Molecular Cell
Short Article

Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability

Julie Sollier,[1] Caroline Townsend Stork,[1] María L. García-Rubio,[2] Renee D. Paulsen,[1] Andrés Aguilera,[2] and Karlene A. Cimprich[1,*]
[1]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA
[2]Centro Andaluz de Biología Molecular y Medicina Regenerativa CABIMER, Universidad de Sevilla, Avenida Américo Vespucio, 41092 Seville, Spain
*Correspondence: cimprich@stanford.edu
http://dx.doi.org/10.1016/j.molcel.2014.10.020

SUMMARY

R-loops, consisting of an RNA-DNA hybrid and displaced single-stranded DNA, are physiological structures that regulate various cellular processes occurring on chromatin. Intriguingly, changes in R-loop dynamics have also been associated with DNA damage accumulation and genome instability; however, the mechanisms underlying R-loop-induced DNA damage remain unknown. Here we demonstrate in human cells that R-loops induced by the absence of diverse RNA processing factors, including the RNA/DNA helicases Aquarius (AQR) and Senataxin (SETX), or by the inhibition of topoisomerase I, are actively processed into DNA double-strand breaks (DSBs) by the nucleotide excision repair endonucleases XPF and XPG. Surprisingly, DSB formation requires the transcription-coupled nucleotide excision repair (TC-NER) factor Cockayne syndrome group B (CSB), but not the global genome repair protein XPC. These findings reveal an unexpected and potentially deleterious role for TC-NER factors in driving R-loop-induced DNA damage and genome instability.

INTRODUCTION

R-loops, structures that contain an RNA-DNA hybrid and displaced single-stranded DNA, can form during transcription when an RNA molecule emerging from the transcription machinery hybridizes with the DNA template. These structures arise naturally in organisms from bacteria to humans, and they have a multitude of roles in the cell (Aguilera and García-Muse, 2012; Skourti-Stathaki and Proudfoot, 2014; Hamperl and Cimprich, 2014). In human cells, R-loops form over switch regions at the immunoglobulin locus to facilitate class switching, a physiological event in which DSBs are initiated through the processing of R-loops (Yu et al., 2003). In addition, R-loops form preferentially at the promoters of genes with a high GC skew to protect these regions from DNA methylation (Ginno et al., 2012). They also form at the termination regions of genes where they promote efficient transcriptional termination (Skourti-Stathaki et al., 2011).

R-loops can form in an unscheduled manner due to defects in RNA processing (Huertas and Aguilera, 2003; Li and Manley, 2005; Paulsen et al., 2009; Stirling et al., 2012; Wahba et al., 2011), and in these situations they are commonly associated with DNA damage. Indeed, R-loops were initially proposed to be the source of the hyperrecombination phenotype in yeast THO/TREX complex mutants, where they form as a result of defects in transcriptional elongation and RNA export (Huertas and Aguilera, 2003). Unscheduled R-loops are also thought to initiate the genomic or epigenomic changes associated with several neurodegenerative diseases, including amyotrophic lateral sclerosis, Fragile X syndrome, and Friedreich's ataxia (Chen et al., 2004; Colak et al., 2014; Groh et al., 2014; Loomis et al., 2014), and they can cause genome instability at trinucleotide repeat sequences and common fragile sites, suggesting that they may contribute to cancer (Haeusler et al., 2014; Helmrich et al., 2011).

Cells utilize diverse mechanisms to regulate the formation of R-loops. These structures can be resolved by RNase H, which specifically degrades the RNA moiety in RNA-DNA hybrids (Wahba et al., 2011), or by helicases such as Senataxin, which unwind RNA-DNA hybrids (Mischo et al., 2011; Skourti-Stathaki et al., 2011). R-loop formation is also suppressed by topoisomerase I, which resolves the negative torsional stress behind RNA polymerase II to prevent annealing of the nascent RNA with the DNA template (Tuduri et al., 2009). Other RNA processing factors also preclude R-loop formation, presumably by binding to RNA as it emerges from RNA polymerase (Li et al., 2007). However, when these mechanisms fail, R-loops may persist or accumulate, ultimately leading to DNA breaks and genome instability (Huertas and Aguilera, 2003; Li and Manley, 2005; Paulsen et al., 2009; Wahba et al., 2011; Tuduri et al., 2009; Stirling et al., 2012).

How DNA damage arises from an R-loop is an unresolved question. Several studies in bacteria, yeast, and human cells suggest that R-loop-induced DNA damage is associated with defects in replication fork progression (Alzu et al., 2012; Gan et al., 2011; Wellinger et al., 2006; Yüce and West, 2013; Tuduri et al., 2009). Whether it is the R-loop itself or the stalled RNA polymerase resulting from R-loop formation that impairs DNA replication and ultimately causes replication fork collapse and DSB formation is not clear. It has also been proposed that DNA damage may arise from the single-stranded DNA in the R-loop, because this DNA is more susceptible to DNA damaging agents (Lindahl, 1993) and could be targeted by enzymes like activation-induced cytidine deaminase (AID) that act at the immunoglobin

locus (Muramatsu et al., 2000). However, AID is not expressed in most cells types, and no other specific factors have been shown to cause DNA damage when R-loops arise in cells. Thus, many questions remain about the mechanisms that underlie the accumulation of DNA damage and genome instability associated with R-loop formation.

Here, we report that R-loops formed in the absence of mRNA processing factors or in the presence of camptothecin, an inhibitor of topoisomerase I, are actively processed by the nucleotide excision repair (NER) endonucleases XPF and XPG. Moreover, the transcription-coupled nucleotide excision repair (TC-NER) protein CSB is required for this processing, suggesting that R-loop processing is coupled to stalled transcription complexes. We also demonstrate that this mechanism is conserved in yeast, where it drives genomic instability. These findings reveal a function for TC-NER factors in R-loop processing and provide molecular insights into the processes underlying R-loop-induced DNA damage.

RESULTS

AQR Knockdown Induces R-Loop-Dependent DNA Damage

To investigate the mechanism of R-loop processing in human cells, we took advantage of the data from a genome-wide siRNA screen we previously carried out to identify factors involved in the maintenance of genome stability; highly enriched among the genes that induced DNA damage when knocked down were RNA processing factors. Surprisingly, overexpression of RNase H reversed the DNA damage observed after depletion of many of these RNA processing factors, suggesting that R-loops might be a source of this damage (Paulsen et al., 2009). We were particularly interested in one of these factors, Aquarius (AQR), a protein which is part of a subfamily of proteins possessing a conserved DEAxQ-like domain with putative RNA/DNA helicase activity (Fairman-Williams et al., 2010; Hirose et al., 2006). Interestingly, this subfamily includes Senataxin (SETX), which is thought to promote efficient transcriptional termination by resolving R-loops formed at specific loci (Skourti-Stathaki et al., 2011), and its yeast ortholog, Sen1, which prevents R-loop-mediated genome instability (Alzu et al., 2012; Mischo et al., 2011).

Knockdown of *AQR* robustly induced the DNA damage response (DDR), as evidenced by the phosphorylation of histone variant H2AX (termed γH2AX), a marker of DNA damage (see Figures S1A–S1C) (Paulsen et al., 2009). We also observed phosphorylation of the transcriptional repressor and DDR target KAP1 (termed P-KAP1) as well as the phosphorylation of CHK1 and RPA-2 (Figure 1A). These findings suggest that *AQR* knockdown ultimately leads to DSB formation and fork stalling. To test whether knockdown of *AQR* produced DSBs or induced DDR signaling by some other mechanism, we performed a neutral comet assay. The significant increase in comet tail moment we observed in AQR-depleted cells provides direct evidence for DSB formation and suggests that *AQR* knockdown does not simply induce DDR signaling (Figures 1B and 1C). Importantly, there was no significant difference in cell cycle progression upon *AQR* knockdown (Figure S1D). After prolonged knockdown, however, AQR-depleted cells accumulate in G2, consistent with the observed DSB formation and checkpoint activation (Figure S1E).

RNase H1 overexpression in AQR-depleted cells decreases γH2AX (Paulsen et al., 2009), and we found that it reduces P-KAP1 as well (Figure S1F). This finding suggests that RNA-DNA hybrids induced by the knockdown of *AQR* lead to DNA damage. To directly determine whether RNA-DNA hybrids accumulate upon *AQR* knockdown, we used a monoclonal antibody (S9.6) that specifically detects these hybrids (Boguslawski et al., 1986) to probe genomic DNA extracted from wild-type and AQR-depleted cells. We observed a 2-fold enrichment of RNA-DNA hybrids in AQR-depleted cells, which was abolished by pretreatment of the DNA with RNase H1 (Figure 1D). We also measured the nuclear S9.6 signal using confocal microscopy. Strikingly, high S9.6 signal was present in the nucleolus and mitochondria even before *AQR* knockdown. Although this is consistent with the known presence of RNA-DNA hybrids in these cellular compartments (El Hage et al., 2010; Aguilera and García-Muse, 2012), we also found that the nucleolar S9.6 signal persisted after RNase H1 treatment. This could be due to the presence of RNA species that are resistant to RNase H1, such as more structured RNA-DNA hybrids, or incomplete action of the nuclease in the nucleolus, where RNA-DNA hybrids are abundant. More importantly, in the absence of AQR, we observed an enrichment of nuclear RNA-DNA hybrids (Figure 1E), which we quantified after subtraction of the nucleolar signal (Figure 1F), and this enrichment could be reversed by treatment with RNase H1. Together, these data strongly suggest that the DNA damage observed in the absence of AQR results from the accumulation of R-loops.

R-Loops Induced upon *AQR* Knockdown Are Processed by the NER Endonucleases XPF and XPG

R-loops are thought to be open DNA structures with flap extremities, and we speculated that this structure might be recognized and processed by nucleases that act on related structures in the cell. The NER pathway uses two structure-specific flap endonucleases, XPF and XPG, to repair bulky lesions in the genome caused by a variety of DNA damaging agents. This process requires over 30 proteins that collectively recognize distortions caused by the lesion, excise a small lesion-containing oligonucleotide, and fill in the resulting gap by repair synthesis (Fagbemi et al., 2011). Intriguingly, purified XPF and XPG were previously shown to cleave R-loop structures formed at S regions of the immunoglobulin locus in vitro (Tian and Alt, 2000). Surprisingly, however, these nucleases were also shown to be dispensable for class switch recombination (CSR), suggesting XPF and XPG have no role in the processing of R-loops formed at S regions (Tian et al., 2004a, 2004b). Indeed, it has been shown that this processing depends upon AID (Muramatsu et al., 2000).

Because R-loops formed upon depletion of AQR induce unprogrammed DNA damage and are likely distinct from those formed at the immunoglobin locus, we wondered whether XPF or XPG might be able to act on R-loops induced by *AQR* knockdown in cells. To test this hypothesis, we assessed whether DDR signaling is induced upon knockdown of *AQR* in the absence of

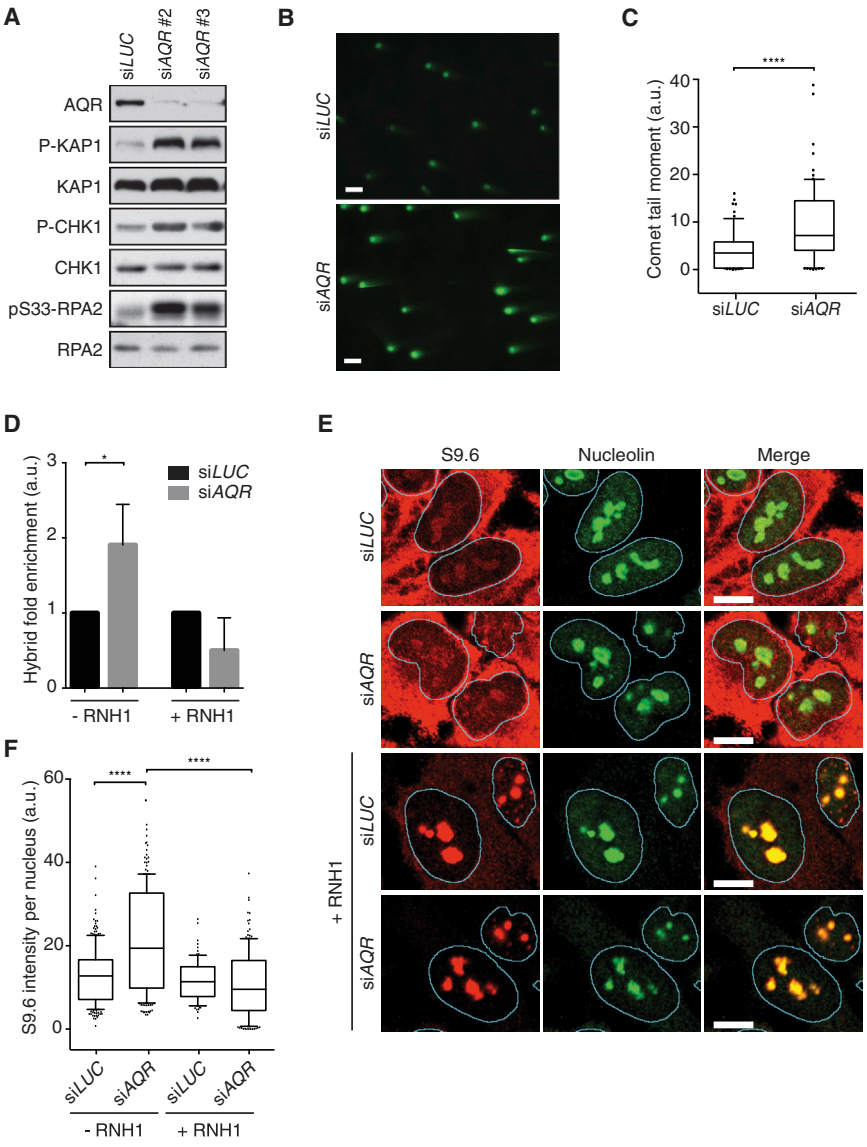


Figure 1. *AQR* Knockdown Leads to DSB Formation and R-Loop Accumulation

(A) P-KAP1, P-CHK1, and pS33-RPA2 levels in HeLa cells transfected with si*LUC* and two siRNAs directed against *AQR* for 72 hr.

(B) Neutral comet assay in HeLa cells transfected with si*LUC* or si*AQR*. Scale bar represents 50 μm.

(C) Quantification of comet tail moment for the experiment described in (B). a.u., arbitrary units. ****$p < 0.0001$.

(D) Quantification of RNA-DNA hybrids detected by slot blot with S9.6 antibody in HeLa cells, with fold enrichment relative to si*LUC* signal. Errors bars represent SEM of three biological replicates (*$p < 0.05$ by Student's t test).

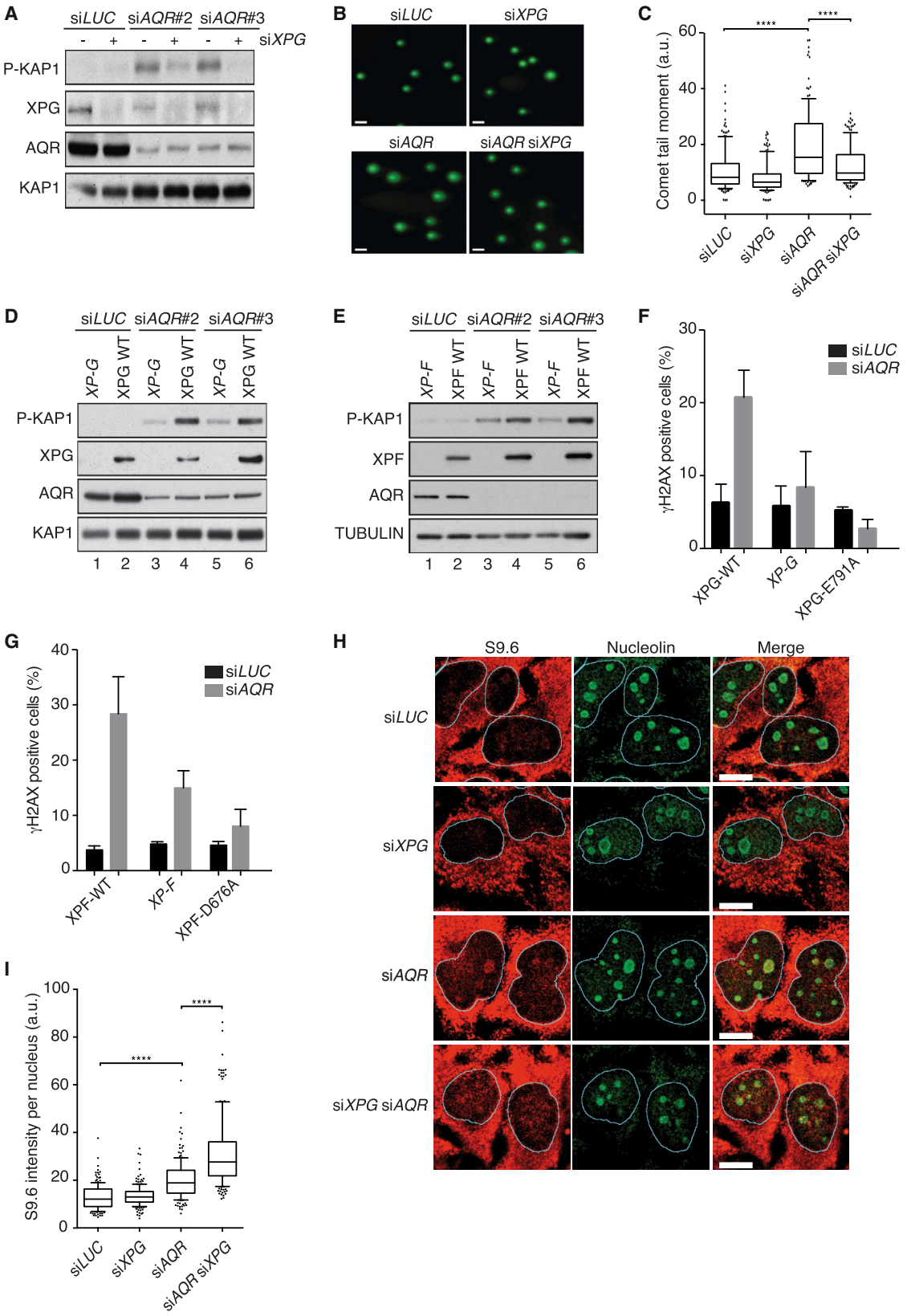
(E) Immunostaining with S9.6 (red) and nucleolin (green) antibodies in HeLa cells transfected with siRNA and fixed after 48 hr. The nucleus (stained with Hoechst) is outlined. Scale bar represents 10 μm. The levels of all panels were adjusted equally in Adobe Photoshop.

(F) Quantification of S9.6 signal per nucleus after nucleolar removal for the experiment described in (E), shown as box and whiskers plot. ****$p < 0.0001$. See also Figure S1.

XPG. Strikingly, double knockdown of both *XPG* and *AQR* dramatically reduced P-KAP1 observed after knockdown of *AQR* alone (Figure 2A). Importantly, DSB formation was also reduced, as indicated by a neutral comet assay (Figures 2B and 2C). Next, we investigated the effects of XPG in AQR-depleted cells by using immortalized XPG-deficient fibroblasts from a xeroderma pigmentosum (XP) patient, and an isogenic cell line complemented with wild-type XPG. *AQR* knockdown induced P-KAP1 in the XP patient cell line complemented with XPG (lanes 4 and 6 in Figure 2D). This phosphorylation was dramatically reduced when AQR was downregulated in the noncomplemented XPG-deficient cells (lanes 3 and 5 in Figure 2D). We also tested the role of XPF in this process using an analogous strategy. We found that P-KAP1 was also reduced when AQR was downregulated in XP patient cells deficient in XPF (Figure 2E). Thus, DSBs resulting from *AQR* knockdown are dependent upon both XPF and XPG. These findings suggest that XPF and XPG may be able to process R-loops resulting from *AQR* knockdown in cells.

To further test this idea, we asked whether the endonuclease activities of XPG and XPF are required to generate the DSBs observed after *AQR* knockdown or whether XPG and XPF simply play a structural role. To do so, we analyzed γH2AX levels by immunofluorescence in immortalized XPG- and XPF-deficient patient fibroblasts complemented with nuclease-dead forms of XPG or XPF proteins, XPG-E791A and XPF-D676A, respectively (Fagbemi et al., 2011). *AQR* knockdown in XPG- and XPF-deficient fibroblasts complemented with the wild-type form of XPG or XPF showed a high level of γH2AX compared to the control siRNA-transfected cell lines. In contrast, the γH2AX signal was much lower when *AQR* was knocked down in cells complemented with the nuclease-dead form of XPG or XPF (Figures 2F and 2G). Thus, the nuclease activities of both XPG and XPF are required to generate the DSBs observed upon AQR depletion.

Our data are consistent with the possibility that DSBs result from the direct processing of R-loops by XPF and XPG. However, another explanation is that decreases in transcription associated with *XPF* and *XPG* knockdown (Le May et al., 2010) reduce the levels of R-loops and, consequently, DNA damage. Alternatively, XPF and XPG might process DNA lesions that arise in the single-stranded DNA of the nontemplate strand, lesions that would require NER for repair after the RNA-DNA hybrid is resolved. To distinguish between



(legend on next page)

these models, we monitored the fate of RNA-DNA hybrids by confocal microscopy after codepletion of both XPG and AQR. We observed more RNA-DNA hybrids when both *XPG* and *AQR* were knocked down compared to knockdown of *AQR* alone (Figures 2H and 2I). If XPG was processing DNA lesions induced by the formation of R-loops through the classical NER pathway, R-loops would not be expected to accumulate. Similarly, decreases in transcription would not lead to R-loop accumulation. This result therefore suggests that XPG can induce DNA damage in cells by directly processing RNA-DNA hybrids.

R-Loops Induced by Defects in mRNA Processing or Camptothecin Treatment Are Processed by XPG

We then asked whether XPG acts exclusively on R-loops induced by *AQR* knockdown or if our findings could be extended to other factors which perturb R-loop dynamics. First, we tested the relationship between XPG and two mRNA processing factors that have been linked to R-loop formation and strong DDR activation, the splicing factor ASF/SF2 (Li and Manley, 2005) and the AQR-related helicase SETX (Alzu et al., 2012; Skourti-Stathaki et al., 2011). Concurrent knockdown of *XPG* in ASF- or SETX-depleted cells abrogated the DDR response, indicating that XPG processes R-loops induced by depletion of these factors (Figures 3A and 3B).

We also tested the effect of *XPG* knockdown on DNA damage induced by camptothecin (CPT), an inhibitor of topoisomerase I. Previous work showed that overexpression of RNase H1 decreased the induction of γH2AX following CPT treatment in postmitotic primary neurons and noncycling HeLa cells (Sordet et al., 2009), indicating that CPT generates R-loops that induce DSBs. To determine if this is also true when cells are cycling, we took advantage of a stable HeLa cell line we generated expressing tetracycline-inducible FLAG-tagged RNase H1. This allowed us to examine the DDR specifically in FLAG-positive cells. We found that overexpression of RNase H1 decreased the induction of γH2AX. We also found that *XPG* knockdown decreased the induction of γH2AX to a slightly greater extent than RNase H overexpression (Figure 3C). Moreover, the combined effect of *XPG* knockdown and RNase H1 expression was similar to that of *XPG* knockdown alone. These data indicate that R-loops formed in response to CPT treatment are processed into DSBs by XPG. Thus, XPG action is not limited to specific R-loops associated with AQR loss. Rather, XPG has a general role in the processing of R-loops into DSBs.

The Processing of R-Loops by XPG Drives Genome Instability in Yeast

Next, we considered whether XPG's influence on the DNA damage associated with R-loop formation is conserved in other species. The THO protein complex composed of the Tho2, Hpr1, Mft1, and Thp2 proteins is involved in transcription and RNA export (Chávez et al., 2000), and previous studies suggest that yeast THO mutants have a hyperrecombination phenotype that is dependent on the formation of R-loops (Huertas and Aguilera, 2003). Thus, we asked whether mutating *RAD2*, the yeast XPG homolog, would suppress this phenotype. To test this, we measured the frequency of R-loop-induced recombination (Prado and Aguilera, 1995; see Figure S2A for the assay description) in a single mutant of the THO complex *mft1Δ* and in the double mutant *rad2Δ mft1Δ*. The formation of recombinants markedly increased in the single mutant *mft1Δ* compared to the wild-type strain (Chávez et al., 2000). However, the recombination frequency returned to the wild-type level in the *rad2Δ mft1Δ* double mutant (Figure 3D), indicating that Rad2 promotes genome instability in the absence of mRNA processing factors. We confirmed this result using an alternative recombination assay (Figure S2B). These findings suggest that the molecular mechanisms responsible for the processing of R-loop are broadly conserved.

TC-NER Factors Are Required for the Processing of R-Loops in Mammalian Cells

Lastly, we considered whether XPF and XPG act on R-loops through an NER-like pathway or through another mechanism involving their nuclease activities. During canonical NER, XPA plays a critical role in positioning NER factors, including the XPF and XPG nucleases, for incision and repair (Fagbemi et al., 2011). Thus, we asked whether XPA can contribute to R-loop processing. We found that XPA depletion clearly suppressed the DDR activation observed upon knockdown of *AQR* (Figure 4A). Next, we asked whether the TFIIH complex is required for the processing of R-loops. Two components of TFIIH, XPB and XPD, play an essential role during NER, using their respective DNA-dependent ATPase and helicase activities to open the DNA around the lesion (Fagbemi et al., 2011). We found that concurrent knockdown of either *XPB* or *XPD* with *AQR* dramatically reduced P-KAP1 compared to knockdown of *AQR* alone (Figure 4B). Similar results were obtained in immortalized XPD-deficient patient fibroblasts, and an isogenic cell line complemented with a wild-type form of XPD; *AQR* knockdown induced γH2AX in the complemented cell line, while the

Figure 2. R-Loop-Induced DNA Damage Depends on XPF and XPG

(A) P-KAP1 level in HeLa cells transfected with si*XPG* or si*LUC* 24 hr prior to transfection with si*LUC* or si*AQR*.

(B) Neutral comet assay in HeLa cells treated as in Figure 2A. Scale bar represents 50 μm.

(C) Quantification of comet tail moment for the experiment described in (B). a.u., arbitrary units. ****$p < 0.0001$.

(D and E) P-KAP1 level in XPG and XPF patient cell lines either complemented or not with the corresponding wild-type proteins, and transfected with si*LUC*, si*AQR*#2, or si*AQR*#3.

(F and G) Quantification of percent γH2AX-positive cells in XPG and XPF patient cell lines either complemented or not with the wild-type or nuclease-dead proteins, and transfected with indicated siRNA (SEM, n = 3).

(H) Immunostaining with S9.6 (red) and nucleolin (green) antibodies in HeLa cells transfected with si*XPG* or si*LUC* 24 hr before transfection with si*LUC* or si*AQR*. A merge of the two channels is shown, with the nucleus (stained with Hoechst) outlined. Scale bar represents 10 μm. The levels of all panels were adjusted equally in Adobe Photoshop.

(I) Quantification of S9.6 signal per nucleus after nucleolar removal for the experiment described in (H), shown as box and whiskers plot. ****$p < 0.0001$.

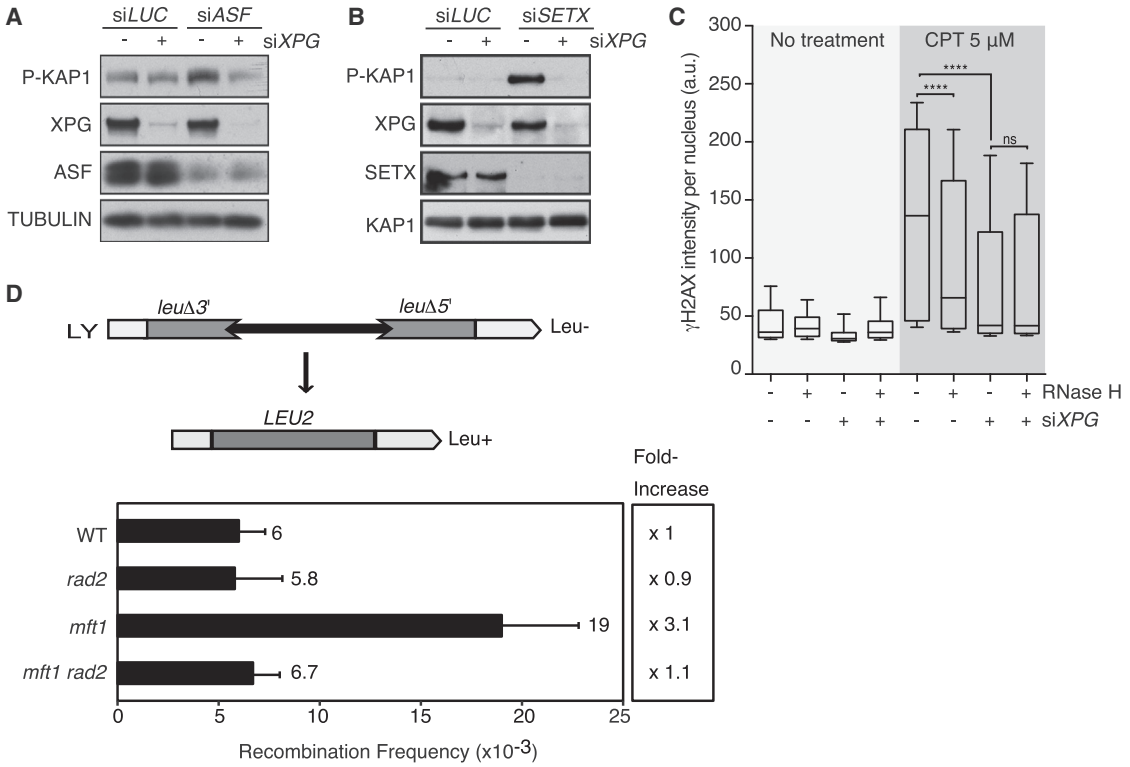


Figure 3. The Processing of R-Loops by the Endonucleases XPF and XPG Is a General and Conserved Mechanism

(A and B) P-KAP1 level in HeLa cells transfected with si*XPG* or si*LUC* 24 hr before transfection with si*ASF* or si*SETX*.

(C) γH2AX intensity in HeLa-TetON-RNase H1 cells transfected with si*LUC* or si*XPG* for 48 hr and treated for 2 hr with 5 μM camptothecin. Doxycycline (500 ng/μl) was added in combination with siRNAs where indicated. a.u., arbitrary units.

(D) Frequencies of recombination in the LY direct-repeat system. Data represent the median of three to four independent experiments. Error bars represent the standard error of the median (SEM, n = 3, 4). See also Figure S2 and Table S1.

γH2AX signal was reduced in the noncomplemented cell line (Figure S3A). These findings indicate that R-loop processing is a concerted action that requires the NER factors, XPA, XPB, and XPD. They also indicate that R-loop processing is not due to the unregulated activity of these endonucleases on flap structures but instead is a result of classical NER-like events.

The endonucleases XPF and XPG, as well as the factors XPA, XPB, and XPD, are involved in two forms of NER, global genome repair (GG-NER) and transcription-coupled repair (TC-NER) (Svejstrup, 2002). To delineate which of these pathways was responsible for R-loop-induced DNA damage or if both pathways are involved, we examined the effect of depleting factors specific to each form of NER. XPC recognizes helix-distorting lesions during GG-NER and subsequently mobilizes NER proteins (Sugasawa et al., 1998). We found that P-KAP1 did not decrease upon depletion of XPC when *AQR* was knocked down (Figure 4C), suggesting that R-loop processing is not mediated by the GG-NER pathway. Similar results were observed in immortalized XPC-deficient fibroblasts from an XP patient, and in an isogenic cell line complemented with a wild-type form of XPC; *AQR* knockdown induced γH2AX and P-KAP1 in both the complemented and noncomplemented XP patient cell lines (Figures S3B and S3C). To probe the role of TC-NER in R-loop processing, we examined the effect of depleting CSB, which recruits downstream NER proteins when a DNA lesion is encountered on the transcribed strand (Fousteri et al., 2006). Surprisingly, we found that knockdown of *CSB* reduced the high P-KAP1 observed after knockdown of *AQR* (Figure 4D). The P-KAP1 induced upon knockdown of *SETX* was strongly reduced upon depletion of CSB as well (Figure S3D). These data suggest that the TC-NER factor CSB is required for the processing of R-loops into DSBs. Since CSB but not XPC is required for the processing of R-loops, we propose that TC-NER factors act in a noncannonical manner to generate DNA damage when R-loops form during transcription.

DISCUSSION

Our findings reveal a conserved molecular mechanism by which R-loops are actively processed to DSBs, and they indicate that DSBs do not simply result from the collision of a replication fork with an R-loop. Surprisingly, this processing involves several NER factors, including XPA, the TFIIH subunits XPB and XPD, and the endonucleases XPF and XPG. Because we also observe that the TC-NER protein CSB, but not XPC, is required for this processing, we conclude that R-loop processing requires the concerted action of TC-NER factors. The involvement of XPF and XPG in the processing of R-loops induced by depletion of

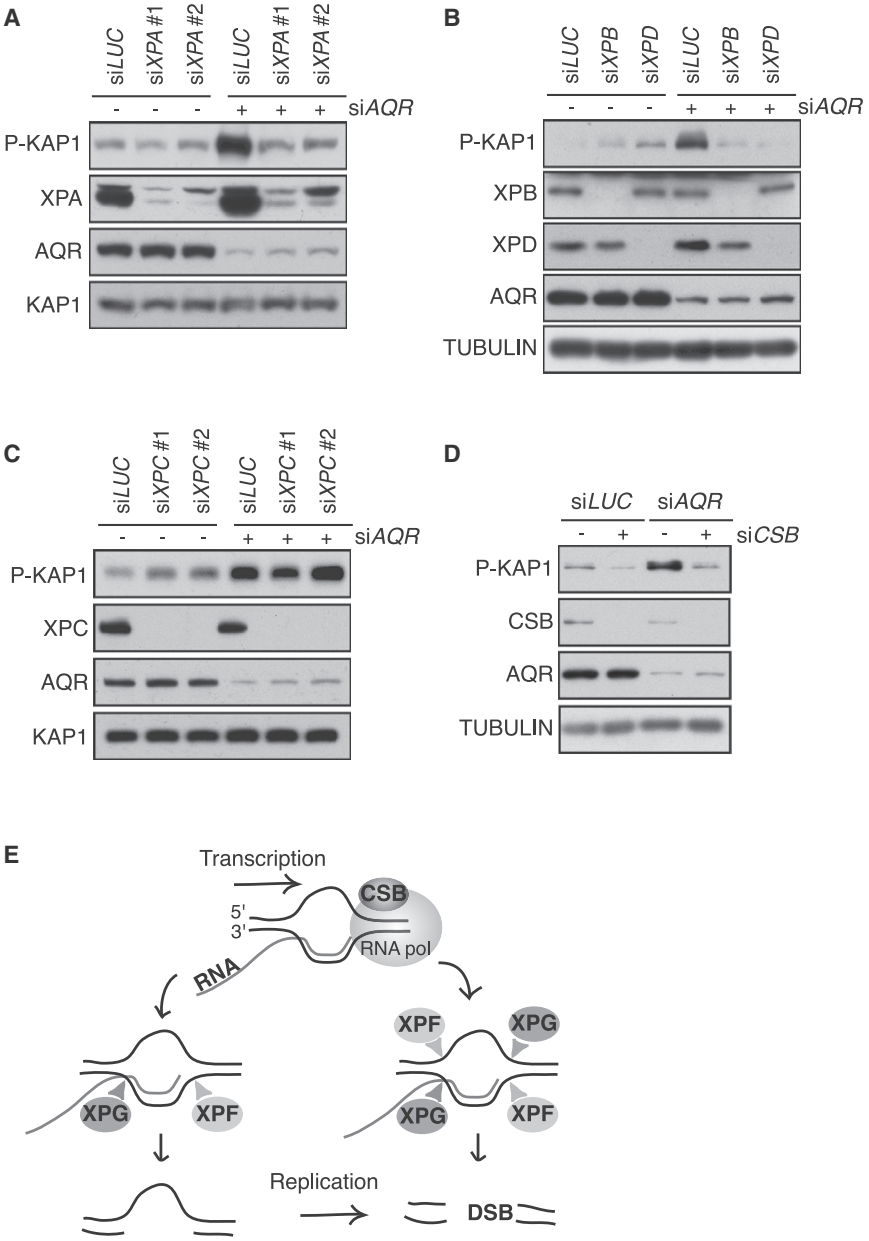


Figure 4. R-Loop Processing Requires TC-NER Factors but Not XPC

(A) P-KAP1 level in HeLa cells transfected with si*XPA* or si*LUC* 24 hr before transfection with si*LUC* or si*AQR*.

(B) P-KAP1 level in HeLa cells transfected with si*XPB*, si*XPD*, or si*LUC* 24 hr before transfection with si*LUC* or si*AQR*.

(C) P-KAP1 level in HeLa cells transfected with si*XPC* or si*LUC* 24 hr before transfection with si*LUC* or si*AQR*.

(D) P-KAP1 level in HeLa cells transfected with si*CSB* or si*LUC* 24 hr before transfection with si*LUC* or si*AQR*.

(E) Model for how an R-loop is processed into a DSB. The stalling of the RNA polymerase allows CSB to recruit the endonucleases XPF and XPG. XPF and XPG generate a gap that can be converted into a DSB through DNA replication, and/or XPF and XPG cleave the R-loop on both strands, producing a DSB. See also Figure S3.

RNA processing factors or CPT treatment, and not R-loops involved in CSR, indicates there are distinct pathways for processing different types of R-loop structures. We suggest that R-loops associated with paused transcription complexes may be the target of TC-NER factors. This is a new and unexpected role for these factors outside the transcription-coupled repair of DNA damage and suggests that TC-NER factors affect genome stability in diverse, and potentially detrimental, ways.

There are a variety of distinctions between the R-loop structure and the lesion-containing structures processed by XPF and XPG during classical NER which raise interesting questions about precisely how R-loops are cut by TC-NER factors. Because CSB is required, a parsimonious explanation would be that R-loop processing is triggered upon stalling of the RNA polymerase complex by an R-loop; this pause would allow CSB to recruit XPF and XPG for processing, as it does when it stalls at a DNA lesion. However, our observation that NER factors lead to DSB formation is unexpected in the classical NER context, because NER factors typically generate a single-strand DNA gap during excision of a lesion. In the context of R-loops, there are potential substrates for the flap endonucleases on both the transcribed and nontranscribed strand, and at both ends of the R-loop. Thus, processing could lead directly to DSB formation, or to the formation of nicks or gaps which are known to ultimately cause fork collapse and DSBs formation in S phase (Figure 4E). Indeed, diverse studies suggest that DNA replication is required for R-loop-induced genome instability, and the activation of ATR suggested by the phosphorylation of RPA and CHK1 indicates there may be effects of these structures or their processed intermediates in S phase (Alzu et al., 2012; Gan et al., 2011; Tuduri et al., 2009; Wellinger et al., 2006; Yüce and West, 2013). It is also possible that R-loop-associated TC-NER factors act directly on forks that collide with R-loop structures or that processing of the R-loop is coordinated with DNA replication. Regardless of the precise structure processed, the requirement for TC-NER factors in generating the DSBs associated with R-loop formation reveals the molecular mechanism of R-loop-induced genome instability.

Importantly, R-loops further accumulate in cells depleted for AQR when *XPG* is also knocked down. This suggests that XPG clears R-loops in the absence of efficient mRNA processing, and it also raises the possibility that NER factors may be a clearance pathway for naturally occurring R-loops. However,

whether TC-NER-dependent R-loop processing is beneficial to cells is still unclear. Although TC-NER-dependent processing leads to DNA damage, it is possible that this is preferred to the continued persistence of R-loops, which could pose additional problems for the cells. R-loop processing may be a cost that comes with the ability to rapidly repair lesions during transcription. Given that the effect of XPG on R-loops is observed in both yeast and humans, we speculate that this processing is conserved and has some long-term benefit to the organism.

R-loops have been found to cover a substantial portion of the genome and to play fundamental roles in various cellular processes (Chan et al., 2014; Ginno et al., 2012, 2013; Skourti-Stathaki and Proudfoot, 2014), suggesting that not all R-loops induce genome instability. Thus, our work raises the fascinating question of how certain types of R-loops are specifically protected from the deleterious effects of the TC-NER machinery and what drives DSB formation and genome instability at other R-loop sites. In the absence of splicing factors or in the presence of the topoisomerase inhibitor CPT, we and others have shown that R-loops accumulate (Alzu et al., 2012; Li and Manley, 2005; Skourti-Stathaki et al., 2011; Sordet et al., 2009). Thus, one possibility is that these unscheduled R-loops saturate the clearance pathways that normally act to resolve these structures or prevent their formation, thereby allowing TC-NER factors to cleave R-loops aberrantly. Another possibility is that unscheduled R-loops exist in different chromosomal contexts which may affect their processing. Consistent with this idea, R-loops arise at new genomic loci in yeast mutants of RNaseH or Senataxin (Chan et al., 2014). Lastly, the dynamics of unscheduled R-loops may differ from those of regulatory R-loops, allowing the latter to escape the deleterious effect of TC-NER factors. Indeed, the dynamic formation and resolution of R-loops are needed for efficient transcriptional termination (Skourti-Stathaki et al., 2011).

In summary, we demonstrate that the processing of unscheduled R-loops by the TC-NER pathway poses a threat to genome stability. R-loops have been observed at some common fragile sites (Helmrich et al., 2011) as well as the proto-oncogene MYC (Duquette et al., 2005), and based on our results, we propose that TC-NER-dependent R-loop processing contributes to genome instability and cancer progression by stimulating recombination at R-loop sites. Recently, the accumulation of R-loops has also been implicated in the silencing of critical genes that are associated with neurodegenerative diseases and which contain repeated DNA sequences (Colak et al., 2014; Groh et al., 2014; Haeusler et al., 2014; Loomis et al., 2014). Thus, the formation of R-loops may be detrimental in different ways; R-loop processing by TC-NER factors might promote genome rearrangements leading to cancer, while R-loop stabilization may be more relevant to neurodegenerative diseases. Lastly, because TC-NER factors may also play a role in the clearance of R-loops, it is tempting to speculate that some phenotypes observed in XP/CS patients result from a defect in R-loop processing. The mechanistic insights provided by this work may ultimately point the way to strategies for the modulation of R-loop formation and processing that could be used for the treatment of these and others human diseases.

EXPERIMENTAL PROCEDURES

Cell culture, western blotting, and immunofluorescence were performed using standard methods. Neutral comet assay was performed as described previously (Tuduri et al., 2009). For slot blot, total genomic DNA was blotted on nylon membrane and probed with either S9.6 antibody or denatured and probed with the single-strand DNA antibody. The yeast LY recombination assays were performed as previously described (Luna et al., 2005). Detailed information regarding methodology and any associated references are available in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2014.10.020.

AUTHOR CONTRIBUTIONS

K.A.C. and J.S. designed the study and analyzed the data. J.S., C.T.S., and M.L.G.-R. performed the experiments and analyzed the data. R.D.P. generated the preliminary data motivating the study. A.A. analyzed the data and edited the manuscript. J.S. and K.A.C. wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank Julie Rageul (Stony Brook University), Frédéric Chédin (University of California, Davis), and Xialu Li (National Institute of Biological Sciences, Beijing) for providing helpful reagents and advice. We would also like to thank Yea-Lih Lin (Institute of Human Genetics, France) for technical advice, and members of the Cimprich lab for thoughtful discussion. This work was supported by the Spanish Ministry of Economy and Competitiveness (grants BFU2010-16372 and CSD2007-00015), the European Union (FEDER) to A.A., and an NIH grant (GM100489) to K.A.C.

Received: July 29, 2014
Revised: October 7, 2014
Accepted: October 22, 2014
Published: November 26, 2014

REFERENCES

Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol. Cell *46*, 115–124.

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., and Liberi, G. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell *151*, 835–846.

Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., and Carrico, R.J. (1986). Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. J. Immunol. Methods *89*, 123–130.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y.T., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet. *10*, e1004288.

Chávez, S., Beilharz, T., Rondón, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q., Lithgow, T., and Aguilera, A. (2000). A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. EMBO J. *19*, 5824–5834.

Chen, Y.-Z., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M.L., Rabin, B.A., et al. (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). Am. J. Hum. Genet. *74*, 1128–1135.

Colak, D., Zaninovic, N., Cohen, M.S., Rosenwaks, Z., Yang, W.Y., Gerhardt, J., Disney, M.D., and Jaffrey, S.R. (2014). Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. Science *343*, 1002–1005.

Duquette, M.L., Pham, P., Goodman, M.F., and Maizels, N. (2005). AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. Oncogene *24*, 5791–5798.

El Hage, A., French, S.L., Beyer, A.L., and Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev. *24*, 1546–1558.

Fagbemi, A.F., Orelli, B., and Schärer, O.D. (2011). Regulation of endonuclease activity in human nucleotide excision repair. DNA Repair (Amst.) *10*, 722–729.

Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Curr. Opin. Struct. Biol. *20*, 313–324.

Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H.F. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. Mol. Cell *23*, 471–482.

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., and Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes Dev. *25*, 2041–2056.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell *45*, 814–825.

Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I., and Chédin, F. (2013). GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. *23*, 1590–1600.

Groh, M., Lufino, M.M.P., Wade-Martins, R., and Gromak, N. (2014). R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. PLoS Genet. *10*, e1004318.

Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A.J., Shaw, P.G., Kim, M.-S., Maragakis, N.J., Troncoso, J.C., Pandey, A., Sattler, R., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature *507*, 195–200.

Hamperl, S., and Cimprich, K.A. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. DNA Repair (Amst.) *19*, 84–94.

Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol. Cell *44*, 966–977.

Hirose, T., Ideue, T., Nagai, M., Hagiwara, M., Shu, M.-D., and Steitz, J.A. (2006). A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. Mol. Cell *23*, 673–684.

Huertas, P., and Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell *12*, 711–721.

Le May, N., Mota-Fernandes, D., Vélez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010). NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. Mol. Cell *38*, 54–66.

Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell *122*, 365–378.

Li, X., Niu, T., and Manley, J.L. (2007). The RNA binding protein RNPS1 alleviates ASF/SF2 depletion-induced genomic instability. RNA *13*, 2108–2115.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709–715.

Loomis, E.W., Sanz, L.A., Chédin, F., and Hagerman, P.J. (2014). Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. PLoS Genet. *10*, e1004294.

Luna, R., Jimeno, S., Marín, M., Huertas, P., García-Rubio, M., and Aguilera, A. (2005). Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. Mol. Cell *18*, 711–722.

Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol. Cell *41*, 21–32.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell *102*, 553–563.

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., et al. (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell *35*, 228–239.

Prado, F., and Aguilera, A. (1995). Role of reciprocal exchange, one-ended invasion crossover and single-strand annealing on inverted and direct repeat recombination in yeast: different requirements for the RAD1, RAD10, and RAD52 genes. Genetics *139*, 109–123.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. *28*, 1384–1396.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell *42*, 794–805.

Sordet, O., Redon, C.E., Guirouilh-Barbat, J.E.E., Smith, S., Solier, S.E.P., Douarre, C.E.L., Conti, C., Nakamura, A.J., Das, B.B., Nicolas, E., et al. (2009). Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks. EMBO Rep. *10*, 887–893.

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., and Hieter, P. (2012). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes Dev. *26*, 163–175.

Sugasawa, K., Ng, J.M., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A.P., Hanaoka, F., Bootsma, D., and Hoeijmakers, J.H. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol. Cell *2*, 223–232.

Svejstrup, J.Q. (2002). Mechanisms of transcription-coupled DNA repair. Nat. Rev. Mol. Cell Biol. *3*, 21–29.

Tian, M., and Alt, F.W. (2000). Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. J. Biol. Chem. *275*, 24163–24172.

Tian, M., Jones, D.A., Smith, M., Shinkura, R., and Alt, F.W. (2004a). Deficiency in the nuclease activity of xeroderma pigmentosum G in mice leads to hypersensitivity to UV irradiation. Mol. Cell. Biol. *24*, 2237–2242.

Tian, M., Shinkura, R., Shinkura, N., and Alt, F.W. (2004b). Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. Mol. Cell. Biol. *24*, 1200–1205.

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., et al. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol. *11*, 1315–1324.

Wahba, L., Amon, J.D., Koshland, D., and Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol. Cell *44*, 978–988.

Wellinger, R.E., Prado, F., and Aguilera, A. (2006). Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. Mol. Cell. Biol. *26*, 3327–3334.

Yu, K., Chédin, F., Hsieh, C.-L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. *4*, 442–451.

Yüce, O., and West, S.C. (2013). Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. Mol. Cell. Biol. *33*, 406–417.