in human mismatch repair. Mol. Cell 12, 1077–1086. 10.1016/s1097-2765(03)00428-3. [PubMed: 14636568]
70. Ramilo C, Gu L, Guo S, Zhang X, Patrick SM, Turchi JJ, and Li G-M (2002). Partial reconstitution of human DNA mismatch repair in vitro: characterization of the role of human replication protein A. Mol. Cell. Biol. 22, 2037–2046. 10.1128/MCB.22.7.2037-2046.2002. [PubMed: 11884592]
71. Slean MM, Reddy K, Wu B, Nichol Edamura K, Kekis M, Nelissen FHT, Aspers RLEG, Tessari M, Schärer OD, Wijmenga SS, et al. (2013). Interconverting conformations of slipped-DNA junctions formed by trinucleotide repeats affect repair outcome. Biochemistry 52, 773–785. 10.1021/bi301369b. [PubMed: 23339280]
72. Goula A-V, Berquist BR, Wilson DM, Wheeler VC, Trottier Y, and Merienne K (2009). Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. PLoS Genet. 5, e1000749. 10.1371/journal.pgen.1000749. [PubMed: 19997493]

73. Goula A-V, Pearson CE, Della Maria J, Trottier Y, Tomkinson AE, Wilson DM, and Merienne K (2012). The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excisi on repair outcome at CAG/CTG repeats. Biochemistry 51, 3919–3932. 10.1021/bi300410d. [PubMed: 22497302]
74. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, McMurray CT, and Wilson SH (2009). Coordination between Polymerase β and FEN1 Can Modulate CAG Repeat Expansion *. J. Biol. Chem. 284, 28352–28366. 10.1074/jbc.M109.050286. [PubMed: 19674974]
75. Liu Y, Prasad R, and Wilson SH (2010). HMGB1: roles in base excision repair and related function. Biochim. Biophys. Acta 1799, 119–130. 10.1016/j.bbagrm.2009.11.008. [PubMed: 20123074]
76. Kim C, and Wold MS (1995). Recombinant human replication protein A binds to polynucleotides with low cooperativity. Biochemistry 34, 2058–2064. 10.1021/bi00006a028. [PubMed: 7849064]
77. Lao Y, Lee CG, and Wold MS (1999). Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. Biochemistry 38, 3974–3984. 10.1021/bi982371m. [PubMed: 10194309]
78. Hartwig A (2001). Role of magnesium in genomic stability. Mutat. Res. 475, 113–121. 10.1016/s0027-5107(01)00074-4. [PubMed: 11295157]
79. Tsutakawa SE, Lafrance-Vanasse J, and Tainer JA (2014). The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. DNA Repair 19, 95–107. 10.1016/j.dnarep.2014.03.022. [PubMed: 24754999]
80. Tam M, Erin Montgomery S, Kekis M, David Stollar B, Price GB, and Pearson CE (2003). Slipped (CTG)·(CAG) Repeats of the Myotonic Dystrophy Locus: Surface Probing with Anti-DNA Antibodies. J. Mol. Biol. 332, 585–600. 10.1016/S0022-2836(03)00880-5. [PubMed: 12963369]
81. Myler LR, Gallardo IF, Zhou Y, Gong F, Yang S-H, Wold MS, Miller KM, Paull TT, and Finkelstein IJ (2016). Single-molecule imaging reveals the mechanism of Exo1 regulation by single-stranded DNA binding proteins. Proc. Natl. Acad. Sci. U. S. A. 113, E1170–1179. 10.1073/pnas.1516674113. [PubMed: 26884156]
82. Nguyen B, Sokoloski J, Galletto R, Elson EL, Wold MS, and Lohman TM (2014). Diffusion of human replication protein A along single-stranded DNA. J. Mol. Biol. 426, 3246–3261. 10.1016/j.jmb.2014.07.014. [PubMed: 25058683]
83. Gibb B, Ye LF, Gergoudis SC, Kwon Y, Niu H, Sung P, and Greene EC (2014). Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. PloS One 9, e87922. 10.1371/journal.pone.0087922. [PubMed: 24498402]
84. Sears RM, May DG, and Roux KJ (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. In Enzyme-Mediated Ligation Methods Methods in Molecular Biology., Nuijens T and Schmidt M, eds. (Springer), pp. 299–313. 10.1007/978-1-4939-9546-2_15.
85. Roux KJ, Kim DI, Burke B, and May DG (2018). BioID: A Screen for Protein-Protein Interactions. Curr. Protoc. Protein Sci. 91, 19.23.1–19.23.15. 10.1002/cpps.51.
86. Varnaitė R, and MacNeill SA (2016). Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. PROTEOMICS 16, 2503–2518. 10.1002/pmic.201600123. [PubMed: 27329485]
87. Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, Motamedchaboki K, and Roux KJ (2016). An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell 27, 1188–1196. 10.1091/mbc.E15-12-0844. [PubMed: 26912792]
88. Hulsen T, de Vlieg J, and Alkema W (2008). BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9, 488. 10.1186/1471-2164-9-488. [PubMed: 18925949]
89. Wang Q-M, Yang Y-T, Wang Y-R, Gao B, Xi X, and Hou X-M (2019). Human replication protein A induces dynamic changes in single-stranded DNA and RNA structures. J. Biol. Chem. 294, 13915–13927. 10.1074/jbc.RA119.009737. [PubMed: 31350334]
90. Manfrini N, Trovesi C, Wery M, Martina M, Cesena D, Descrimes M, Morillon A, d'Adda di Fagagna F, and Longhese MP (2015). RNA-processing proteins regulate Mec1/ATR

activation by promoting generation of RPA-coated ssDNA. EMBO Rep. 16, 221–231. 10.15252/embr.201439458. [PubMed: 25527408]

91. Lee H-C, Aalto AP, Yang Q, Chang S-S, Huang G, Fisher D, Cha J, Poranen MM, Bamford DH, and Liu Y (2010). The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. PLoS Biol. 8, e1000496. 10.1371/journal.pbio.1000496. [PubMed: 20957187]
92. Nguyen HD, Yadav T, Giri S, Saez B, Graubert TA, and Zou L (2017). Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. Mol. Cell 65, 832–847.e4. 10.1016/j.molcel.2017.01.029. [PubMed: 28257700]
93. Feng S, and Manley JL (2021). Replication protein A associates with nucleolar R loops and regulates rRNA transcription and nucleolar morphology. Genes Dev. 35, 1579–1594. 10.1101/gad.348858.121. [PubMed: 34819354]
94. Dueva R, and Iliakis G (2020). Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. NAR Cancer 2, zcaa022. 10.1093/narcan/zcaa022. [PubMed: 34316690]
95. Maréchal A, and Zou L (2015). RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. Cell Res. 25, 9–23. 10.1038/cr.2014.147. [PubMed: 25403473]
96. Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 49, D325–D334. 10.1093/nar/gkaa1113. [PubMed: 33290552]
97. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29. 10.1038/75556. [PubMed: 10802651]
98. Nakatani R, Nakamori M, Fujimura H, Mochizuki H, and Takahashi MP (2015). Large expansion of CTG•CAG repeats is exacerbated by MutSβ in human cells. Sci. Rep. 5, 11020. 10.1038/srep11020. [PubMed: 26047474]
99. van den Broek WJAA, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJTA, and Wieringa B (2002). Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch–repair proteins. Hum. Mol. Genet. 11, 191–198. 10.1093/hmg/11.2.191. [PubMed: 11809728]
100. Tomé S, Simard JP, Slean MM, Holt I, Morris GE, Wojciechowicz K, te Riele H, and Pearson CE (2013). Tissue-specific mismatch repair protein expression: MSH3 is higher than MSH6 in multiple mouse tissues. DNA Repair 12, 46–52. 10.1016/j.dnarep.2012.10.006. [PubMed: 23228367]
101. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, and Monckton DG (2004). Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. Hum. Mol. Genet. 13, 1815–1825. 10.1093/hmg/ddh186. [PubMed: 15198993]
102. Lin Y, and Wilson JH (2009). Diverse Effects of Individual Mismatch Repair Components on Transcription-Induced CAG Repeat Instability in Human Cells. DNA Repair 8, 878–885. 10.1016/j.dnarep.2009.04.024. [PubMed: 19497791]
103. Jung J, and Bonini N (2007). CREB-binding protein modulates repeat instability in a Drosophila model for polyQ disease. Science 315, 1857–1859. 10.1126/science.1139517. [PubMed: 17332375]
104. Su XA, and Freudenreich CH (2017). Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. 114, E8392–E8401. 10.1073/pnas.1711283114. [PubMed: 28923949]
105. Lin Y, and Wilson JH (2007). Transcription-Induced CAG Repeat Contraction in Human Cells Is Mediated in Part by Transcription-Coupled Nucleotide Excision Repair. Mol. Cell. Biol. 27, 6209–6217. 10.1128/MCB.00739-07. [PubMed: 17591697]
106. Dutta A, Ruppert JM, Aster JC, and Winchester E (1993). Inhibition of DNA replication factor RPA by p53. Nature 365, 79–82. 10.1038/365079a0. [PubMed: 8361542]

107. Li R, and Botchan MR (1993). The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate in vitro BPV-1 DNA replication. Cell 73, 1207–1221. 10.1016/0092-8674(93)90649-b. [PubMed: 8390328]

108. He Z, Brinton BT, Greenblatt J, Hassell JA, and Ingles CJ (1993). The transactivator proteins VP16 and GAL4 bind replication factor A. Cell 73, 1223–1232. 10.1016/0092-8674(93)90650-f. [PubMed: 8513504]

109. Miller SD, Moses K, Jayaraman L, and Prives C (1997). Complex formation between p53 and replication protein A inhibits the sequence-specific DNA binding of p53 and is regulated by single-stranded DNA. Mol. Cell. Biol. 17, 2194–2201. 10.1128/MCB.17.4.2194. [PubMed: 9121469]

110. Bochkareva E, Kaustov L, Ayed A, Yi G-S, Lu Y, Pineda-Lucena A, Liao JCC, Okorokov AL, Milner J, Arrowsmith CH, et al. (2005). Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc. Natl. Acad. Sci. U. S. A. 102, 15412–15417. 10.1073/pnas.0504614102. [PubMed: 16234232]

111. Nakamori M, Pearson CE, and Thornton CA (2011). Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. Hum. Mol. Genet. 20, 580–588. 10.1093/hmg/ddq501. [PubMed: 21088112]

112. Taniguchi JB, Kondo K, Fujita K, Chen X, Homma H, Sudo T, Mao Y, Watase K, Tanaka T, Tagawa K, et al. (2016). RpA1 ameliorates symptoms of mutant ataxin-1 knock-in mice and enhances DNA damage repair. Hum. Mol. Genet. 25, 4432–4447. 10.1093/hmg/ddw272. [PubMed: 28173122]

113. Dion V, Lin Y, Hubert L Jr., Waterland RA, and Wilson JH (2008). Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. Hum. Mol. Genet. 17, 1306–1317. 10.1093/hmg/ddn019. [PubMed: 18252747]

114. Hubert L, Lin Y, Dion V, and Wilson JH (2011). Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. Hum. Mol. Genet. 20, 4822–4830. 10.1093/hmg/ddr421. [PubMed: 21926083]

115. Watase K, Weeber EJ, Xu B, Antalffy B, Yuva-Paylor L, Hashimoto K, Kano M, Atkinson R, Sun Y, Armstrong DL, et al. (2002). A long CAG repeat in the mouse Sca1 locus replicates SCA1 features and reveals the impact of protein solubility on selective neurodegeneration. Neuron 34, 905–919. 10.1016/s0896-6273(02)00733-x. [PubMed: 12086639]

116. Watase K, Venken KJT, Sun Y, Orr HT, and Zoghbi HY (2003). Regional differences of somatic CAG repeat instability do not account for selective neuronal vulnerability in a knock-in mouse model of SCA1. Hum. Mol. Genet. 12, 2789–2795. 10.1093/hmg/ddg300. [PubMed: 12952864]

117. Hashida H, Goto J, Kurisaki H, Mizusawa H, and Kanazawa I (1997). Brain regional differences in the expansion of a CAG repeat in the spinocerebellar ataxias: dentatorubral-pallidoluysian atrophy, Machado-Joseph disease, and spinocerebellar ataxia type 1. Ann. Neurol. 41, 505–511. 10.1002/ana.410410414. [PubMed: 9124808]

118. Outwin E, Carpenter G, Bi W, Withers MA, Lupski JR, and O'Driscoll M (2011). Increased RPA1 gene dosage affects genomic stability potentially contributing to 17p13.3 duplication syndrome. PLoS Genet. 7, e1002247. 10.1371/journal.pgen.1002247. [PubMed: 21901111]

119. Dimitrova DS, and Gilbert DM (2000). Stability and nuclear distribution of mammalian replication protein A heterotrimeric complex. Exp. Cell Res. 254, 321–327. 10.1006/excr.1999.4770. [PubMed: 10640430]

120. Loo YM, and Melendy T (2000). The majority of human replication protein A remains complexed throughout the cell cycle. Nucleic Acids Res. 28, 3354–3360. 10.1093/nar/28.17.3354. [PubMed: 10954605]

121. Lee J-M, Zhang J, Su AI, Walker JR, Wiltshire T, Kang K, Dragileva E, Gillis T, Lopez ET, Boily M-J, et al. (2010). A novel approach to investigate tissue-specific trinucleotide repeat instability. BMC Syst. Biol. 4, 29. 10.1186/1752-0509-4-29.

122. Lange W (1975). Cell number and cell density in the cerebellar cortex of man and some other mammals. Cell Tissue Res. 157, 115–124. 10.1007/BF00223234. [PubMed: 804353]

123. Castaldo I, De Rosa M, Romano A, Zuchegna C, Squitieri F, Mechelli R, Peluso S, Borrelli C, Del Mondo A, Salvatore E, et al. (2019). DNA damage signatures in peripheral blood cells

as biomarkers in prodromal huntington disease. Ann. Neurol. 85, 296–301. 10.1002/ana.25393. [PubMed: 30549309]

124. Askeland G, Dosoudilova Z, Rodinova M, Klempir J, Liskova I, Kuśnierczyk A, Bjørås M, Nesse G, Klungland A, Hansikova H, et al. (2018). Increased nuclear DNA damage precedes mitochondrial dysfunction in peripheral blood mononuclear cells from Huntington's disease patients. Sci. Rep. 8, 9817. 10.1038/s41598-018-27985-y. [PubMed: 29959348]
125. Illuzzi J, Yerkes S, Parekh-Olmedo H, and Kmiec EB (2009). DNA breakage and induction of DNA damage response proteins precede the appearance of visible mutant huntingtin aggregates. J. Neurosci. Res. 87, 733–747. 10.1002/jnr.21881. [PubMed: 18831068]
126. Giuliano P, de Cristofaro T, Affaitati A, Pizzulo GM, Feliciello A, Criscuolo C, De Michele G, Filla A, Avvedimento EV, and Varrone S (2003). DNA damage induced by polyglutamine-expanded proteins. Hum. Mol. Genet. 12, 2301–2309. 10.1093/hmg/ddg242. [PubMed: 12915485]
127. Lahiri M, Gustafson TL, Majors ER, and Freudenreich CH (2004). Expanded CAG Repeats Activate the DNA Damage Checkpoint Pathway. Mol. Cell 15, 287–293. 10.1016/j.molcel.2004.06.034. [PubMed: 15260979]
128. Peng S, Guo P, Lin X, An Y, Sze KH, Lau MHY, Chen ZS, Wang Q, Li W, Sun JK-L, et al. (2021). CAG RNAs induce DNA damage and apoptosis by silencing NUDT16 expression in polyglutamine degeneration. Proc. Natl. Acad. Sci. 118, e2022940118. 10.1073/pnas.2022940118. [PubMed: 33947817]
129. Sundararajan R, and Freudenreich CH (2011). Expanded CAG/CTG Repeat DNA Induces a Checkpoint Response That Impacts Cell Proliferation in Saccharomyces cerevisiae. PLOS Genet. 7, e1001339. 10.1371/journal.pgen.1001339. [PubMed: 21437275]
130. Sundararajan R, Gellon L, Zunder RM, and Freudenreich CH (2010). Double-Strand Break Repair Pathways Protect against CAG/CTG Repeat Expansions, Contractions and Repeat-Mediated Chromosomal Fragility in Saccharomyces cerevisiae. Genetics 184, 65–77. 10.1534/genetics.109.111039. [PubMed: 19901069]
131. Gasset-Rosa F, Chillon-Marinas C, Goginashvili A, Atwal RS, Artates JW, Tabet R, Wheeler VC, Bang AG, Cleveland DW, and Lagier-Tourenne C (2017). Polyglutamine-Expanded Huntingtin Exacerbates Age-Related Disruption of Nuclear Integrity and Nucleocytoplasmic Transport. Neuron 94, 48–57.e4. 10.1016/j.neuron.2017.03.027. [PubMed: 28384474]
132. Binz SK, Sheehan AM, and Wold MS (2004). Replication protein A phosphorylation and the cellular response to DNA damage. DNA Repair 3, 1015–1024. 10.1016/j.dnarep.2004.03.028. [PubMed: 15279788]
133. Zou Y, Liu Y, Wu X, and Shell SM (2006). Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. J. Cell. Physiol. 208, 267–273. 10.1002/jcp.20622. [PubMed: 16523492]
134. Handler HP, Duvick L, Mitchell JS, Cvetanovic M, Reighard M, Soles A, Mather KB, Rainwater O, Serres S, Nichols-Meade T, et al. (2023). Decreasing mutant ATXN1 nuclear localization improves a spectrum of SCA1-like phenotypes and brain region transcriptomic profiles. Neuron 111, 493–507.e6. 10.1016/j.neuron.2022.11.017. [PubMed: 36577403]
135. DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, and Aronin N (1997). Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science 277, 1990–1993. 10.1126/science.277.5334.1990. [PubMed: 9302293]
136. Li SH, and Li XJ (1998). Aggregation of N-terminal huntingtin is dependent on the length of its glutamine repeats. Hum. Mol. Genet. 7, 777–782. 10.1093/hmg/7.5.777. [PubMed: 9536080]
137. Becher MW, Kotzuk JA, Sharp AH, Davies SW, Bates GP, Price DL, and Ross CA (1998). Intranuclear neuronal inclusions in Huntington's disease and dentatorubral and pallidoluysian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. Neurobiol. Dis. 4, 387–397. 10.1006/nbdi.1998.0168. [PubMed: 9666478]
138. Li H, Li SH, Cheng AL, Mangiarini L, Bates GP, and Li XJ (1999). Ultrastructural localization and progressive formation of neuropil aggregates in Huntington's disease transgenic mice. Hum. Mol. Genet. 8, 1227–1236. 10.1093/hmg/8.7.1227. [PubMed: 10369868]

139. Li H, Li SH, Johnston H, Shelbourne PF, and Li XJ (2000). Amino-terminal fragments of mutant huntingtin show selective accumulation in striatal neurons and synaptic toxicity. Nat. Genet. 25, 385–389. 10.1038/78054. [PubMed: 10932179]

140. Carty N, Berson N, Tillack K, Thiede C, Scholz D, Kottig K, Sedaghat Y, Gabrysiak C, Yohrling G, Kammer H von der, et al. (2015). Characterization of HTT Inclusion Size, Location, and Timing in the zQ175 Mouse Model of Huntingtoń s Disease: An In Vivo High-Content Imaging Study. PLOS ONE 10, e0123527. 10.1371/journal.pone.0123527. [PubMed: 25859666]

141. Cummings CJ, Reinstein E, Sun Y, Antalffy B, Jiang Y, Ciechanover A, Orr HT, Beaudet AL, and Zoghbi HY (1999). Mutation of the E6-AP Ubiquitin Ligase Reduces Nuclear Inclusion Frequency While Accelerating Polyglutamine-Induced Pathology in SCA1 Mice. Neuron 24, 879–892. 10.1016/S0896-6273(00)81035-1. [PubMed: 10624951]

142. Rosche WA, Jaworski A, Kang S, Kramer SF, Larson JE, Geidroc DP, Wells RD, and Sinden RR (1996). Single-stranded DNA-binding protein enhances the stability of CTG triplet repeats in Escherichia coli. J. Bacteriol. 178, 5042–5044. 10.1128/jb.178.16.5042-5044.1996. [PubMed: 8759875]

143. Khristich AN, Armenia JF, Matera RM, Kolchinski AA, and Mirkin SM (2020). Large-scale contractions of Friedreich's ataxia GAA repeats in yeast occur during DNA replication due to their triplex-forming ability. Proc. Natl. Acad. Sci. U. S. A. 117, 1628–1637. 10.1073/pnas.1913416117. [PubMed: 31911468]

144. Andreoni F, Darmon E, Poon WCK, and Leach DRF (2010). Overexpression of the single-stranded DNA-binding protein (SSB) stabilises CAG*CTG triplet repeats in an orientation dependent manner. FEBS Lett. 584, 153–158. 10.1016/j.febslet.2009.11.042. [PubMed: 19925793]

145. Dragileva E, Hendricks A, Teed A, Gillis T, Lopez ET, Friedberg EC, Kucherlapati R, Edelmann W, Lunetta KL, MacDonald ME, et al. (2009). Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. Neurobiol. Dis. 33, 37–47. 10.1016/j.nbd.2008.09.014. [PubMed: 18930147]

146. Zhao X-N, Lokanga R, Allette K, Gazy I, Wu D, and Usdin K (2016). A MutSβ-Dependent Contribution of MutSα to Repeat Expansions in Fragile X Premutation Mice? PLoS Genet. 12, e1006190. 10.1371/journal.pgen.1006190. [PubMed: 27427765]

147. Mason AG, Tomé S, Simard JP, Libby RT, Bammler TK, Beyer RP, Morton AJ, Pearson CE, and La Spada AR (2014). Expression levels of DNA replication and repair genes predict regional somatic repeat instability in the brain but are not altered by polyglutamine disease protein expression or age. Hum. Mol. Genet. 23, 1606–1618. 10.1093/hmg/ddt551. [PubMed: 24191263]

148. Maiuri T, Mocle AJ, Hung CL, Xia J, van Roon-Mom WMC, and Truant R (2017). Huntingtin is a scaffolding protein in the ATM oxidative DNA damage response complex. Hum. Mol. Genet. 26, 395–406. 10.1093/hmg/ddw395. [PubMed: 28017939]

149. Gao R, Chakraborty A, Geater C, Pradhan S, Gordon KL, Snowden J, Yuan S, Dickey AS, Choudhary S, Ashizawa T, et al. (2019). Mutant huntingtin impairs PNKP and ATXN3, disrupting DNA repair and transcription. eLife 8, e42988. 10.7554/eLife.42988. [PubMed: 30994454]

150. Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu Y-Z, Gohler H, Wanker EE, Bates GP, Housman DE, and Thompson LM (2000). The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc. Natl. Acad. Sci. U. S. A. 97, 6763–6768. [PubMed: 10823891]

151. Kennedy MA, Greco TM, Song B, and Cristea IM (2022). HTT-OMNI: A Web-based Platform for Huntingtin Interaction Exploration and Multi-omics Data Integration. Mol. Cell. Proteomics MCP 21, 100275. 10.1016/j.mcpro.2022.100275. [PubMed: 35932982]

152. Hong EP, Chao MJ, Massey T, McAllister B, Lobanov S, Jones L, Holmans P, Kwak S, Orth M, Ciosi M, et al. (2021). Association Analysis of Chromosome X to Identify Genetic Modifiers of Huntington's Disease. J. Huntingt. Dis. 10, 367–375. 10.3233/JHD-210485.

153. Martins S, Pearson CE, Coutinho P, Provost S, Amorim A, Dubé M-P, Sequeiros J, and Rouleau GA (2014). Modifiers of (CAG)(n) instability in Machado-Joseph disease (MJD/SCA3) transmissions: an association study with DNA replication, repair and recombination genes. Hum. Genet. 133, 1311–1318. 10.1007/s00439-014-1467-8. [PubMed: 25026993]

154. O'Driscoll M, Dobyns WB, van Hagen JM, and Jeggo PA (2007). Cellular and clinical impact of haploinsufficiency for genes involved in ATR signaling. Am. J. Hum. Genet. 81, 77–86. 10.1086/518696. [PubMed: 17564965]

155. Blazejewski SM, Bennison SA, Smith TH, and Toyo-Oka K (2018). Neurodevelopmental Genetic Diseases Associated With Microdeletions and Microduplications of Chromosome 17p13.3. Front. Genet. 9, 80. 10.3389/fgene.2018.00080. [PubMed: 29628935]

156. Yoshihara T, Ishida M, Kinomura A, Katsura M, Tsuruga T, Tashiro S, Asahara T, and Miyagawa K (2004). XRCC3 deficiency results in a defect in recombination and increased endoreduplication in human cells. EMBO J. 23, 670–680. 10.1038/sj.emboj.7600087. [PubMed: 14749735]

157. Sharma R, Sahoo SS, Honda M, Granger SL, Goodings C, Sanchez L, Künstner A, Busch H, Beier F, Pruett-Miller SM, et al. (2022). Gain-of-function mutations in RPA1 cause a syndrome with short telomeres and somatic genetic rescue. Blood 139, 1039–1051. 10.1182/blood.2021011980. [PubMed: 34767620]

158. Maurin M-L, Arfeuille C, Sonigo P, Rondeau S, Vekemans M, Turleau C, Ville Y, and Malan V (2017). Large Duplications Can Be Benign Copy Number Variants: A Case of a 3.6-Mb Xq21.33 Duplication. Cytogenet. Genome Res. 151, 115–118. 10.1159/000460278. [PubMed: 28273668]

159. Pillai JA, Hansen LA, Masliah E, Goldstein JL, Edland SD, and Corey-Bloom J (2012). Clinical severity of Huntington's disease does not always correlate with neuropathologic stage. Mov. Disord. Off. J. Mov. Disord. Soc. 27, 1099–1103. 10.1002/mds.25026.

160. Caramins M, Halliday G, McCusker E, and Trent RJ (2003). Genetically confirmed clinical Huntington's disease with no observable cell loss. J. Neurol. Neurosurg. Psychiatry 74, 968–970. 10.1136/jnnp.74.7.968. [PubMed: 12810795]

161. Mizuno H, Shibayama H, Tanaka F, Doyu M, Sobue G, Iwata H, Kobayashi H, Yamada K, Iwai K, Takeuchi T, et al. (2000). An autopsy case with clinically and molecular genetically diagnosed Huntington's disease with only minimal non-specific neuropathological findings. Clin. Neuropathol. 19, 94–103. [PubMed: 10749290]

162. Giardina F, Lanza G, Calì F, and Ferri R (2020). Late-onset oro-facial dyskinesia in Spinocerebellar Ataxia type 2: a case report. BMC Neurol. 20, 156. 10.1186/s12883-020-01739-8. [PubMed: 32340607]

163. Yang L, Dong Y, Ma Y, Ni W, and Wu Z-Y (2021). Genetic profile and clinical characteristics of Chinese patients with spinocerebellar ataxia type 2: A multicenter experience over 10 years. Eur. J. Neurol. 28, 955–964. 10.1111/ene.14601. [PubMed: 33070405]

164. Horimoto Y, Hayashi E, Ito Y, Iida A, Goto Y, Kato S, Okita K, Kako T, Sato C, Tajima T, et al. (2020). Dopaminergic function in spinocerebellar ataxia type 6 patients with and without parkinsonism. J. Neurol. 267, 2692–2696. 10.1007/s00415-020-09908-y. [PubMed: 32440919]

165. Mätlik K, Baffuto M, Kus L, Deshmukh AL, Davis DA, Paul MR, Carroll TS, Caron M-C, Masson J-Y, Pearson CE, et al. (2023). Cell Type Specific CAG Repeat Expansions and Toxicity of Mutant Huntingtin in Human Striatum and Cerebellum. 2023.04.24.538082. 10.1101/2023.04.24.538082.

166. Pressl C, Mätlik K, Kus L, Darnell P, Luo J-D, Paul MR, Weiss AR, Liguore W, Carroll TS, Davis DA, et al. (2023). Selective Vulnerability of Layer 5a Corticostriatal Neurons in Huntington's Disease. 2023.04.24.538096. 10.1101/2023.04.24.538096.

167. Olmos V, Gogia N, Luttik K, Haidery F, and Lim J (2022). The extra-cerebellar effects of spinocerebellar ataxia type 1 (SCA1): looking beyond the cerebellum. Cell. Mol. Life Sci. CMLS 79, 404. 10.1007/s00018-022-04419-7. [PubMed: 35802260]

168. Luttik K, Olmos V, Owens A, Khan A, Yun J, Driessen T, and Lim J (2022). Identifying Disease Signatures in the Spinocerebellar Ataxia Type 1 Mouse Cortex. Cells 11, 2632. 10.3390/cells11172632. [PubMed: 36078042]

169. Chaudhuri J, Khuong C, and Alt FW (2004). Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature 430, 992–998. 10.1038/nature02821. [PubMed: 15273694]

170. Slean MM, Panigrahi GB, Ranum LP, and Pearson CE (2008). Mutagenic roles of DNA “repair” proteins in antibody diversity and disease-associated trinucleotide repeat instability. DNA Repair 7, 1135–1154. 10.1016/j.dnarep.2008.03.014. [PubMed: 18485833]

171. Shi B, Xue J, Yin H, Guo R, Luo M, Ye L, Shi Q, Huang X, Liu M, Sha J, et al. (2019). Dual functions for the ssDNA-binding protein RPA in meiotic recombination. PLoS Genet. 15, e1007952. 10.1371/journal.pgen.1007952. [PubMed: 30716097]

172. Hinch AG, Becker PW, Li T, Moralli D, Zhang G, Bycroft C, Green C, Keeney S, Shi Q, Davies B, et al. (2020). The Configuration of RPA, RAD51, and DMC1 Binding in Meiosis Reveals the Nature of Critical Recombination Intermediates. Mol. Cell 79, 689–701.e10. 10.1016/j.molcel.2020.06.015. [PubMed: 32610038]

173. Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543–D552. 10.1093/nar/gkab1038. [PubMed: 34723319]

174. Waldvogel HJ, Curtis MA, Baer K, Rees MI, and Faull RLM (2006). Immunohistochemical staining of post-mortem adult human brain sections. Nat. Protoc. 1, 2719–2732. 10.1038/nprot.2006.354. [PubMed: 17406528]

175. Hung CL-K, Maiuri T, Bowie LE, Gotesman R, Son S, Falcone M, Giordano JV, Gillis T, Mattis V, Lau T, et al. (2018). A patient-derived cellular model for Huntington’s disease reveals phenotypes at clinically relevant CAG lengths. Mol. Biol. Cell 29, 2809–2820. 10.1091/mbc.E18-09-0590. [PubMed: 30256717]

176. Henricksen LA, Umbricht CB, and Wold MS (1994). Recombinant replication protein A: expression, complex formation, and functional characterization. J. Biol. Chem. 269, 11121–11132. [PubMed: 8157639]

177. Binz SK, Dickson AM, Haring SJ, and Wold MS (2006). Functional assays for replication protein A (RPA). Methods Enzymol. 409, 11–38. 10.1016/S0076-6879(05)09002-6. [PubMed: 16793393]

178. Maity R, Pauty J, Krietsch J, Buisson R, Genois M-M, and Masson J-Y (2013). GST-His purification: a two-step affinity purification protocol yielding full-length purified proteins. J. Vis. Exp. JoVE, e50320. 10.3791/50320. [PubMed: 24193370]

179. Wieczorek S, Combes F, Lazar C, Giai Gianetto Q, Gatto L, Dorffer A, Hesse A-M, Couté Y, Ferro M, Bruley C, et al. (2017). DAPAR & ProStaR: software to perform statistical analyses in quantitative discovery proteomics. Bioinforma. Oxf. Engl. 33, 135–136. 10.1093/bioinformatics/btw580.

180. Choi H, Larsen B, Lin Z-Y, Breitkreutz A, Mellacheruvu D, Fermin D, Qin ZS, Tyers M, Gingras A-C, and Nesvizhskii AI (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods 8, 70–73. 10.1038/nmeth.1541. [PubMed: 21131968]

181. Babeu J-P, Wilson SD, Lambert É, Lévesque D, Boisvert F-M, and Boudreau F (2019). Quantitative Proteomics Identifies DNA Repair as a Novel Biological Function for Hepatocyte Nuclear Factor 4α in Colorectal Cancer Cells. Cancers 11, E626. 10.3390/cancers11050626.

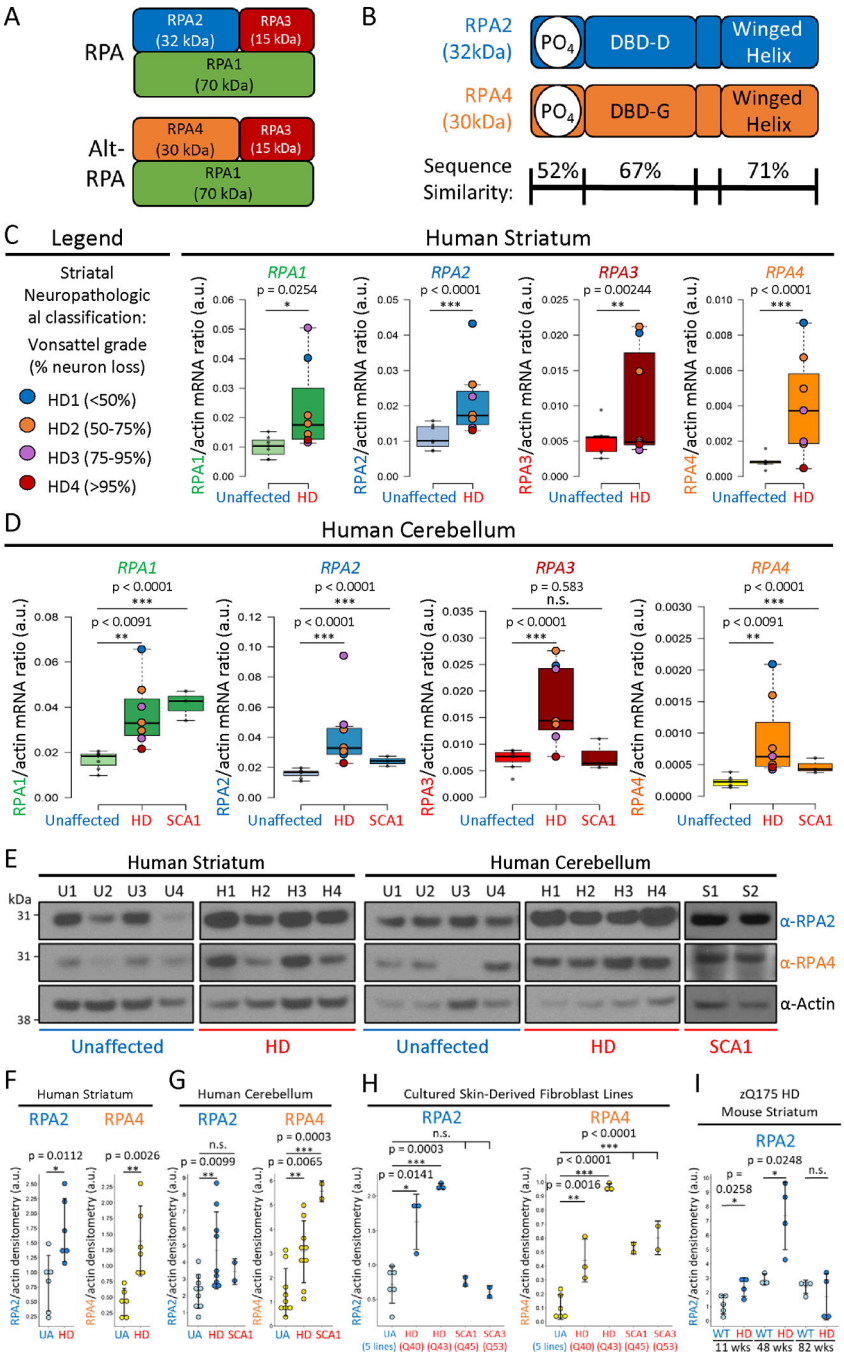


Figure 1: RPA1, 2, 3 &4 are upregulated in patient brain tissues, cell lines, and mouse brain tissues.

A) canonical RPA (RPA1, 2, 3) versus Alt-RPA (RPA1, 4, 3). B) RPA2 and RPA4 have homologous DNA-binding domains (DBD), winged helix domains, and a less similar N-terminal region. ddPCR data of all RPA subunits from C) striatum of HD patients (n=7 individuals/group/tissue, 3 replicates/person) and D) cerebellum of HD and SCA1 patients. (n=7 HD & unaffected individuals/group/tissue, and 3 SCA1 patients/group/tissue, 3 replicates/person). Colored dots indicate striatal neuropathological grade (HD 1–4). E) Representative western blots for RPA2 and RPA4 expression levels in HD and SCA1 patient

striatum and cerebellum relative to unaffected control tissues. Actin loading control. F-I) densitometric quantification of RPA2 and RPA4 levels versus actin in striatum, cerebellum, cultured fibroblasts, and zQ175 HD mouse striatum. Unpaired student t-test comparing means. Data are represented as median ± Tukey whisker extent (box plots) or mean ± SD (dot plots).

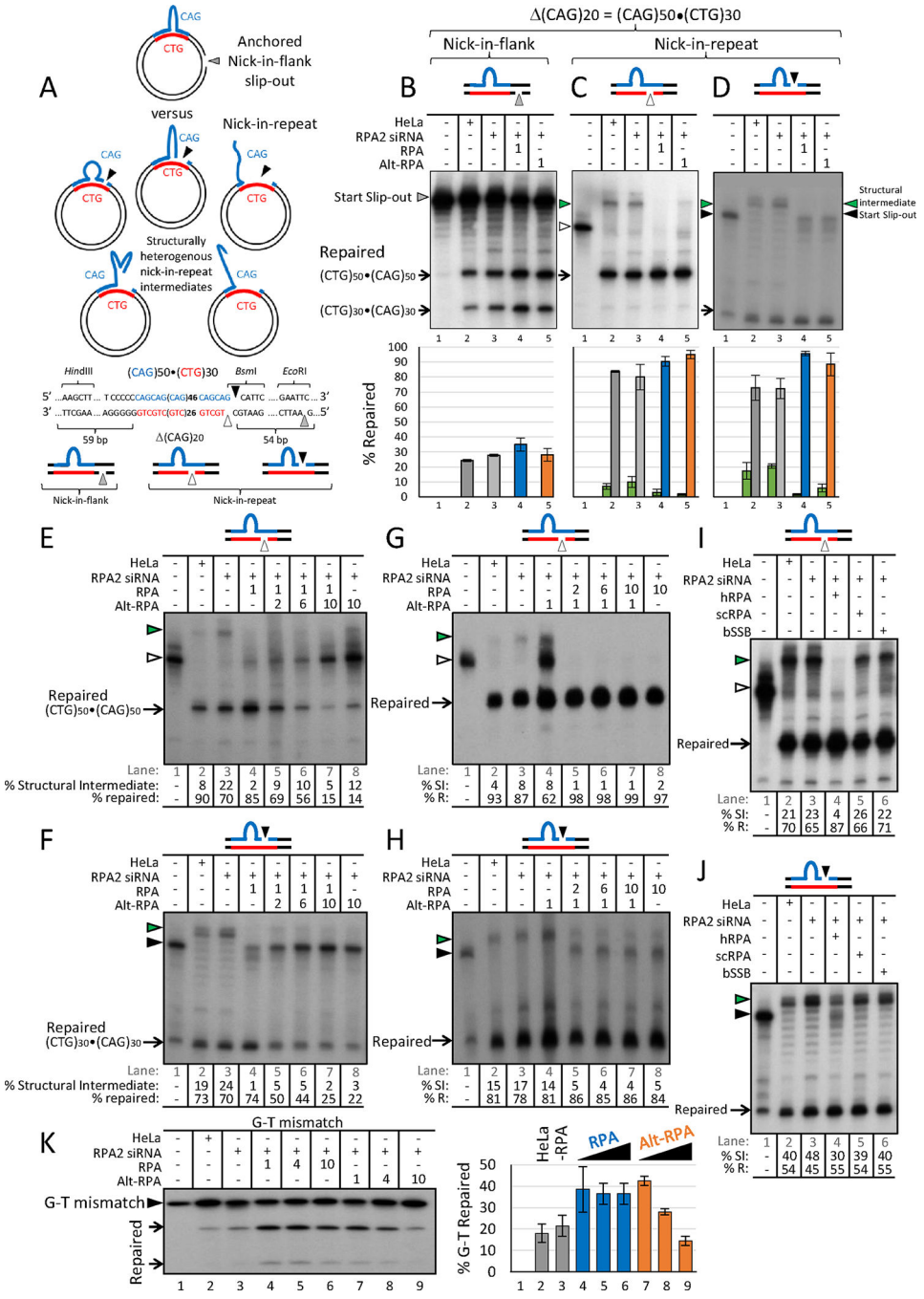


Figure 2: RPA and Alt-RPA competitively modulate slip-out repair.

A) Nick-in-flank substrates are structurally stable; whereas nick-in-repeat substrates permit multiple heterogenous structures. Nicks created by cutting and annealing strands of (CAG)30 and (CAG)50 DNAs (nicks-in-flank, *Eco*RI, and nicks-in-repeat, *Bsm*I, grey, white, and black triangles). B-D) Repair reactions with 25 μg of HeLa or RPA-/Alt-RPA-depleted (RPA2 siRNA) HeLa extracts, dNTPs, and purified (6 μg) RPA or Alt-RPA. Reaction products were purified, repeat-containing fragments released (*Eco*RI/*Hin*dIII), electrophoretically resolved on native 4% PAGE, and visualized by Southern. Structural

intermediates (SI; green triangles) and fully-duplexed repair products (black arrows). Densitometric quantification of 3–6 replicates (normalized for background). Graphed SI levels (green bars) and correct repair products (grey, blue, and orange bars). E-H) Repair reactions with indicated slip-out or G-T mismatch, cell extract, and purified protein(s), as in B-D. Where indicated reactions included 0.6 μg RPA (lane 4) or Alt-RPA (lane 4). Lanes 5–7 shows processing by RPA-deficient extract supplemented by 0.6 μg RPA or Alt-RPA at ratios of 1:0, 1:2, 1:6, or 1:10 Alt-RPA or RPA. Lane 8 shows substrate processed by RPA-deficient extract supplemented by only 6 μg RPA or Alt-RPA. I-J). Repair reactions were performed as above, included 0.6 μg of human RPA, bacterial bSSB or yeast scRPA. Quantitation of E-J as for B-D. Data are the mean ± SD.

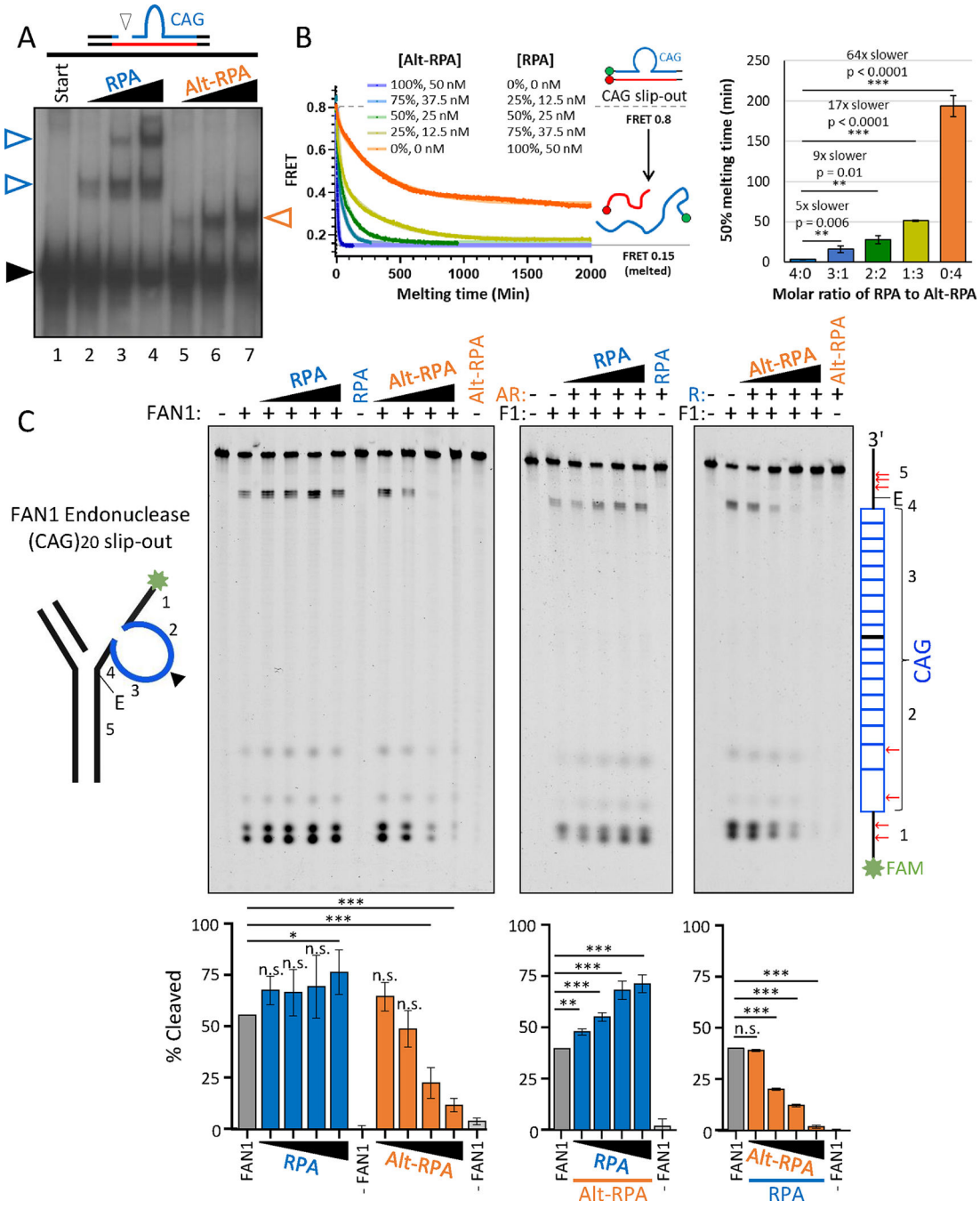


Figure 3: Alt-RPA has altered binding, poorly melts slipped-DNAs, and inhibits FAN1 cleavage.

A) γ-^{32}P-radiolabelled DNA substrates were incubated with 600 ng (lanes 2, 5), 1200 ng (lanes 3, 6), or 2400 ng (lanes 4, 7) of RPA or Alt-RPA, resolved on 4% native acrylamide. Free-DNA (black triangles); protein-DNA complexes (blue arrowheads (RPA) or orange arrowhead (Alt-RPA)). B) Time-course of melting of fluorescently labelled slipped-CAG oligonucleotides by RPA and/or Alt-RPA. Time needed to reach 0.475 FRET (i.e. half the DNA being melted) quantified rates of melting by each complex based on molar ratio of RPA:Alt-RPA (bar graph; n=3 replicates). Statistics: unpaired t-test comparing means. C)

FAN1 *endo*-nuclease activity is enhanced by RPA and inhibited by Alt-RPA. Purified FAN1, RPA € and/or Alt-RPA (AR) were incubated with FAM-labelled oligonucleotides (mimics slipped-CAG; schematic at left) (- = no FAN1 added and 200 nM of RPA or Alt-RPA, + = 50 nM FAN1; triangles: 25 nM, 50 nM, 100 nM, or 200 nM RPA or Alt-RPA; straight line = 50 nM RPA or Alt-RPA) and resolved on a 4% denaturing gel. Schematic (right) indicates migration positions for labeled CAG strand and FAN1 *endo*-nucleolytic cleavage positions (red arrows). Nuclease activity was quantified densitometrically (cleavage products intensity/full-length substrate intensity, n=3 replicates. *=$p<0.05$, ***=$p<0.001$). Data are the mean ± SD.

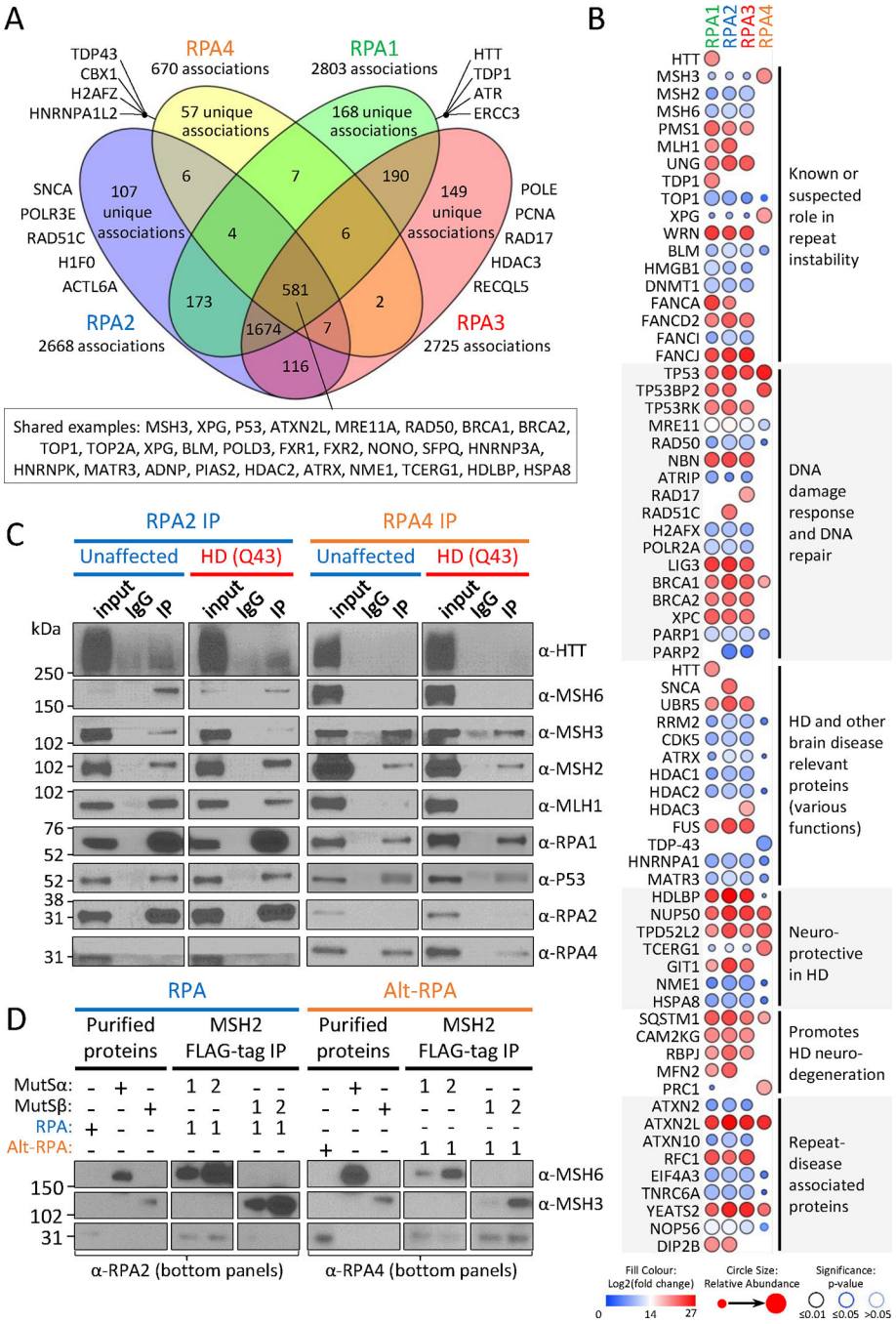


Figure 4: RPA and Alt-RPA BioID interactomes.

A) Shared and unique BioID hits for RPA and Alt-RPA subunits from HEK293T cells. B) Dot plot outlining subunit associated proteins (functionally categorized). Fill-color shows log2 fold enrichment versus BirA-only control (non-specific interactions): protein enrichment versus whole-dataset minimum (dark blue) and maximum enrichment (dark red) value. Dot size indicates relative abundance of interaction versus other subunits. Dot outline color shows significance (black=p<0.01, blue=p<0.05, and light blue, p>0.05). C) Co-IP of RPA (RPA2) or Alt-RPA (RPA4) from unaffected and HD patient-derived fibroblasts, or D)

Co-Ips of purified RPA or Alt-RPA complexes with purified MutSα or MutSβ complexes using FLAG-tagged MSH2. In C and D co-Ips visualised by Western, n=3–6 replicates; representative images.

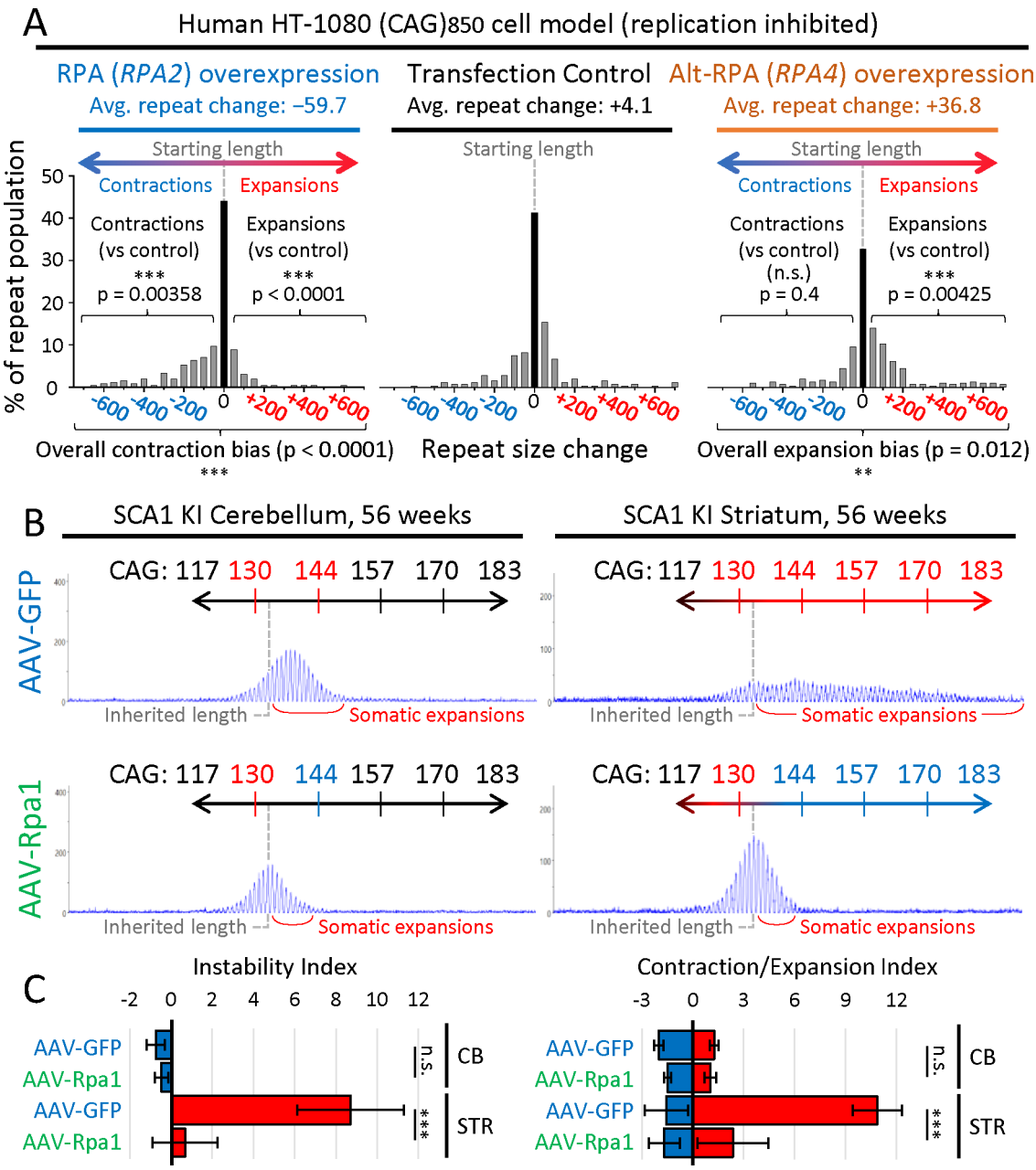


Figure 5: Upregulation of RPA inhibits while Alt-RPA enhances somatic repeat expansions.

A) Small-pool PCR (spPCR) quantification of RPA (left panel) or Alt-RPA (right panel) overexpression versus transfection control (middle panel) in serum-starved HT1080 (CAG)850 cells grown10-days. (Light gray line=starting repeat length, gray bars=expansions (right) and contractions (left). Statistical analysis, with indicated comparisons, were by χ-square test[27,57]. B) Representative fragment length analysis scans of GFP- and Rpa1-overexpressing 56-week-old SCA1 mouse cerebellum and striatum. (Gray bar=inherited repeat length, red brackets=ongoing expansions). C) Average instability and expansion/contraction indices for all GFP- and Rpa1-overexpressing SCA1 mouse cerebellum and striatum ***=p<0.001. Data are represented as mean ± SD.

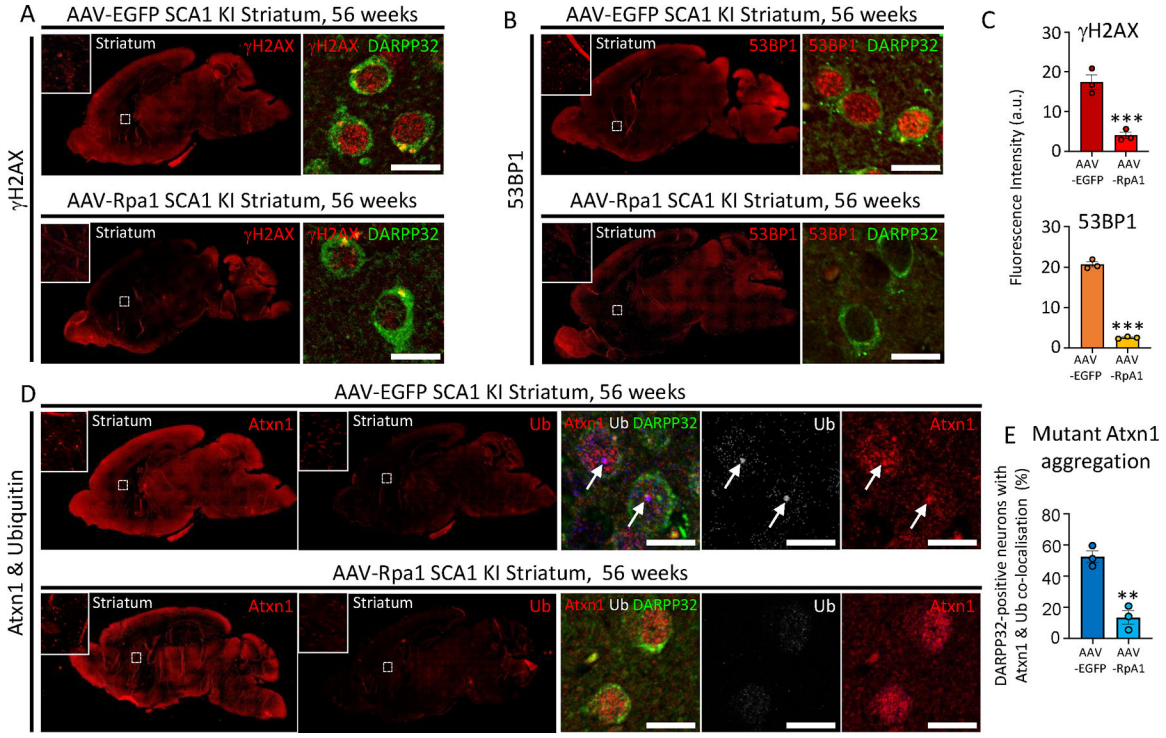


Figure 6: Canonical RPA upregulation in SCA1 mouse striatum reduces neuronal DNA damage and mutant Ataxin-1 aggregation.

A-B) Representative IF images of γ-H2AX and 53BP1 in DARPP32-posititive striatal medium spiny neurons (MSNs) and C) quantification (n=3 mice/group, 3 replicates with ≥30 neurons/replicate, ***=p<0.001). D) Representative images of Ataxin-1 and ubiquitin co-staining within MSNs and E) quantification (n=3 mice/group, 3 replicates with ≥50 neurons/replicate. **=p<0.01). Data are mean ± SD.

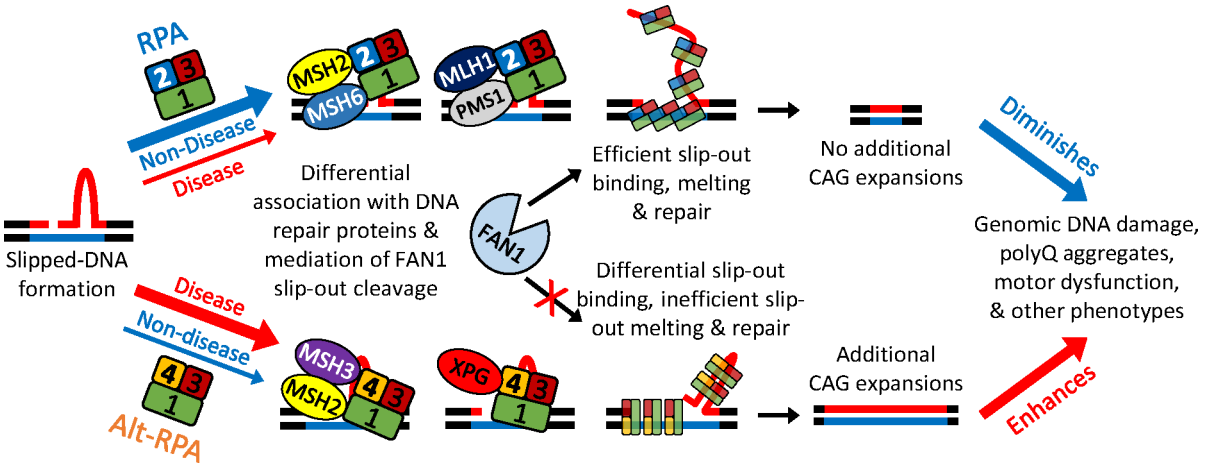


Figure 7: Working model of Alt-RPA↔RPA antagonistic interactions in somatic expansions.

(Top) Canonical RPA enhances correct repair by rapidly melting slip-outs, enhancing FAN1-mediated slip-out excision reducing CAG expansions and diminishing downstream events. (Bottom) Alt-RPA inhibits repair by differential binding, inefficiently melting slip-outs, inhibiting FAN1-mediated excision, leading to expansions and exacerbating downstream events. For non-disease state, the top pathway may predominate whereas for diseased state (dysregulated RPA/Alt-RPA levels) the bottom path may predominate. Differential associations of RPA/Alt-RPA with DNA repair proteins known to modulate repeat expansions also contribute to instability. RPA preferentially interacts with MutSα (MSH2-MSH6) and MutLβ (MLH1-PMS1). Alt-RPA preferentially interacts with MutSβ (MSH2-MSH3) and XPG. Downstream events shown in this study include brain DNA damage and neuronal polyglutamine-aggregates (Figure 6A–B and S7), motor phenotypes, neuron morphology, dysregulated transcriptome, spliceosome, and cell cycle[112].

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Western blotting primary: Monoclonal Mouse Anti-RPA2 (Clone 9H8)	Abcam	Catalogue #ab2175
Western blotting primary and co-IP: Sheep Serum Anti-RPA4	This paper	N/A
Western blotting primary: monoclonal mouse Anti-Actin	BD Transduction Laboratory	Catalogue #612657
Western blotting secondary: Peroxidase-AffiniPure Sheep Anti-Mouse IgG H+L	Cedarlane Labs	Catalogue #515035062
Western blotting secondary: Sheep IgG (H+L) Highly Cross-Adsorbed Donkey anti-Ovine HRP	ThermoFisher Scientific	Catalogue #A16047
Co-IP: Polyclonal rabbit Anti-RPA2	Thermo Fisher Scientific	Catalogue #A300–244A
Co-IP isotype control: Normal Rabbit control IgG	New England Biolabs	Catalogue #2729S
Co-IP isotype control: Normal Sheep control IgG	Thermo Fisher Scientific	Catalogue #31243
IF primary: Monoclonal Mouse Anti-phospho-H2AX (γ-H2AX) (Clone JBW301)	Millipore Sigma	Catalogue #05–636
IF primary: Polyclonal Rabbit Anti-53BP1	Novus bio	Catalogue #NB100–304SS
IF primary: Monoclonal Mouse Anti-ATXN1 (Clone N76/8)	EMD Millipore	Catalogue #MABN37
IF primary: Monoclonal Mouse Anti-ubiquitin (Clone P4D1)	Cell Signaling Technology	Catalogue #3936S
IF primary: Monoclonal Rabbit Anti-DARPP32 (Clone 19A3)	Cell Signaling Technology	Catalogue # 2306
IF primary: Monoclonal Rabbit Anti-calbindin (clone EG-20)	EMD Millipore	Catalogue #05–636
IF primary: Monoclonal Mouse Anti-calbindin (clone CB-955)	Millipore Sigma	Catalogue #C9848
IF secondary: Goat Anti-Mouse IgG (H+L) Superclonal Recombinant Secondary Antibody Alexa Fluor 555	ThermoFisher Scientific	Catalogue #A28180
IF secondary: Goat Anti-Rabbit IgG H&L Alexa Fluor 488	Abcam	Catalogue #ab150077
IF secondary: Goat Anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody Alexa Fluor 568	ThermoFisher Scientific	Catalogue #A-11011
IF secondary: Goat Anti-Mouse IgG H&L Alexa Fluor 488 preabsorbed	Abcam	Catalogue #ab150117
Bacterial and virus strains		
AAV1-CMV-Rpa1-EmGFP	Taniguchi *et al.*[112]	N/A
AAV1-CMV-EmGFP	Taniguchi *et al.*[112]	N/A
Biological samples		
Human Postmortem Frozen Brain Samples (HD and unaffected cohort #1) – see Table S1 for more details	Neurological Foundation Human Brain Bank	https://www.brainbank.ac.nz/
Human Postmortem Frozen Brain Samples (HD and unaffected cohort #2) – see Table S1 for more details	La Spada lab brain bank	N/A
Human Postmortem Frozen Brain Samples (SCA1) – see Table S1 for more details	National Ataxia Foundation Biobank	https://www.ataxia.org/
Chemicals, peptides, and recombinant proteins		
Recombinant Human RPA and Alt-RPA purified from *E. Coli*	Purified as previously described in *Binz et al.*[243]	N/A
Recombinant Human FAN1 purified from Sf9 insect cells	Purified as previously described in Deshmukh *et al.*[57]	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Critical commercial assays		
Direct-zol RNA purification kit	Zymo research	Catalog # R2071
SuperScript IV First-Strand Synthesis System kit	ThermoFisher Scientific	Catalogue #18091050
Streptavidin Sepharose High Performance beads	Sigma-Aldrich	Catalogue # GE17–5113-01
Protein G magnetic Dynabeads Immunoprecipitation kit	Thermo Fisher Scientific	Catalogue #10007D
Pierce Anti-FLAG magnetic beads	Millipore Sigma	Catalogue #M8823
Deposited data		
Mass spectrometry proteomics data collected for BioID deposited to the ProteomeXchange Consortium via the PRIDE partner repository[173]	This paper	Dataset identifier: PXD044158
Experimental models: Cell lines		
hTERT immortalized Q43 and Q40 HD patient derived fibroblasts and control cell line #1 – see Table S1 for more details	Laboratory of Ray Truant (previously described in Hung *et al.*[240])	N/A
Primary Q43 HD patient derived fibroblasts – see Table S1 for more details	Coriell Biorepository	GM02191
Primary Q45 SCA1 patient derived fibroblasts – see Table S1 for more details	Coriell Biorepository	GM06927
Primary Q53 SCA3 patient derived fibroblasts – see Table S1 for more details	Coriell Biorepository	GM06153
Primary control cell line #2 and #3 – see Table S1 for more details	Laboratory of Guy Rouleau	N/A
Primary control cell line #4 and #5 – see Table S1 for more details	Laboratory of Elise Heon	N/A
Stable integration of BioID constructs in HEK293-Flp-In-T-Rex cells	ThermoFisher Scientific	Catalogue #R78007
Experimental models: Organisms/strains		
Mouse: zQ175 heterozygous knock-in: B6J.129S1-Htttm1Mfc/190ChdiJ. The background mice C57BL/6 were used for breeding.	The Jackson Laboratory	RRID:IMSR_JAX:027410
Mouse: Atxn1–154Q, heterozygous knock-in. The background mice C57BL/6 were used for breeding.	Original mouse developed by laboratory of Huda Zoghbi[115], mice used for this paper were previously generated and described by Taniguchi *et al.*[112]	N/A
Oligonucleotides		
siRPA2 (commercial)	Santa Cruz Biotechnology	Catalogue #sc-38230
HT1080-(CAG)850 repeat sizing: See Table S2	This paper	N/A
ATXN1 repeat sizing: See Table S2	This paper	N/A
FAN1 CAG slip-out nuclease substrate: See Table S2	This paper	N/A
FAN1 unstructured non-repeat sequence nuclease substrate: See Table S2	This paper	N/A
EMSA ssDNA non-repeat unstructured substrate: See Table S2	This paper	N/A
EMSA and FRET duplex non-repeat unstructured substrate: See Table S2	This paper	N/A
EMSA and FRET 8nt non-repeat bubble substrate: See Table S2	This paper	N/A
EMSA 20nt non-repeat bubble substrate: See Table S2	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
EMSA 20nt CAG bubble substrate: See Table S2	This paper	N/A
FRET 20nt non-repeat slip-out substrate: See Table S2	This paper	N/A
FRET 20nt CAG repeat slip-out substrate: See Table S2	This paper	N/A
Recombinant DNA		
BioID plasmid constructs: RPA1, RPA2, RPA3, or RPA4 cDNA cloned into pgLAP1–3MYC-BioID2 vector	This paper	N/A
Slipped-CAG DNA repair and EMSA substrates	Previously described[59,60]	N/A
Human RPA2 and RPA4 overexpression constructs (for human cell line expression)	Haring *et al.*[40]	N/A
Software and algorithms		
QuantaSoft Software	Bio-Rad	https://www.bio-rad.com/en-ca/life-science/digital-pcr/qx200-droplet-digital-pcr-system/quantasoft-software-regulatory-edition
ggpmisc R package	The R foundation	https://cran.r-project.org/web/packages/ggpmisc/index.html
MaxQuant version 1.6.17.0 software392	Laboratory of Jürgen Cox/Max Planck Institute of Biochemistry	https://www.maxquant.org/
Prostar software	Prostar-proteomics	https://www.prostar-proteomics.org/
Peak Scanner 2 software	ThermoFisher Scientific	https://www.thermofisher.com/ca/en/home/life-science/sequencing/fragment-analysis/fragment-analysis-fundamentals/fragment-analysis-software-data-analysis.html
Graphpad Prism	Dotmatics	https://www.graphpad.com/
Prohits-viz	Laboratory of Anne-Claude Gingras of the University of Toronto	https://prohits-viz.org/
ImageJ software	ImageJ	https://imagej.net/ij/index.html