Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids

Petermann, Eva; Lan, Li; Zou, Lee

DOI:
10.1038/s41580-022-00474-x

Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids

Eva Petermann[1,2], Li Lan[3,4], and Lee Zou[5]

1. Institute of Cancer and Genomic Sciences,
University of Birmingham,
Birmingham, UK

2. Birmingham Centre for Genome Biology
University of Birmingham,
Birmingham, UK

3. Massachusetts General Hospital Cancer Center,
Harvard Medical School,
Charlestown, MA, USA

4. Department of Radiation Oncology,
Massachusetts General Hospital,
Harvard Medical School,
Boston, MA, USA

5. Department of Pathology,
Massachusetts General Hospital,
Harvard Medical School,
Boston, MA, USA

Corresponding author: zou.lee@mgh.harvard.edu

Abstract

RNA–DNA hybrids are generated during transcription, DNA replication and DNA repair, and are crucial intermediates in these processes. When RNA–DNA hybrids are stably formed in double-stranded DNA, they displace one of the DNA strands and give rise to a three-stranded structure called an R-loop. R-loops are widespread in the genome and enriched at active genes. R-loops have important roles in regulating gene expression and chromatin structure, but they also pose a threat to genomic stability, especially during DNA replication. To keep the genome stable, cells have evolved a slew of mechanisms to prevent aberrant R-loop accumulation. Although R-loops can cause DNA damage, they are also induced by DNA damage and act as key intermediates in DNA repair, such as transcription-coupled and RNA-templated DNA break repair. When the regulation of R-loops goes awry, pathological R-loops accumulate, which contribute to diseases such as neurodegeneration and cancer. In this Review, we discuss the current understanding of the sources of R-loops and other RNA–DNA hybrids, mechanisms that suppress and resolve these structures, the impact of these structures on DNA repair and genomic stability, and opportunities to therapeutically target pathological R-loops.

Introduction

R-loops are a three-stranded nucleic acid structure consisting of RNA–DNA hybrid and a displaced strand of DNA. R-loops and other RNA–DNA hybrids, such as those formed at DNA replication forks and at processed double-stranded DNA (dsDNA) ends, are important intermediates in transcription, DNA replication and DNA repair. Studies in multiple organisms have revealed that R-loops have crucial roles in cellular processes such as chromatin organization and chromosome segregation[1-4]. The RNA–DNA hybrids formed at DNA replication forks and processed DNA ends do not displace single-stranded DNA (ssDNA) and thus do not constitute R-loops, but are also important for DNA replication and repair[4,5]. Paradoxically, emerging evidence suggests that R-loops are a major source of genomic instability, highlighting the importance of proper regulation of R-loops.

Recent genetic and biochemical studies have uncovered numerous regulators of R-loops, providing insights into how R-loops are generated and resolved in various cellular and chromosomal

contexts. The functions of R-loops in different cellular processes, especially in DNA repair are also increasingly appreciated. New technical advances in mapping R-loops, such as DNA–RNA immunoprecipitation sequencing (DRIP-seq) and ribonucleases H1 (RNaseH1) chromatin immunoprecipitation (R-ChIP)[6,7], have allowed us to assess the distribution and dynamics of R-loops throughout the genome and detect changes in pathological contexts. Notably, alterations of R-loop levels are detected in cancer models and are associated with transcription–replication conflicts **[G]** and genomic instability[8-11], offering insights into the source of genomic instability in cancer cells and a new opportunity for cancer therapy.

In this Review, we first discuss new findings of the sources of R-loops and RNA–DNA hybrids in both physiological and pathological conditions. We then discuss the factors and pathways that generate or resolve R-loops and other RNA–DNA hybrids in various cellular contexts, and effects of R-loops and RNA–DNA hybrids on the genome, including positive and negative effects at active genes, DNA replication forks and DNA damage sites, are also discussed. Finally, we discuss the causes and consequences of aberrant R-loops in cancer cells, and the opportunity to target aberrant R-loops in cancer therapy. Through this Review, we aim to provide an updated and integrated view of the fast-evolving research of R-loops and RNA–DNA hybrids.

[H1] Sources and distribution of hybrids

The sources of R-loops and RNA–DNA hybrids, as well as their functions, are distinct in various cellular, chromosomal and sequence contexts. In this section, we discuss how R-loops and RNA–DNA hybrids form naturally during transcription, DNA replication and DNA repair throughout the genome. Furthermore, R-loops accumulate in specific regions of the genome, such as at telomeres and centromeres, and in mitochondrial DNA.

[H2] Co-transcriptional R-loops

During replication, the nascent RNA very transiently anneals to the DNA template within the active site of the the RNA polymerase, giving rise to a short, transient RNA–DNA hybrid, which is resolved by release of the nascent RNA through a dedicated channel[12]. R-loops form when nascent RNA transiently re-anneals back onto the template behind the RNA polymerase. At genes transcribed by RNA polymerase II (Pol II), which include the protein-coding genes, co-

transcriptional R-loops mainly accumulate at promoters and transcription start sites (TSSs)[13] and at transcription termination sites (TTSs)[14,15]. R-loops at promoters are involved in transcription activation[16], and R-loops at Pol II pause sites near G-rich 'terminator' elements aid transcription termination and formation of gene-repressive chromatin[14,15] (Fig. 1a). In immunoglobulin genes, R-loops also accumulate in the highly repetitive, GC-rich switch regions, where they aid class switch recombination **[G]** [17].

Both increased Pol II transcription[9,18,19] and conversely Pol II pausing[20], have been linked to R-loop accumulation. Increased Pol II transcription generates more RNA that can form R-loops, whereas Pol II pausing increases the chance of RNA hybridization with the DNA template. Active ribosomal DNA repeats are highly transcribed by Pol I, and accumulate R-loops that can interfere with Pol I transcription even in normal conditions[21]. High levels of R-loops, particularly in the 5' regions of the 18S genes, correlate with Pol I pile-up[21]. Pol II transcription of the intergenic spacers between ribosomal DNA genes also generates R-loops that limit Pol I transcription of intergenic non-coding RNAs, which could otherwise disrupt nucleolar structure[22]. The 5S rRNA genes, which are transcribed by Pol III, are also prone to R-loop formation in normal growth conditions, and other Pol III genes such as tRNA genes accumulate R-loops only in absence of RNase H activity (see below)[23].

[H2] R-loops at telomeres and centromeres

Telomeric repeat-containing RNA (TERRA) is a long non-coding RNA (lncRNA) transcribed by Pol II from telomeres and sub-telomeric regions[24,25]. TERRA contains the UUAGGG repeats, which can hybridize with the C-rich strand of telomere DNA **[G]**. In human cells, TERRA associates with telomeres[24], interacts with a number of telomere-binding proteins and chromatin modulators such as telomere repeat binding factor 2 (TRF2) and ATRX, and is important for telomere maintenance[26]. TERRA-generated R-loops are detected at telomeres in human cells[27] (Fig. 1b). A recent study showed that RNA consisting of the UUAGGG repeats is sufficient to form telomeric R-loops in *trans* through a RAD51 recombinase-mediated mechanism, and that the formation of telomeric R-loops increases telomere fragility[28]. Thus, telomeric R-loops allow TERRA to associate with and maintain telomeres, but they also generate genomic instability. The

proper regulation of TERRA and telomeric R-loops is likely important for the proper function of telomeres.

Similar to telomeres, centromeres are also transcribed by Pol II[29]. RNA transcripts of centromeric and pericentric satellite DNA are detected in human cells. The RNA transcripts of α satellite repeats associate with centromeres[30,31], suggesting that they form R-loops in *cis* (Fig. 1b). Centromeric RNAs are required for the association of histone H3-like centromeric protein A (CENPA) and centromere protein C (CENP-C) with centromeres[30,31], rendering these RNAs important structural components of centromeres. Centromeric R-loops are readily detected in mitotic chromosomes[32]. The kinase ataxia telangiectasia mutated and Rad3-related (ATR) is recruited to centromeres during mitosis in an R-loop-dependent manner, allowing ATR to promote Aurora B activation through CHK1[32] (Fig. 1b). This mitosis-specific and centromere-specific mechanism of ATR activation ensures accurate microtubule–kinetochore attachment and faithful chromosome segregation. However, although centromeric R-loops are important for centromere assembly and function, they are also a source of replication stress. In S phase, CENPA is required for suppressing transcription and R-loops at centromeres, thereby protecting centromeres from DNA damage[33]. Thus, centromeric R-loops have both positive and negative effects on centromeres, and the levels of centromeric R-loops and their associated proteins may be intricately orchestrated during the cell cycle.

[H2] Hybrids generated during replication

The most prominant source of replication-associated RNA–DNA hybrids is lagging-strand replication, during which the DNA polymerase α (Pol α)-primase complex synthesizes an 8–10-nucleotide-long RNA primer every ~200 nucleotides[34] (Fig. 2a). However, these hybrids are routinely removed during Okazaki fragment maturation, when the RNA primers and the majority of their Polα-synthesized DNA extensions are displaced by DNA polymerase δ and the resulting flaps are removed by flap endonuclease 1 (FEN1), acting alone or with DNA2 (Ref.[35]). RNA-primer-based RNA–DNA hybrids might also be produced during bypass of DNA lesions and repriming by Polα-primase or DNA-directed primase/polymerase protein (PrimPol), although in this case the mechanism of RNA–DNA hybrid removal is much less certain[36,37] (Fig. 2a). Another form of RNA–DNA hybrid that is frequently generated during DNA replication is a

misincorporated ribonucleotide. The high concentration of ribonucleotides in the cell leads to their misincorporation by replicative DNA polymerases at an estimated rate of once every 7.6 kb, which would equate to 1 million sites per cell division[38]. Misincorporated ribonucleotides are proposed to be the most frequent genomic lesion[38] and require ribonucleotide excision repair (RER) for their removal[39] (Fig. 2a).

[H2] DNA damage induced hybrids

In human cells, RNA–DNA hybrids are detected at sites of DNA double-stranded breaks (DSBs). These hybrids are formed by RNA–DNA hybridization on the ssDNA overhangs of DSBs or by invasion of RNA into dsDNA. Several types of RNA contribute to the hybrids at DSBs. Dicer-generated and Drosha-generated small RNAs are reported to form hybrids around DSBs[40,41]. DNA damage-induced long non-coding RNAs (dilncRNAs) also form hybrids at DSBs[42]. The synthesis of dilncRNAs is linked to Pol II recruited to DSBs by the MRE11–RAD50–NBS1 (MRN) complex[42,43] (Fig. 2b). Pol III also contributes to RNA synthesis at DSBs[44]. These findings suggest that *de novo* RNA synthesis at DSBs promotes the formation of RNA–DNA hybrids. However, genome-wide mapping of RNA–DNA hybrids following induction of DSBs at specific sites reveals that the increase in hybrid levels primarily occurs in transcriptionally active regions[45], suggesting that preexisting RNA transcripts have a predominant role in the formation of damage-induced RNA–DNA hybrids (Fig. 2b). Consistently, when DSBs and single-stranded breaks (SSBs) are induced by reactive oxygen species (ROS) at a specific chromosomal locus, hybrids are only detected in the presence of local transcription[46]. It was proposed that the pausing of Pol II by DSBs and SSBs leads to formation of R-loops[4,46]. How Pol II is paused by DSBs and SSBs is not understood. The exposure of ssDNA in R-loops at active genes may also allow *de novo* RNA synthesis[4], thereby contributing to the formation of RNA–DNA hybrids at DNA damage sites.

[H2] Mitochondrial R-loops

R-loops also accumulate in mitochondrial DNA (mtDNA), which consists of a heavy (H) strand and a light (L) strand with distinct nucleotide compositions[47]. R-loops are detected across the control region of mtDNA **[G]**, and the 3' ends of the RNAs in R-loops coincide with ori-H, the origin of heavy-strand synthesis. The formation of R-loops at ori-H displaces ssDNA and initiates RNA priming for heavy-strand synthesis. After the nascent heavy strand is extended beyond ori-

L, the synthesis of the light strand is initiated from ori-L[48], allowing mtDNA to be fully duplicated. Ribonuclease H1 (RNaseH1) is crucial for R-loop processing and DNA replication in mitochondria[49]. Although R-loops are important for mtDNA replication, high levels of R-loops cause mtDNA instability. The mitochondrial degradosome complex, which consists of the helicase SUV3 and the ribonuclease polyribonucleotide nucleotidyltransferase 1, prevents the accumulation of pathological R-loops in mitochondria[50].

[H2] R-loops generated during genome editing

In addition to the naturally occurring R-loops and RNA–DNA hybrids, R-loops can be artificially generated during genome editing by CRISPR–Cas[51,52]. In CRISPR–Cas systems, Cas proteins in complex with small guide RNAs (gRNAs) recognize and cleave target DNA sequences. During the recognition of target DNA, the gRNA invades dsDNA and forms RNA–DNA hybrids. The RNA–DNA hybrids formed by Cas9 and gRNAs are about 16–20-nucleotide long. The crystal structure of the Cas9–gRNA–dsDNA complex shows that the RNA–DNA hybrid is largely buried inside the Cas9 protein[53]. The formation of these RNA–DNA hybrids displaces the non-target strand of DNA and positions the RNA–DNA hybrid and the ssDNA near the active sites of the HNH and RuvC nuclease domains of Cas9, respectively[53]. The ability of CRISPR–Cas systems to recognize target DNA by forming R-loops also enables them to mediate base editing and prime editing[54,55]. Other Cas proteins also use R-loops to recognize and cleave targets[56,57].

[H1] Impact of hybrids on genomic stability

Although R-loops and RNA–DNA hybrids have contributing roles in many cellular processes, they are also a source of genomic instability when they accumulate at abnormally high levels, or in the wrong context. For example, aberrant R-loops impose a threat to replication forks and increase the nucleolytic cleavage of DNA.

[H2] R-loop–replication fork collisions

Collisions between R-loops and replication forks are a source of DNA damage in bacteria, yeast and human cells[58-60]. Using R-loop-forming episomes carrying a unidirectional Epstein-Barr virus (EBV) replication origin, head-on collisions of R-loops and replication forks were shown to generate more DNA damage than co-directional collisions[59] (Fig. 3a). Head-on collisions also

increase R-loop levels in the episomes and in the genome, whereas co-directional collisions decrease R-loop levels[59]. Furthermore, the intermediates generated by head-on collisions or co-directional collisions are processed differently[59]. Head-on collisions were suggested to lead to stalling of replication forks, whereas co-directional collisions allow the CDC45–MCM–GINS (CMG) replicative helicase to unwind RNA–DNA hybrids. Notably, only head-on collisions activate the ATR pathway, which orchestrates the cellular response to DNA replication stress[59]. Subsequent studies revealed that the activation of ATR by R-loops requires replication fork reversal **[G]** and the structure-specific endonuclease MUS81, suggesting that replication forks are forced to undergo reversal and MUS81-mediated processing when they encounter R-loops[61]. ATR inhibition in cells with high levels of R-loops leads to increased MUS81-dependent DNA damage[61], suggesting that R-loop-induced DNA damage arises from MUS81 cleavage of stalled forks, R-loops, or both.

R-loops could interfere with replication fork progression through a number of mechanisms. The RNA–DNA hybrids and paused Pol II in R-loops or the positive supercoiling induced by R-loops may interfere with the progression of replication forks and cause replication fork collapse. Topoisomerase 1 (TOP1), which relaxes DNA supercoiling, suppresses DNA damage at TTSs where head-on collisions occur[62] (Fig. 3b). The displaced ssDNA in R-loops can potentially form various types of secondary structures. In sequences with high GC skew, the displaced G-rich ssDNA could form G-quadruplexes **[G]** (G4s), which are barriers to replication forks[63]. At telomeres, G4s are a source of replication stress; defects in helicases that resolve G4, such as regulator of telomere elongation helicase 1 (RTEL1), Fanconi anemia group J protein (FANCJ) and Bloom syndrome protein (BLM) increase genomic instability[64-66]. R-loops also induce genomic instability at DNA trinucleotide repeats, possibly by promoting the formation of DNA secondary structures[67,68].

R-loops are associated with gene-repressive histone modifications and chromatin compaction, which could impose replication stress and genomic instability. At TTSs, which are prone to head-on collisions[62], R-loops induce gene-repressive di-methylation of histone H3 Lys9 and recruitment of heterochromatin protein 1 (HP1)[15] (Fig. 3c). R-loops are also associated with high levels of phosphorylated histone H3 Ser10 and with chromatin compaction at centromeres and

pericentromeric regions[69] (Fig. 3c), which are prone to replication stress[33]. Notably, depletion of the linker histone H1, which promotes chromatin compaction, suppresses R-loops and replication-associated genomic instability in *Drosophila melanogaster*, supporting the idea that R-loop-associated chromatin compaction contributes to collisions with replication forks and to genomic instability[70].

[H2] Nucleolytic processing of R-loops

Xeroderma pigmentosum group F (XPF) and XPG, two structure-specific endonucleases involved in transcription-coupled nucleotide excision repair (TC-NER), generate DNA damage in an R-loop-dependent manner[71]. Other TC-NER factors, such as XPB and XPD, are also required for the formation of R-loop-dependent DNA damage[71]. These results suggest that TC-NER proteins have a role in processing R-loops. Camptothecin induces transcription-dependent DSBs through XPF and XPG, suggesting that they cleave co-transcriptional R-loops[72]. XPG is not required for the activation of ATR by R-loops[61], indicating that it acts independently of replication.

The displaced ssDNA in R-loops is a potential substrate of DNA modifying enzymes. During immunoglobulin class-switch recombination in B cells, activation induced cytidine deaminase (AID) converts cytosines to uracil in the ssDNA exposed by R-loops in the switch regions[17]. The uracil generated by AID is processed by the uracil-DNA glycosylase into abasic sites, which directly stall DNA polymerases and also give rise to SSBs through base excision repair, which are converted into DSBs during replication (Fig. 3d). In addition to AID, other cytidine deaminases such as APOBEC3A and APOBEC3B may also attack the ssDNA in R-loops and generate abasic sites and DNA breaks[73].

[H1] R-loop suppression and resolution

To ensure that R-loops are properly regulated and avoid genomic instability, multiple pathways act in concert to prevent aberrant R-loop formation and remove excessive R-loops.

[H2] RNaseH1 and RNaseH2

RNase H proteins are endoribonucleases that specifically cleave the RNA strand in RNA–DNA hybrids. In eukaryotes, this activity is shared between RNaseH1 and RNaseH2 (Fig. 4a, b; Table

1). RNaseH1 removes R-loops in the nucleus and in the mitochondria, where it is essential for mitochondrial DNA replication[74]. RNaseH2 is purely nuclear, consists of three subunits, and interacts with proliferating cell nuclear antigen (PCNA), which may help to recruit it to sites of DNA replication and repair[75]. The RNaseH2A subunit provides the enzymatic activity, and the RNaseH2B subunit is required for PCNA interaction and nuclear localization[75,76]. RNaseH2 can remove R-loops and misincorporated ribonucleotides from DNA, the latter through its action in the RER pathway[39]. RER is initiated by ribonucleotide incision by RNaseH2, followed by DNA repair by the replication machinery, which involves strand displacement DNA synthesis by DNA polymerases δ (Pol δ) and Pol ε, flap cleavage by FEN1 or exonuclease 1 (EXO1), and DNA ligation by DNA ligase I (LIG1)[39]. Although FEN1 provides the dominant RNA–primer removing activity, RNaseH2 is also implicated in RNA primer removal[75]. Work in budding yeast suggests that RNaseH2 is cell cycle regulated and more highly expressed and active in late S and G2 phase, when it is required for RER and R-loop removal[77]. RNaseH1 is active independently of cell cycle stage, but may require R-loop-induced stress for its recruitment to chromatin[77].

[H2] R-loop suppression in transcription

RNA processing suppresses R-loop accumulation at all stages of transcription (Fig. 4a; Table 1). During RNA synthesis, the intrinsic transcript-cleavage activity of Pol II, which requires transcription elongation factor S-II (TFIIS; also known as TCEA1), suppresses R-loop formation by reducing Pol II pausing and backtracking **[G]**[78]. During co-transcriptional RNA processing, splicing factors such as serine/arginine-rich splicing factor 1 (SRSF1)[79], splicing factor proline and glutamine rich (SFPQ)[80] and splicing factor 3B subunit 1 (SF3B1)[81] bind the nascent RNA as it leaves the RNA polymerase active site and prevent it from re-annealing to the template. The transcription export (TREX) complex, which contains the core components of the THO complex and additional RNA binding proteins[82], suppresses R-loop formation by promoting the assembly and nuclear export of the messenger ribonucleoprotein[83]. The TREX-2 complex also suppresses R-loops by promoting mRNA export[84]. Processing of lncRNAs also prevents R-loop formation. Cleavage of lncRNAs mediated by SPT6 and the Integrator complex **[G]** prevents R-loop formation by extended lncRNA transcripts[85]. Degradation of some enhancer RNAs by the RNA exosome **[G]** prevents R-loop accumulation at the corresponding enhancers[86].

TOP1 resolves both positive and negative supercoiling during transcription and DNA replication (Table 1). Loss of TOP1 results in R-loop accumulation in active genes, possibly because the increase of negative supercoiling behind elongating Pol II favors RNA–DNA hybrid formation[87] (Fig. 4a). The TOP1 inhibitor camptothecin induces cotranscriptional R-loops[72]. TOP1 suppresses R-loop-associated DNA damage at TTSs where head-on collisions of R-loops and replication forks occur[62].

The helicase senataxin is also involved in R-loop suppression at TTSs (Fig. 4a; Table 1)[14,88]. The stability of senataxin is regulated by the deubiquitinase by the ubiquitin carboxyl-terminal hydrolase USP11, which also controls R-loop levels[89]. Depletion of the RNA helicase aquarius increases R-loops and genomic instability, suggesting that Aquarius unwinds RNA–DNA hybrids[71]. Several helicases of the DEAD-box (DDX) and DEAH box (DHX) families are involved in R-loop suppression. DDX39B unwinds RNA–DNA hybrids *in vitro* and suppresses co-transcriptional R-loops genome-wide[90]. DDX21 also unwinds RNA–DNA hybrids and it cooperates with sirtuin-7 to suppress R-loops at specific genes[91]. DHX9 unwinds R-loops *in vitro*, associates with R-loops in cells, and suppresses camptothecin-induced cotranscriptional R-loops[92,93]. Paradoxically, DHX9 was shown to also promote R-loop formation in cells with impaired splicing[80]. In fission and budding yeasts, the senataxin homolog Sen1[94] and the helicases Pif1 and Rrm3 (Ref.[95,96]) are required for proper transcription termination at tRNA genes and for maintaining genomic stability, although the functions of these proteins are not linked with R-loops[94,95]. Together, these findings suggest that cells use multiple RNA–DNA helicases to suppress cotranscriptional R-loops and their associated genomic instability (Table 1). Whether and how these helicases function in a context-specific manner and how they are regulated remains largely unknown.

A number of DNA repair proteins are also involved in the suppression of cotranscriptional R-loops (Table 1). Breast cancer type 1 susceptibility protein (BRCA1) associates with TTSs and recruits senataxin to suppress R-loops[88]. In human breast luminal epithelial cells with *BRCA1* mutations, R-loops accumulate at promoter-proximal Pol II pause sites; and suppression of Pol II pausing alleviates R-loop accumulation in mouse BRCA1 deficient cells[97]. Loss of BRCA2 also leads to increased Pol II pausing and R-loop accumulation at promoter-proximal sites[98]. BRCA2 binds the

TREX-2 complex to suppress R-loops[84]. In addition, BRCA2 recruits DDX5 to suppress R-loops in transcribed regions[99]. In BRCA1-deficient or BRCA2-deficient cancer cells, DHX9 is recruited to R-loops by RNF168 to keep R-loops at tolerable levels[10]. Several proteins of the Fanconi anemia repair pathway, such as FANCD2, FANCI and FANCA, are important for R-loop suppression[100,101]. The FANCD2–FANCI complex interacts with RNA–DNA hybrids and recruits RNA-processing factors, such as heterogeneous nuclear ribonucleoprotein U and DDX47, to remove R-loops[102 103]. The ssDNA-binding replication protein A (RPA) is also present at R-loops, possibly on the displaced ssDNA[104]. RPA recruits and stimulates RNaseH1 to prevent excessive R-loop accumulation. Structure-specific nucleases such as XPF, XPG, FEN1 and CtIP (also known as RBBP8) are also involved in the processing and removal of cotranscriptional R-loops[72,105].

[H2] R-loop responses during DNA replication

The suppression of cotranscriptional R-loops is important for preventing collisions of R-loops and replication forks. For example, TOP1 resolves R-loop associated DNA supercoiling at TTSs to suppress the induction of DNA damage by head-on collisions of replication forks and R-loops[62]. A number of DNA repair proteins are involved in the response to collisions of replication forks with R-loops (Table 1). The helicase FANCM unwinds telomeric RNA–DNA hybrids *in vitro* and suppresses replication stress during alternative lengthening of telomeres (ALT)[106,107]. BLM also unwinds RNA–DNA hybrids *in vitro* and suppresses R-loops in cells[108]. BRCA1, FANCM and BLM act in concert to resolve R-loop-induced replication stress at ALT telomeres[109]. The helicase senataxin associates with replication forks and protects forks in transcribed genes[110]. In response to DSB formation in such genes, senataxin is recruited to DSBs to remove RNA–DNA hybrids[45]. DDX1, DDX5, DDX21 and DHX9 interact with ATPase family AAA domain-containing protein 5 (ATAD5), which offloads PCNA replication forks[111]. Depletion of this group of helicases increases R-loops in S phase and reduces DNA synthesis, suggesting that they remove R-loops during replication. DDX19 enters the nucleus following DNA damage in an ATR-dependent manner to remove R-loops[112] (Fig. 4b). Structure-specific nucleases such as MUS81 and MRE11 are also involved in the suppression of R-loops, possibly by cleaving DNA structures induced by collisions of R-loops and replication forks[61,113].

Several proteins involved in DNA damage responses, such as ataxia telangiectasia mutated (ATM), ATR, CHK1 and CHK2, suppress R-loops[114]. These checkpoint proteins were proposed to be important for preventing collisions of replication forks with R-loops or for repairing collapsed forks[114]. Head-on collisions between replication forks and R-loops activate ATR signaling, whereas codirectional collisions activate ATM, suggesting that there could be different replication-restart pathways for head-on versus codirectional collisions[59]. Replication forks undergo reversal at R-loops, which generates substrates for nucleolytic processing by MUS81–EME2 (the MUS81 complex most active in S phase)[61,115] (Fig. 4b). MUS81 processing of reversed forks at R-loops activates ATR[61], allows Pol II to passage through stalled forks, and eventually promotes fork restart using strand annealing by RAD52 and nick ligation by LIG4[115] (Fig. 4c). Such a process is reminiscent of, but distinct from break-induced replication **[G]** (BIR)[115].

There is evidence that replication forks can bypass R-loops, either through RNA–DNA hybrid unwinding by the replicative helicase or by repriming replication downstream of the R-loop (Fig. 4b). As discussed above, codirectional collisions of replication forks and R-loops reduce R-loop levels[59], possibly because the helicase MCM (part of the CMG complex) can unwind RNA–DNA hybrids on the leading strand. PrimPol promotes replication fork progression through R-loops at GAA repeats; R-loops could promote the formation of a secondary structures such as hybrid triplex structure formed between a GAA repeat and the opposite CTT repeat. If the GAA repeat was on the leading strand template such a triplex structure could stall the leading-strand DNA polymerase[116] (Fig. 4b). The repriming by PrimPol on the leading strand allows replication forks to avoid stalling. PrimPol also limits the formation of R-loops in GAA repeats, suggesting that in absence of PrimPol, excessive ssDNA could become accessible for RNA polymerase to produce more RNA–DNA hybrids[116]. As ssDNA generally forms at stalled forks, RNA polymerases could also promote R-loop formation in response to other replication-blocking lesions. Behind the progressing fork, the PCNA offloader ATAD5 can restrict R-loop formation, possibly by preventing PCNA from interfering with transcription, as well as by recruiting RNA–DNA helicases as described above[111].

[H2] Chromatin modulators

In addition to factors that function in transcription, replication and DNA repair machineries, several chromatin reorganizing activities are implicated in preventing R-loop accumulation (Table 1). The SWI/SNF[100,117] and INO80[118] chromatin remodeling complexes and the histone chaperone complex facilitates chromatin transcription (FACT)[119] aid replisome progression in presence of R-loops (Fig. 4b). FACT is proposed to replace histones at sites of transcription–replication conflicts, possibly in response to R-loop-induced formation of repressive histone modifications[69,119]. INO80 is similarly suggested to resolve R-loops by facilitating chromatin decompaction but could also help resolve R-loops through its interaction with RNA–DNA helicases such as DDX5 or by facilitating removal of stalled Pol II[118]. SWI/SNF might suppress R-loops by controlling the accessibility of chromatin to the DNA repair machinery[100]. The mechanism is still unclear, as SWI/SNF increases chromatin accessibility genome-wide, but it limits chromatin accessibility at sites of transcription–replication conflicts[100]. SWI/SNF acts in the Fanconi anemia pathway together with senataxin and FANCD2, and in absence of SWI/SNF R-loops accumulate specifically in S phase. This result suggests that the function of SWI/SNF is specific to replicating cells, such as removing R-loops at stalled replication forks[100]. SWI/SNF is also implicated in recruiting TOP2A to R-loop sites[117].

Several chromatin modifiers including Polycomb repressive complex 1 (PRC1)[120,121], the histone Lys acetyltransferase KAT8[122] and the histone deacetylase (HDAC) SIN3A complex[83] have been implicated in suppressing R-loops (Fig. 4; Table 1). As PRC1 and SIN3A function as transcription repressors, R-loop accumulation in their absence could directly result from increased transcription. However, SIN3A also suppresses R-loop accumulation through a physical interaction with the THO complex subunit THOC1 (Ref.[83]). The acetyllysine readers bromodomain-containing protein 2 (BRD2) and BRD4 suppress R-loop formation, either by preventing Pol II pausing (BRD4) or through recruitment of TOP1 (BRD2)[123-125]. The majority of these factors have known roles in DNA replication and the DNA damage response (reviewed in[126]) (Fig. 4a,b), which promotes R-loop processing during collisions with replication forks. As the cited studies were performed in cycling cells, it is uncertain to what extent the above factors also directly and specifically suppress R-loops during transcription. This question could be further investigated using cells in G1 phase or non-cycling cells.

[H1] Functions of hybrids in DNA repair

Although R-loops can induce genomic instability, R-loops and RNA–DNA hybrids are also induced by DNA damage. The R-loops and RNA–DNA hybrids induced by DNA damage may have several distinct roles in promoting the removal of DNA damage in different contexts.

[H2] Repair of double-strand DNA breaks

Both the RNA generated in transcribed regions before DNA damage and the RNAs induced by DNA damage contribute to the formation of RNA–DNA hybrids at DSBs. The RNA–DNA hybrids at DSBs may promote homologous recombination (HR) through several distinct mechanisms.

[H3] Recruitment of repair proteins. Several types of RNA–DNA hybrids have been proposed to recruit DNA repair proteins to DSBs. DSB-induced small RNAs (diRNAs) or DNA damage response small RNAs (DDRNAs) are small non-coding RNAs generated by Dicer at DSBs[40,41]. diRNAs form a complex with Argonaute-2 and RAD51, which promotes the localization of RAD51 to DSBs, possibly through diRNA hybridization with DNA[127] (Fig. 5a). The DDRNAs are derived from damage-induced long non-coding RNAs (dilncRNAs), which are transcribed at DSBs[128]. dilncRNAs form RNA–DNA hybrids at DSBs, thereby promoting the recruitment of the HR proteins BRCA1, BRCA2 and RAD51[129] (Fig. 5a). BRCA1 was shown to bind RNA–DNA hybrids, and BRCA1 foci in cells in S and G2 phases were reduced by RNase H treatment, suggesting that BRCA1 may be a sensor of R-loops[129]. When ROS are locally induced at a specific chromosomal locus, R-loops are generated in a transcription-dependent manner[46], suggesting that RNA transcripts hybridize with DNA at sites of DNA damage (Fig. 5a). The repair of ROS-induced DSBs requires a non-canonical HR pathway mediated by Cockayne syndrome group B (CSB), a protein also involved in TC-NER, and RAD52 and RAD51[46]. Both CSB and RAD52 directly bind RNA–DNA hybrids *in vitro* and localize to sites of DNA damage in an R-loop-dependent manner in cells[46,130], suggesting that they are also sensors of DNA damage-induced R-loops. In G2 cells exposed to ionizing radiation or laser microirradiation, RAD52 is recruited to DNA damage sites in a transcription and RNA–DNA hybrid-dependent manner[131]. Furthermore, RPA may act as a sensor of R-loops by recognizing the displaced ssDNA[104]. Collectively, these findings suggest that R-loops or RNA–DNA hybrids at DSBs may directly or indirectly promote

the recruitment of a number of DNA repair proteins, which may enhance DSB repair by increasing the local concentrations of these repair factors.

DNA end resection **[G]** is important for the choice of DNA repair pathways. Whereas resection is required for HR, it antagonizes canonical non-homologous end joining. Whether and how the R-loops and RNA–DNA hybrids at DSBs affect the resection of DNA ends is still controversial. In yeast, overexpression of RNaseH1 destabilizes RNA–DNA hybrids at DSBs and induces excessive resection[132], suggesting that hybrids restrict resection. In human cells, however, RAD52, which localizes at DSBs in a hybrid-dependent manner, recruits the XPG to cleave R-loops and promotes ssDNA formation (Fig. 5b)[131]. Loss of Pol III, which generates RNA–DNA hybrids at DSBs, also reduces resection (Fig. 5b)[44]. Furthermore, DHX9 recruits BRCA1 to Pol II-generated RNA at DSBs to enhance resection[133]. These findings suggest that hybrids promote resection in human cells. Nonetheless, depletion of senataxin did not significantly affect resection[45]. Notably, reducing hybrids at DSBs by RNaseH1 overexpression in yeast or depletion of Pol III in human cells results in sequence loss around DSBs[44,132], suggesting that hybrids protect ssDNA overhangs.

[H3] RNA-templated or bridged repair. In budding yeast, RNA transcripts can serve as template for DSB repair when RNaseH and reverse transcriptase activities are absent[130]. The RNA-templated repair of DSBs requires Rad52, which can promote the formation of RNA–DNA hybrids and strand exchange between homologous ssRNA and ssDNA[130,134]. RNA transcripts were hypothesized to hybridize with 3' overhangs at DSBs, thereby allowing RNA-templated DNA synthesis (Fig. 5c). The resulting extended ssDNA overhang then captures the ssDNA from the opposite DSB end, joining the two ends together. In human cells, the repair of ROS-induced DSBs, which is dependent on RNA–DNA hybrids, can occur in G0 and in G1 (before DNA replication), suggesting that it may also involve RNA templates[135]. Using purified proteins and RNA, RAD52 was shown to promote RNA-templated DNA recombination *in vitro*[136]. A recent study shows that DNA polymerase θ, which is required for a DSB repair pathway termed theta-mediated DNA end joining can efficiently synthesize DNA using RNA as a template and promote RNA-templated DSB repair in a cell-based reporter assay[137]. Although RNA–DNA hybrids can enable RNA-templated repair of DSBs, whether this type of repair counts for a significant portion of repair activity in cells remains unknown. In addition, RNA transcripts could potentially function as a

molecular bridge to hold two DNA ends together. In the presence of RAD52, an RNA oligo that can anneal with ssDNA on both sides of a DSB enables the joining of DNA ends *in vitro*[136]. Whether RNA can function as a bridge during DSB repair in cells and how important this mechanism is *in vivo* remains unclear.

[H3] DR-loops. Using HR reporter assays, in which local transcription at DSB sites can be switched on and off, local transcription was shown to substantially stimulate HR[138]. Importantly, even in the absence of local transcription, tethering homologous RNA transcripts to the vicinity of DSBs by nuclease-inactive Cas9 enhances HR, recapitulating the effects of transcription[138]. These findings suggest that preexisting RNA transcripts at transcriptionally active sites have an important role in HR once DSBs emerge. Furthermore, the stimulation of HR by transcription or tethered RNA transcripts is dependent on RAD51 associated protein 1 (RAD51AP1) and USP1 associated factor 1 (UAF1), which form a complex that promotes R-loop formation[138,139]. Surprisingly, DSB-induced and RAD51AP1-generated RNA–DNA hybrids are detected not only at DSBs but also in the DNA serving as the donor DNA **[G]** for repair, suggesting that DNA–DNA and RNA–DNA hybrids co-exist in donor DNA. Indeed, purified RAD51AP1 and RAD51 can promote the invasion of both ssRNA and ssDNA into donor dsDNA *in vitro*, generating DR-loops, which contain both DNA–DNA and RNA–DNA hybrids (Fig. 5d). The formation of DR-loops is associated with enhanced RAD51-mediated ssDNA invasion, suggesting that DR-loops are HR intermediates favoring RAD51 function. These findings on DR-loops suggest that RNA is not only a regulator of DNA repair proteins, but also a direct participant in recombination. This role of RNA in HR may be crucial for protecting transcribed regions of the genome.

[H3] RNA modifications. The N^6-methyladenosine (m^6A) modification of mRNA is associated with R-loops and promotes R-loop suppression and resolution[140,141], although it may also facilitate transcription termination through R-loops[142]. In response to ultraviolet irradiation damage, m^6A is rapidly and transiently generated by methyltransferase 3, N^6-adenosine-methyltransferase complex catalytic subunit (METTL3) at DNA damage sites[143]. In the absence of m^6A, the translesion synthesis **[G]** DNA polymerase Polκ cannot be recruited efficiently to DNA, and ultraviolet damage repair is impaired[143]. METTL3 catalyzes the m^6A modification of RNA at DSBs, to which

the m^6A 'reader' YTHDC1 is recruited to stabilize RNA–DNA hybrids[144]. Loss of METTL3 leads to defective HR[144].

In response to ROS-induced DNA damage, the 5-methylcytosine (m^5C) modification of RNA is rapidly induced at DNA damage sites[145] (Fig. 5a). This induction requires local transcription, suggesting that local mRNA is modified by m^5C[145]. RNA–DNA hybrids are directly bound by tRNA aspartic acid methyltransferase 1 (TRDMT1; also known as tRNA (cytosine(38)-C(5))-methyltransferase), which is required for efficient m^5C formation at DNA damage sites and repair of ROS-induced DSBs. Both RAD52 and RAD51 are recruited to sites of ROS-induced DNA damage in a TRDMT1-dependent manner[145] (Fig. 5a). RAD52, which is required for recruiting RAD51 to ROS-induced DSBs, has higher affinity for RNA–DNA hybrids containing m^5C-modified RNA. These results suggest that a TRDMT1–m^5C–RAD52 axis operates on R-loops to promote the repair of ROS-induced DSBs. In addition to RNA methylation, adenosine-to-inosine RNA editing **[G]** is increased near DSBs[146]. This response is mediated by adenosine deaminases acting on RNA 2 (ADAR2; also known as double-stranded RNA-specific editase 1), which is required for efficient HR and cellular resistance to DNA damage.

[H2] Repair of collapsed replication forks

As discussed above, BIR or related pathways are involved in restarting replication forks stalled by R-loops. At sites of DNA damage or genomic instability, DNA damage-induced R-loops may stall replication forks and induce BIR, allowing BIR to alleviate the replication stress. For example, in yeast, TERRA-generated R-loops accumulate at critically short telomeres, promoting the DDR to extend telomeres[147]. In human cancer cells, telomeric R-loops promote the BIR-related ALT pathway to extend telomeres[148,149]. Local induction of ROS at telomeres triggers an R-loop-dependent CSB–RAD52–DNA polymerase delta subunit 3 (POLD3) pathway to remove DNA damage at telomeres, thereby linking R-loops with POLD3-mediated BIR[150]. A number of proteins involved in these BIR (and related) pathways, such as RAD52, RAD51AP1 and CSB, can directly bind RNA–DNA hybrids. Although damage-induced R-loops may induce BIR by stalling replication forks, whether and how R-loops directly participate in the process of BIR is still unknown.

[H2] Repair of single-strand DNA breaks

A complex of BRCA1 with RNAi machinery was recently shown to generate single-stranded DNA-damage-associated small RNAs (sdRNAs) 24–40 nucleotides-long from RNA transcripts in response to DNA damage[151] (Fig. 5e). These sdRNAs interact with partner and localizer of BRCA2 (PALB2) and RAD52 to promote the repair of SSBs at the TTSs of active genes. R-loops were speculated to be involved in the formation and function of sdRNAs. Consistent with this hypothesis, sdRNAs function at the loci from which they are generated, and R-loops are detected at these loci. Further studies are needed to determine whether R-loops directly participate in sdRNA generation and function.

[H2] Resolution of hybrids during DNA repair

Although RNA–DNA hybrids have positive roles in DNA repair, they must be removed before the completion of repair. In yeast, RNaseH is required for suppressing hybrids at DSBs and for efficient HR[132,152]. However, other studies argue that RNaseH is not required for DSB repair[153]. In human cells, BRCA2 recruits DDX5 and RNaseH2 to DSBs to remove RNA–DNA hybrids[129,154]. Furthermore, senataxin promotes RAD51 recruitment to DSBs and reduces illegitimate rejoining of distant DSBs, supporting the idea that hybrid removal is important for accurate DSB repair[45].

[H1] Pathological R-loops

Aberrant accumulation of R-loops can cause increased replications stress and/or DNA damage, which can underlie cancer susceptibility or neurodegeneration. R-loop levels could be elevated in both types of disorders owing to loss of transcription regulation or R-loop suppressing activities. The genomic location of R-loop accumulation might vary depending on the nature of the genetic alterations[20], and on whether the affected tissue is proliferating or post-mitotic.

[H2] Cancer-associated R-loops

A number of oncogenic events can increase the levels of R-loops and associated genomic instabiltiy. Activated oncogenes and bacterial carcinogens can increase R-loop formation, which elevates replication stress and genomic instability (Fig. 6a). Oncogenic estrogen signaling[18], the $HRAS^{V12}$ oncoprotein[19] and the EWS-FLI1 fusion oncoprotein[9], all increase R-loop levels genome-wide, promoting replication stress and DNA damage. HRAS-induced and estrogen-

induced R-loop formation is linked with increased transcription, especially at the target genes of estrogen and RAS signaling[18,19]. EWS-FLI1, a fusion of EWS RNA binding protein 1 (EWSR1) and Friend leukemia integration 1 transcription factor (FLI1), is also proposed to increase transcription by stimulating Pol II activity; depletion of wild-type EWSR1 to mimic the effects of EWS-FLI1 fusion induced R-loops especially in highly expressed regions[9] (Fig. 6a). Expression of the SS18-SSX1 oncoprotein, a fusion of the SS18 (also known as SSXT) subunit of BAF chromatin remodeling complex (mammalian SWI/SNF) with SSX1, which drives synovial sarcoma, also increased R-loop levels[11]. Although the underlying mechanism is not known, it is of note that SS18-SSX alters SWI/SNF targeting to chromatin[155]. *Heliobacter pylori* infections, which act as a carcinogen promoting gastric cancer, also increase RNA synthesis by activating NF-κB signaling and thereby inducing R-loop-dependent replication stress[156].

However, elevated RNA synthesis is not the only way in which oncogenes can increase R-loop levels. Suspected oncogenic mutations in the genes encoding the splicing factors SRSF2 and U2 small RNA auxiliary factor 1 (U2AF1; also known as U2AF35), which are prevalent in myelodysplastic syndrome (MDS)[20,104,157], were found to induce R-loops and replication stress, possibly through increasing Pol II pausing or other mechanisms[20] (Fig. 6b). Finally, the oncogenes MYCN[158] and MDM2 (Ref.[121]) have been reported to suppress R-loop formation, acting in concert with BRCA1[158] or the chromatin modifier PRC1 (Ref.[121]), respectively. The R-loop suppressing effects of MYCN and MDM2 may allow cancer cells to tolerate high levels of transcription and replication stress. The discrepancy between MYCN and other oncogenes such as $HRAS^{V12}$ might be explained by timing, as MYCN suppresses R-loops in the first few hours following oncogene induction[158], whereas $HRAS^{V12}$-induced R-loops were detected after three days of induction[19]. It will be interesting to investigate the effects of prolonged MYCN overexpression on R-loop levels.

Factors involved in the repair of replication-associated DNA damage such as FANCM, the MRN complex[113,159], BRCA1[88] and BRCA2[84], and factors required for activation of cell cycle checkpoints such as ATM[160], function as tumor suppressors. The R-loop suppressing functions of these proteins[84,88,113,159,160] could be important in preventing conflicts between R-loops and replication, thereby limiting cancer-promoting genomic instability in proliferating cells (Fig. 6b).

The regulation of TERRA and telomeric R-loops is altered in a subset of human cancers. TERRA is generally upregulated in cancer cells that use the ALT pathway to extend telomeres[161]. Loss of ATRX, which is a chromatin remodeller acting at subtelomeres and frequently mutated in ALT-positive cancers, increases TERRA levels[162] (Fig. 6c). Whether ATRX loss is the primary cause of TERRA upregulation in ALT-positive cancers remains unclear[161]. Consistent with the upregulation of TERRA, telomeric R-loops are more abundant in ALT-positive cancer cells[27]. Loss of ATRX also increases formation of G-quadruplexes at telomeres[163,164], which may indirectly contribute to R-loop formation. A recent study showed that inhibition of TERRA transcription diminishes ALT activity[148], supporting the idea that telomeric R-loops drive ALT by inducing replication stress[27,148,165,166] (Fig. 1b). The helicase FANCM has the ability to unwind telomeric R-loops, and it restricts ALT activity to a tolerable level in ALT-positive cancer cells[106,107] (Fig. 6d). BRCA1 directly binds TERRA and suppresses telomeric R-loops and DNA damage[167] (Fig. 6d).

Cancer cells can contain variants in the telomeric repeat sequence, and A–C mismatches between DNA and TERRA RNA arising from this variation inhibit RNaseH2 activity. RNA editing by the adenosine-to-inosine editing enzyme ADAR1 (also known as double-stranded RNA-specific adenosine deaminase) converts A–C mismatches to I–C matched base pairs, thereby allowing RNA processing by RNaseH2 specifically at telomeres, which is required for the proliferation of non-ALT or telomerase-activated cancer cells[168]. Interestingly, depletion of the helicase RTEL1, which binds the G-quadruplexes formed by TERRA, reduces telomeric R-loops, but increases telomere instability[169]. These results suggest that telomeric R-loops need to be maintained at an appropriate level to keep telomeres stable or to not activate ALT above a tolerable level.

[H2] R-loops in the central nervous system

Excessive R-loop formation in post-mitotic neurons could promote transcription-associated DNA damage and genomic instability independently of conflicts with DNA replication, thereby contributing to neurodegenerative diseases and neurodevelopmental disorders. Mutations in DNA repair genes such as *ATM* and *SETX* underlie cerebellar ataxias **[G]** and are linked to increased DNA damage in neurons[170,171]. At least one of the affected genes, *SETX*, encodes senataxin, a protein with direct roles in R-loop resolution[14]. Increased R-loop formation could plausibly

contribute to DNA damage in these disorders[170]. Mouse models of cerebellar ataxias, including ataxia telangiectasia (caused by mutations in *ATM*), ataxia oculomotor apraxia type 1 (AOA1; mutations in aprataxin), AOA2 (mutations in senataxin), and spinocerebellar ataxia with axonal neuropathy (mutations in tyrosyl-DNA phosphodiesterase) all display increased levels of R-loops in proliferating tissues[160] (Fig. 6d). However, in line with the observation that mice tend to be poor models for neurodegeneration that frequently fail to develop the neurological symptoms of the human disorders, R-loops are not detected in the brains of these mice[160]. By contrast, human non-proliferating AOA2-derived neural progenitor cells do display increased R-loop signals in nucleoli and the cytoplasm, as well as increased oxidative DNA damage[170]. Other *SETX* mutations underlie amyotrophic lateral sclerosis 4 (ALS4), in which R-loops are reduced in gene promoters, which reduces the expression of affected genes and may contribute to disease etiology[172].

Several neurodegenerative diseases such as amyotrophic lateral sclerosis and frontotemporal dementia (ALS/FTD) and Friedreich ataxia are linked with expansions of short guanine-containing repeat sequences, which are prone to forming stable R-loops[173,174]. Several mechanisms have been proposed to explain how increased R-loop formation in one gene might promote neurodegeneration. R-loops at expanded repeats in *C9orf72*, which underlies ALS, and in the frataxin gene, which underlies Friedreich ataxia, have been linked to impaired Pol II transcription and gene silencing, again suggesting that R-loops could promote pathology by posing blocks to gene expression[173,174] (Fig. 6d). In addition, RNA binding proteins such as TDP-43 form dysfunctional aggregates in ALS and FTD, which can further promote R-loop accumulation[175]. Repeat expansions in *C9orf72* are linked also with P62 accumulation, causing defective ATM signaling, which exacerbates R-loop-mediated DNA damage[176]. Mutations in splicing factors can also underlie neurodegenerative disorders. R-loops accumulate specifically in the embryonic neurons of zebrafish lacking the splicing factor SF3B1, where they promote DNA damage and apoptosis[81].

R-loop formation in expanded repeats is implicated also in the neurodevelopmental disorder Fragile X syndrome[173]. Mutations in RNaseH2 or in the deoxynucleoside triphosphate triphosphohydrolase SAMHD1, which are factors involved in R-loop prevention[177], cause Aicardi-Goutieres syndrome, a neuro-inflammatory disorder that has been linked to the presence of DNA

in the cytosol and thus to activation of the cGAS–STING pathway **[G]** rather than to R-loop accumulation[178,179].

[H1] Targeting aberrant R-loops in cancer

Although aberrant R-loop accumulation in cancer cells may promote tumorigenesis by altering gene expression and fueling genomic instability, it also confers a vulnerability to cancer cells that can be exploited therapeutically. Recent studies suggest that cancer cells harboring high levels of R-loops can be preferentially killed by drugs that exacerbate R-loop-associated DNA damage and block the R-loop response.

[H2] Inhibition of the ATR checkpoint

An R-loop dependent ATR response is detected in cells depleted of SRSF1, suggesting that ATR is activated in response to aberrant R-loop accumulation[61]. In contrast to the activation of ATR by the replication inhibitor hydroxyurea, ATR activation by R-loops required replication-fork reversal and the endonuclease MUS81–EME2[61]. Inhibition of ATR in cells with high levels of R-loops leads to an increase of MUS81-EME1–generated DNA damage, suggesting that ATR functions in a feedback loop to suppress R-loop-associated DNA damage by limiting MUS81 activity[61]. ATR also suppresses transcription–replication collisions and promotes fork recovery in cells with high levels of R-loops[61]. Together, these results suggest that ATR is a key regulator of the R-loop response.

Several splicing-factor mutations prevalent in MDS and acute myeloid leukemia (AML), such as $U2AF1^{S34F}$ and $SRSF2^{P95H}$, were shown to increase R-loops in cells[20,104,157]. An ATR response was induced by these spliceosome mutants in an R-loop-dependent manner. Furthermore, inhibition of ATR in cells expressing the spliceosome mutants increases R-loop-dependent DNA damage, suggesting that ATR protects the cells against aberrant R-loops. Thus, inhibition of ATR is an attractive strategy to exploit the R-loop-associated vulnerability of cancer cells (Fig. 7a). How R-loops are induced by spliceosome mutants and how R-loops contribute to the ATR inhibitor (ATRi) effects is still not fully understood.

[H2] Inhibition of RNA splicing

The splicing-factor mutations found in MDS affect alternative splicing of many genes. RNA splicing inhibitors were developed to kill MDS cells by exacerbating splicing defects. In particular, inhibitors of SF3B1, such as pladienolide B, E7107, and H3B-8800, have been shown to preferentially kill cells expressing splicing-factor mutants[180]. Splicing inhibitors were also found to increase R-loops and stimulate the ATR response in cells expressing U2AF1^{S34F} (Ref. [157]) (Fig. 7b). Consequently, splicing inhibitors increase the sensitivity of U2AF1^{S34F}-expressing cells to ATRi[157]. Thus, combinations of ATRi and drugs that increase R-loop levels may be effective in treating MDS.

[H2] *Nonsense-mediated mRNA decay inhibition*

Cells expressing MDS-associated splicing-factor mutants, including U2AF1^{S34F} and SF3B1^{K700E}, were recently shown to be sensitive to an inhibitor of serine/threonine-protein kinase SMG1 (SMG1), which is required for nonsense-mediated mRNA decay **[G]** (NMD)[181]. SMG1 inhibition or depletion of the NMD factor UPF1 increase R-loop levels and induce DNA damage in an R-loop-dependent manner, suggesting that inhibition of NMD preferentially kills MDS cells by exacerbating R-loop-associated genomic instability (Fig. 7b).

[H2] R-loops-associated DNA repair defects

As described above, EWS-FLI expression increases Pol II elongation and R-loop formation, and activates the ATR pathway[9]. Furthermore, EWS-FLI reduces HR efficiency[9]. In Ewing's sarcoma cells, BRCA1 associates with elongating Pol II and thus fails to localize at DSBs, conferring HR deficiency[9]. Because ATR is crucial for both HR and the R-loop response and as poly(ADP-ribose) polymerase (PARP) inhibitors selectively kill HR-deficient cells, the HR deficiency of Ewing's sarcoma cells renders them sensitive to inhibitors of ATR and PARPs (Fig. 7c).

[H2] Stabilization of G-quadruplexes

When R-loops are formed in sequences with high GC skew, the displaced G-rich ssDNA can give rise to G4s. Stabilization of G4s by G4 ligands renders R-loops more persistent, thereby enhancing their impacts on genomic stability. Furthermore, G4s themselves can interfere with DNA replication and cause DNA damage. In BRCA2-deficient cells, G4 ligands induce DNA damage in an R-loop-dependent manner[182] (Fig. 7d). G4-stabilizing ligands also interfere with the

resolution of R-loops at sites of ROS-induced damage in transcribed regions, thereby compromising DSB repair[108]. Telomeric R-loops are associated with G4s, and in ALT$^+$ cancer cells, G4-stabilizing ligands enhance ALT activity and its associated telomere instability[183]. These findings suggest that G4-stabilizing drugs can be used to exploit R-loop associated genomic instability in cancer cells.

[H2] Other R-loop-augmenting drugs

Though not originally designed to induce R-loops, several targeted cancer treatments such as TOP1 inhibitors (camptothecin), HDAC inhibitors (romidepsin) and BET protein inhibitors (JQ1) increase R-loop formation, presumably by inhibiting the R-loop suppressing activities of their targets[123-125,173,184]. There are also reports that tumor-treating electric fields induce R-loop accumulation, possibly as a by-product of inhibiting DNA repair or inducing replication stress[185], and the natural compound anti-cancer drug trabectedin and its derivative lurbinectedin promote R-loop-partially dependent DNA damage[186]. Such findings may have implications for combination therapy with other R-loop-targeting treatments.

[H1] Conclusions and future perspective

Our understanding of the sources, resolution and impact of R-loops and other RNA–DNA hybrids has extended considerably in the past few years owing to the development of new analytic tools and technologies, which allowed us to map R-loops genome-wide, to identify R-loop associated proteins, and to follow the effects of R-loops in various biological contexts. It has become clear that R-loops are not a homogenous structure, but a population of distinct and dynamic RNA–DNA–protein assemblies. The effects of R-loops on the genome are also diverse, depending on where, when and how R-loops are generated, how abundant and stable the R-loops are, and which proteins the R-loops interact with. Although numerous regulators of R-loops have been identified, we still know little about how these regulators function in concert in cells. Why do cells need so many R-loop regulators? Do these regulators function in a context-specific manner? If so, what controls the context specificities of R-loop regulators?

Going forward, it will be crucial to investigate how different types of R-loops and RNA–DNA hybrids are generated, processed and resolved, and how they function in specific chromosomal,

cellular and tissue contexts. This effort should include more detailed investigations of R-loop-mediated genomic instability in cycling versus non-cycling cells and of the relative contribution of DNA replication to R-loop-mediated genomic instability. Finally, our understanding of the roles of R-loops in pathological and therapeutic contexts is still far from complete. It will be important to establish clinically useful assays to measure the accumulation of R-loops in patients, to understand how R-loops and associated genomic instability contribute to pathogenesis and to develop new strategies to alleviate or exploit R-loop-associated defects in therapy.

References:

1 Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* **73**, 398-411, doi:10.1016/j.molcel.2019.01.024 (2019).

2 Garcia-Muse, T. & Aguilera, A. R Loops: From Physiological to Pathological Roles. *Cell* **179**, 604-618, doi:10.1016/j.cell.2019.08.055 (2019).

3 Niehrs, C. & Luke, B. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat Rev Mol Cell Biol* **21**, 167-178, doi:10.1038/s41580-019-0206-3 (2020).

4 Marnef, A. & Legube, G. R-loops as Janus-faced modulators of DNA repair. *Nat Cell Biol* **23**, 305-313, doi:10.1038/s41556-021-00663-4 (2021).

5 Stodola, J. L. & Burgers, P. M. Mechanism of Lagging-Strand DNA Replication in Eukaryotes. *Adv Exp Med Biol* **1042**, 117-133, doi:10.1007/978-981-10-6955-0_6 (2017).

6 Sanz, L. A. *et al.* Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol Cell* **63**, 167-178, doi:10.1016/j.molcel.2016.05.032 (2016).

7 Chen, J. Y., Zhang, X., Fu, X. D. & Chen, L. R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. *Nat Protoc* **14**, 1661-1685, doi:10.1038/s41596-019-0154-6 (2019).

8 Chiang, H. C. *et al.* BRCA1-associated R-loop affects transcription and differentiation in breast luminal epithelial cells. *Nucleic Acids Res* **47**, 5086-5099, doi:10.1093/nar/gkz262 (2019).

9 Gorthi, A. *et al.* EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature* **555**, 387-391, doi:10.1038/nature25748 (2018).

10 Patel, P. S. *et al.* RNF168 regulates R-loop resolution and genomic stability in BRCA1/2-deficient tumors. *J Clin Invest* **131**, doi:10.1172/JCI140105 (2021).

11 Jones, S. E. *et al.* ATR Is a Therapeutic Target in Synovial Sarcoma. *Cancer Res* **77**, 7014-7026, doi:10.1158/0008-5472.CAN-17-2056 (2017).

12 Ujvari, A. & Luse, D. S. RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat Struct Mol Biol* **13**, 49-54, doi:10.1038/nsmb1026 (2006).

13 Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chedin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* **45**, 814-825, doi:10.1016/j.molcel.2012.01.017 (2012).

14 Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* **42**, 794-805, doi:10.1016/j.molcel.2011.04.026 (2011).
15 Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N. J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436-439, doi:10.1038/nature13787 (2014).
16 Boque-Sastre, R. *et al.* Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc Natl Acad Sci U S A* **112**, 5785-5790, doi:10.1073/pnas.1421197112 (2015).
17 Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E. & Lieber, M. R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* **4**, 442-451, doi:10.1038/ni919 (2003).
18 Stork, C. T. *et al.* Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *Elife* **5**, e17548 (2016).
19 Kotsantis, P. *et al.* Increased global transcription activity as a mechanism of replication stress in cancer. *Nat Commun* **7**, 13087, doi:10.1038/ncomms13087 (2016).
20 Chen, L. *et al.* The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol Cell* **69**, 412-425 e416, doi:10.1016/j.molcel.2017.12.029 (2018).
21 El Hage, A., French, S. L., Beyer, A. L. & Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* **24**, 1546-1558, doi:10.1101/gad.573310 (2010).
22 Abraham, K. J. *et al.* Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature* **585**, 298-302, doi:10.1038/s41586-020-2497-0 (2020).
23 El Hage, A., Webb, S., Kerr, A. & Tollervey, D. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet* **10**, e1004716, doi:10.1371/journal.pgen.1004716 (2014).
24 Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798-801, doi:10.1126/science.1147182 (2007).
25 Schoeftner, S. & Blasco, M. A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* **10**, 228-236, doi:10.1038/ncb1685 (2008).
26 Luke, B. & Lingner, J. TERRA: telomeric repeat-containing RNA. *EMBO J* **28**, 2503-2510, doi:10.1038/emboj.2009.166 (2009).
27 Arora, R. *et al.* RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* **5**, 5220, doi:10.1038/ncomms6220 (2014).
28 Feretzaki, M. *et al.* RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature* **587**, 303-308, doi:10.1038/s41586-020-2815-6 (2020).
29 Talbert, P. B. & Henikoff, S. Transcribing Centromeres: Noncoding RNAs and Kinetochore Assembly. *Trends Genet* **34**, 587-599, doi:10.1016/j.tig.2018.05.001 (2018).
30 Bobkov, G. O. M., Gilbert, N. & Heun, P. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J Cell Biol* **217**, 1957-1972, doi:10.1083/jcb.201611087 (2018).

31 McNulty, S. M., Sullivan, L. L. & Sullivan, B. A. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev Cell* **42**, 226-240 e226, doi:10.1016/j.devcel.2017.07.001 (2017).
32 Kabeche, L., Nguyen, H. D., Buisson, R. & Zou, L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science* **359**, 108-114, doi:10.1126/science.aan6490 (2018).
33 Giunta, S. *et al.* CENP-A chromatin prevents replication stress at centromeres to avoid structural aneuploidy. *Proc Natl Acad Sci U S A* **118**, e2015634118 (2021).
34 Smith, D. J. & Whitehouse, I. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**, 434-438, doi:10.1038/nature10895 (2012).
35 Gloor, J. W., Balakrishnan, L., Campbell, J. L. & Bambara, R. A. Biochemical analyses indicate that binding and cleavage specificities define the ordered processing of human Okazaki fragments by Dna2 and FEN1. *Nucleic Acids Res* **40**, 6774-6786, doi:10.1093/nar/gks388 (2012).
36 Bianchi, J. *et al.* PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol Cell* **52**, 566-573, doi:10.1016/j.molcel.2013.10.035 (2013).
37 Garcia-Gomez, S. *et al.* PrimPol, an archaic primase/polymerase operating in human cells. *Mol Cell* **52**, 541-553, doi:10.1016/j.molcel.2013.09.025 (2013).
38 Reijns, M. A. *et al.* Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**, 1008-1022, doi:10.1016/j.cell.2012.04.011 (2012).
39 Sparks, J. L. *et al.* RNase H2-initiated ribonucleotide excision repair. *Mol Cell* **47**, 980-986, doi:10.1016/j.molcel.2012.06.035 (2012).
40 Wei, W. *et al.* A role for small RNAs in DNA double-strand break repair. *Cell* **149**, 101-112, doi:10.1016/j.cell.2012.03.002 (2012).
41 Francia, S. *et al.* Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature* **488**, 231-235, doi:10.1038/nature11179 (2012).
42 Pessina, F. *et al.* Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nat Cell Biol* **21**, 1286-1299, doi:10.1038/s41556-019-0392-4 (2019).
43 Sharma, S. *et al.* MRE11-RAD50-NBS1 Complex Is Sufficient to Promote Transcription by RNA Polymerase II at Double-Strand Breaks by Melting DNA Ends. *Cell Rep* **34**, 108565, doi:10.1016/j.celrep.2020.108565 (2021).
44 Liu, S. *et al.* RNA polymerase III is required for the repair of DNA double-strand breaks by homologous recombination. *Cell* **184**, 1314-1329 e1310, doi:10.1016/j.cell.2021.01.048 (2021).
45 Cohen, S. *et al.* Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nat Commun* **9**, 533, doi:10.1038/s41467-018-02894-w (2018).
46 Teng, Y. *et al.* ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB. *Nat Commun* **9**, 4115, doi:10.1038/s41467-018-06586-3 (2018).
47 Holt, I. J. The mitochondrial R-loop. *Nucleic Acids Res* **47**, 5480-5489, doi:10.1093/nar/gkz277 (2019).

48 Falkenberg, M. Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays Biochem* **62**, 287-296, doi:10.1042/EBC20170100 (2018).
49 Lima, W. F. *et al.* Viable RNaseH1 knockout mice show RNaseH1 is essential for R loop processing, mitochondrial and liver function. *Nucleic Acids Res* **44**, 5299-5312, doi:10.1093/nar/gkw350 (2016).
50 Silva, S., Camino, L. P. & Aguilera, A. Human mitochondrial degradosome prevents harmful mitochondrial R loops and mitochondrial genome instability. *Proc Natl Acad Sci U S A* **115**, 11024-11029, doi:10.1073/pnas.1807258115 (2018).
51 Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278, doi:10.1016/j.cell.2014.05.010 (2014).
52 Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, doi:10.1126/science.1258096 (2014).
53 Jiang, F. *et al.* Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867-871, doi:10.1126/science.aad8282 (2016).
54 Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149-157, doi:10.1038/s41586-019-1711-4 (2019).
55 Gaudelli, N. M. *et al.* Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471, doi:10.1038/nature24644 (2017).
56 Xiao, Y. *et al.* Structure Basis for Directional R-loop Formation and Substrate Handover Mechanisms in Type I CRISPR-Cas System. *Cell* **170**, 48-60 e11, doi:10.1016/j.cell.2017.06.012 (2017).
57 Zhang, B. *et al.* Mechanistic insights into the R-loop formation and cleavage in CRISPR-Cas12i1. *Nat Commun* **12**, 3476, doi:10.1038/s41467-021-23876-5 (2021).
58 Gan, W. *et al.* R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* **25**, 2041-2056, doi:10.1101/gad.17010011 (2011).
59 Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* **170**, 774-786 e719, doi:10.1016/j.cell.2017.07.043 (2017).
60 Lang, K. S. *et al.* Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell* **170**, 787-799 e718, doi:10.1016/j.cell.2017.07.044 (2017).
61 Matos, D. A. *et al.* ATR Protects the Genome against R Loops through a MUS81-Triggered Feedback Loop. *Mol Cell* **77**, 514-527 e514, doi:10.1016/j.molcel.2019.10.010 (2020).
62 Promonet, A. *et al.* Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nat Commun* **11**, 3940, doi:10.1038/s41467-020-17858-2 (2020).
63 Lee, W. T. C. *et al.* Single-molecule imaging reveals replication fork coupled formation of G-quadruplex structures hinders local replication stress signaling. *Nat Commun* **12**, 2525, doi:10.1038/s41467-021-22830-9 (2021).
64 Kotsantis, P. *et al.* RTEL1 Regulates G4/R-Loops to Avert Replication-Transcription Collisions. *Cell Rep* **33**, 108546, doi:10.1016/j.celrep.2020.108546 (2020).
65 Drosopoulos, W. C., Kosiyatrakul, S. T. & Schildkraut, C. L. BLM helicase facilitates telomere replication during leading strand synthesis of telomeres. *J Cell Biol* **210**, 191-208, doi:10.1083/jcb.201410061 (2015).
66 Sarkies, P. *et al.* FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic Acids Res* **40**, 1485-1498, doi:10.1093/nar/gkr868 (2012).

67 Neil, A. J., Liang, M. U., Khristich, A. N., Shah, K. A. & Mirkin, S. M. RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)n repeats via break-induced replication. *Nucleic Acids Res* **46**, 3487-3497, doi:10.1093/nar/gky099 (2018).
68 Loomis, E. W., Sanz, L. A., Chedin, F. & Hagerman, P. J. Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet* **10**, e1004294, doi:10.1371/journal.pgen.1004294 (2014).
69 Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol Cell* **52**, 583-590, doi:10.1016/j.molcel.2013.10.006 (2013).
70 Bayona-Feliu, A., Casas-Lamesa, A., Reina, O., Bernues, J. & Azorin, F. Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. *Nat Commun* **8**, 283, doi:10.1038/s41467-017-00338-5 (2017).
71 Sollier, J. *et al.* Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell* **56**, 777-785, doi:10.1016/j.molcel.2014.10.020 (2014).
72 Cristini, A. *et al.* Dual Processing of R-Loops and Topoisomerase I Induces Transcription-Dependent DNA Double-Strand Breaks. *Cell Rep* **28**, 3167-3181 e3166, doi:10.1016/j.celrep.2019.08.041 (2019).
73 Buisson, R., Lawrence, M. S., Benes, C. H. & Zou, L. APOBEC3A and APOBEC3B Activities Render Cancer Cells Susceptible to ATR Inhibition. *Cancer Res* **77**, 4567-4578, doi:10.1158/0008-5472.CAN-16-3389 (2017).
74 Cerritelli, S. M. *et al.* Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* **11**, 807-815, doi:10.1016/s1097-2765(03)00088-1 (2003).
75 Bubeck, D. *et al.* PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Res* **39**, 3652-3666, doi:10.1093/nar/gkq980 (2011).
76 Kind, B. *et al.* Altered spatio-temporal dynamics of RNase H2 complex assembly at replication and repair sites in Aicardi-Goutieres syndrome. *Hum Mol Genet* **23**, 5950-5960, doi:10.1093/hmg/ddu319 (2014).
77 Lockhart, A. *et al.* RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. *Cell Rep* **29**, 2890-2900 e2895, doi:10.1016/j.celrep.2019.10.108 (2019).
78 Zatreanu, D. *et al.* Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. *Mol Cell* **76**, 57-69 e59, doi:10.1016/j.molcel.2019.07.037 (2019).
79 Li, X. & Manley, J. L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122**, 365-378, doi:10.1016/j.cell.2005.06.008 (2005).
80 Chakraborty, P., Huang, J. T. J. & Hiom, K. DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. *Nat Commun* **9**, 4346, doi:10.1038/s41467-018-06677-1 (2018).
81 Sorrells, S. *et al.* Spliceosomal components protect embryonic neurons from R-loop-mediated DNA damage and apoptosis. *Dis Model Mech* **11**, dmm031583 (2018).
82 Puhringer, T. *et al.* Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *Elife* **9**, e61503 (2020).
83 Salas-Armenteros, I. *et al.* Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J* **36**, 3532-3547, doi:10.15252/embj.201797208 (2017).
84 Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**, 362-365, doi:10.1038/nature13374 (2014).

85 Nojima, T. *et al.* Deregulated Expression of Mammalian lncRNA through Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence. *Mol Cell* **72**, 970-984 e977, doi:10.1016/j.molcel.2018.10.011 (2018).
86 Pefanis, E. *et al.* RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* **161**, 774-789, doi:10.1016/j.cell.2015.04.034 (2015).
87 Manzo, S. G. *et al.* DNA Topoisomerase I differentially modulates R-loops across the human genome. *Genome Biol* **19**, 100, doi:10.1186/s13059-018-1478-1 (2018).
88 Hatchi, E. *et al.* BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol Cell* **57**, 636-647, doi:10.1016/j.molcel.2015.01.011 (2015).
89 Jurga, M., Abugable, A. A., Goldman, A. S. H. & El-Khamisy, S. F. USP11 controls R-loops by regulating senataxin proteostasis. *Nat Commun* **12**, 5156, doi:10.1038/s41467-021-25459-w (2021).
90 Perez-Calero, C. *et al.* UAP56/DDX39B is a major cotranscriptional RNA-DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev* **34**, 898-912, doi:10.1101/gad.336024.119 (2020).
91 Song, C., Hotz-Wagenblatt, A., Voit, R. & Grummt, I. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev* **31**, 1370-1381, doi:10.1101/gad.300624.117 (2017).
92 Cristini, A., Groh, M., Kristiansen, M. S. & Gromak, N. RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Rep* **23**, 1891-1905, doi:10.1016/j.celrep.2018.04.025 (2018).
93 Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst)* **10**, 654-665, doi:10.1016/j.dnarep.2011.04.013 (2011).
94 Rivosecchi, J. *et al.* Senataxin homologue Sen1 is required for efficient termination of RNA polymerase III transcription. *EMBO J* **38**, e101955, doi:10.15252/embj.2019101955 (2019).
95 Osmundson, J. S., Kumar, J., Yeung, R. & Smith, D. J. Pif1-family helicases cooperatively suppress widespread replication-fork arrest at tRNA genes. *Nat Struct Mol Biol* **24**, 162-170, doi:10.1038/nsmb.3342 (2017).
96 Tran, P. L. T. *et al.* PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. *Nat Commun* **8**, 15025, doi:10.1038/ncomms15025 (2017).
97 Zhang, X. *et al.* Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. *Nat Commun* **8**, 15908, doi:10.1038/ncomms15908 (2017).
98 Shivji, M. K. K., Renaudin, X., Williams, C. H. & Venkitaraman, A. R. BRCA2 Regulates Transcription Elongation by RNA Polymerase II to Prevent R-Loop Accumulation. *Cell Rep* **22**, 1031-1039, doi:10.1016/j.celrep.2017.12.086 (2018).
99 Gomez-Gonzalez, B., Sessa, G., Carreira, A. & Aguilera, A. A new interaction between BRCA2 and DDX5 promotes the repair of DNA breaks at transcribed chromatin. *Mol Cell Oncol* **8**, 1910474, doi:10.1080/23723556.2021.1910474 (2021).
100 Bayona-Feliu, A., Barroso, S., Munoz, S. & Aguilera, A. The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription-replication conflicts. *Nat Genet* **53**, 1050-1063 (2021).

101 Madireddy, A. *et al.* FANCD2 Facilitates Replication through Common Fragile Sites. *Mol Cell* **64**, 388-404, doi:10.1016/j.molcel.2016.09.017 (2016).
102 Okamoto, Y. *et al.* SLFN11 promotes stalled fork degradation that underlies the phenotype in Fanconi anemia cells. *Blood* **137**, 336-348, doi:10.1182/blood.2019003782 (2021).
103 Liang, Z. *et al.* Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination. *Cell Rep* **26**, 564-572 e565, doi:10.1016/j.celrep.2018.12.084 (2019).
104 Nguyen, H. D. *et al.* Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol Cell* **65**, 832-847 e834, doi:10.1016/j.molcel.2017.01.029 (2017).
105 Makharashvili, N. *et al.* Sae2/CtIP prevents R-loop accumulation in eukaryotic cells. *Elife* **7**, e42733 (2018).
106 Pan, X. *et al.* FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops. *Sci Rep* **9**, 19110, doi:10.1038/s41598-019-55537-5 (2019).
107 Silva, B. *et al.* FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. *Nat Commun* **10**, 2253, doi:10.1038/s41467-019-10179-z (2019).
108 Tan, J., Wang, X., Phoon, L., Yang, H. & Lan, L. Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase. *FEBS Lett* **594**, 1359-1367, doi:10.1002/1873-3468.13738 (2020).
109 Pan, X., Ahmed, N., Kong, J. & Zhang, D. Breaking the end: Target the replication stress response at the ALT telomeres for cancer therapy. *Mol Cell Oncol* **4**, e1360978, doi:10.1080/23723556.2017.1360978 (2017).
110 Alzu, A. *et al.* Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* **151**, 835-846, doi:10.1016/j.cell.2012.09.041 (2012).
111 Kim, S. *et al.* ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork. *Nucleic Acids Res* **48**, 7218-7238, doi:10.1093/nar/gkaa501 (2020).
112 Hodroj, D. *et al.* An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. *EMBO J* **36**, 1182-1198, doi:10.15252/embj.201695131 (2017).
113 Chang, E. Y. *et al.* MRE11-RAD50-NBS1 promotes Fanconi Anemia R-loop suppression at transcription-replication conflicts. *Nat Commun* **10**, 4265, doi:10.1038/s41467-019-12271-w (2019).
114 Barroso, S. *et al.* The DNA damage response acts as a safeguard against harmful DNA-RNA hybrids of different origins. *EMBO Rep* **20**, e47250, doi:10.15252/embr.201847250 (2019).
115 Chappidi, N. *et al.* Fork Cleavage-Religation Cycle and Active Transcription Mediate Replication Restart after Fork Stalling at Co-transcriptional R-Loops. *Mol Cell* **77**, 528-541 e528, doi:10.1016/j.molcel.2019.10.026 (2020).
116 Svikovic, S. *et al.* R-loop formation during S phase is restricted by PrimPol-mediated repriming. *EMBO J* **38**, e99793 (2019).
117 Tsai, S. *et al.* ARID1A regulates R-loop associated DNA replication stress. *PLoS Genet* **17**, e1009238, doi:10.1371/journal.pgen.1009238 (2021).

118 Prendergast, L. *et al.* Resolution of R-loops by INO80 promotes DNA replication and maintains cancer cell proliferation and viability. *Nat Commun* **11**, 4534, doi:10.1038/s41467-020-18306-x (2020).
119 Herrera-Moyano, E., Mergui, X., Garcia-Rubio, M. L., Barroso, S. & Aguilera, A. The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. *Genes Dev* **28**, 735-748, doi:10.1101/gad.234070.113 (2014).
120 Sanchez, A. *et al.* Transcription-replication conflicts as a source of common fragile site instability caused by BMI1-RNF2 deficiency. *PLoS Genet* **16**, e1008524, doi:10.1371/journal.pgen.1008524 (2020).
121 Klusmann, I. *et al.* Chromatin modifiers Mdm2 and RNF2 prevent RNA:DNA hybrids that impair DNA replication. *Proceedings of the National Academy of Sciences* **115**, E11311, doi:10.1073/pnas.1809592115 (2018).
122 Singh, D. K. *et al.* MOF Suppresses Replication Stress and Contributes to Resolution of Stalled Replication Forks. *Mol Cell Biol* **38**, e00484-17 (2018).
123 Kim, J. J. *et al.* Systematic bromodomain protein screens identify homologous recombination and R-loop suppression pathways involved in genome integrity. *Genes Dev* **33**, 1751-1774, doi:10.1101/gad.331231.119 (2019).
124 Lam, F. C. *et al.* BRD4 prevents the accumulation of R-loops and protects against transcription–replication collision events and DNA damage. *Nature Communications* **11**, 4083, doi:10.1038/s41467-020-17503-y (2020).
125 Edwards, D. S. *et al.* BRD4 Prevents R-Loop Formation and Transcription-Replication Conflicts by Ensuring Efficient Transcription Elongation. *Cell Reports* **32**, 108166, doi:https://doi.org/10.1016/j.celrep.2020.108166 (2020).
126 Mognato, M., Burdak-Rothkamm, S. & Rothkamm, K. Interplay between DNA replication stress, chromatin dynamics and DNA-damage response for the maintenance of genome stability. *Mutat Res* **787**, 108346, doi:10.1016/j.mrrev.2020.108346 (2021).
127 Gao, M. *et al.* Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. *Cell Res* **24**, 532-541, doi:10.1038/cr.2014.36 (2014).
128 Michelini, F. *et al.* Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat Cell Biol* **19**, 1400-1411, doi:10.1038/ncb3643 (2017).
129 D'Alessandro, G. *et al.* BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nat Commun* **9**, 5376, doi:10.1038/s41467-018-07799-2 (2018).
130 Keskin, H. *et al.* Transcript-RNA-templated DNA recombination and repair. *Nature* **515**, 436-439, doi:10.1038/nature13682 (2014).
131 Yasuhara, T. *et al.* Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. *Cell* **175**, 558-570 e511, doi:10.1016/j.cell.2018.08.056 (2018).
132 Ohle, C. *et al.* Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* **167**, 1001-1013 e1007, doi:10.1016/j.cell.2016.10.001 (2016).
133 Chakraborty, P. & Hiom, K. DHX9-dependent recruitment of BRCA1 to RNA promotes DNA end resection in homologous recombination. *Nat Commun* **12**, 4126, doi:10.1038/s41467-021-24341-z (2021).

134 Mazina, O. M., Keskin, H., Hanamshet, K., Storici, F. & Mazin, A. V. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Mol Cell* **67**, 19-29 e13, doi:10.1016/j.molcel.2017.05.019 (2017).
135 Wei, L. *et al.* DNA damage during the G0/G1 phase triggers RNA-templated, Cockayne syndrome B-dependent homologous recombination. *Proc Natl Acad Sci U S A* **112**, E3495-3504, doi:10.1073/pnas.1507105112 (2015).
136 McDevitt, S., Rusanov, T., Kent, T., Chandramouly, G. & Pomerantz, R. T. How RNA transcripts coordinate DNA recombination and repair. *Nat Commun* **9**, 1091, doi:10.1038/s41467-018-03483-7 (2018).
137 Chandramouly, G. *et al.* Poltheta reverse transcribes RNA and promotes RNA-templated DNA repair. *Sci Adv* **7**, eabf1771 (2021).
138 Ouyang, J. *et al.* RNA transcripts stimulate homologous recombination by forming DR-loops. *Nature* **594**, 283-288 (2021).
139 Liang, F. *et al.* Promotion of RAD51-Mediated Homologous DNA Pairing by the RAD51AP1-UAF1 Complex. *Cell Rep* **15**, 2118-2126, doi:10.1016/j.celrep.2016.05.007 (2016).
140 Kang, H. J. *et al.* TonEBP recognizes R-loops and initiates m6A RNA methylation for R-loop resolution. *Nucleic Acids Res* **49**, 269-284, doi:10.1093/nar/gkaa1162 (2021).
141 Abakir, A. *et al.* N(6)-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nat Genet* **52**, 48-55, doi:10.1038/s41588-019-0549-x (2020).
142 Yang, X. *et al.* m(6)A promotes R-loop formation to facilitate transcription termination. *Cell Res* **29**, 1035-1038, doi:10.1038/s41422-019-0235-7 (2019).
143 Xiang, Y. *et al.* RNA m(6)A methylation regulates the ultraviolet-induced DNA damage response. *Nature* **543**, 573-576, doi:10.1038/nature21671 (2017).
144 Zhang, C. *et al.* METTL3 and N6-Methyladenosine Promote Homologous Recombination-Mediated Repair of DSBs by Modulating DNA-RNA Hybrid Accumulation. *Mol Cell* **79**, 425-442 e427, doi:10.1016/j.molcel.2020.06.017 (2020).
145 Chen, H. *et al.* m(5)C modification of mRNA serves a DNA damage code to promote homologous recombination. *Nat Commun* **11**, 2834, doi:10.1038/s41467-020-16722-7 (2020).
146 Jimeno, S. *et al.* ADAR-mediated RNA editing of DNA:RNA hybrids is required for DNA double strand break repair. *Nat Commun* **12**, 5512, doi:10.1038/s41467-021-25790-2 (2021).
147 Graf, M. *et al.* Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* **170**, 72-85 e14, doi:10.1016/j.cell.2017.06.006 (2017).
148 Silva, B., Arora, R., Bione, S. & Azzalin, C. M. TERRA transcription destabilizes telomere integrity to initiate break-induced replication in human ALT cells. *Nat Commun* **12**, 3760, doi:10.1038/s41467-021-24097-6 (2021).
149 Zhang, J. M., Yadav, T., Ouyang, J., Lan, L. & Zou, L. Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell Rep* **26**, 955-968 e953, doi:10.1016/j.celrep.2018.12.102 (2019).
150 Tan, J. *et al.* An R-loop-initiated CSB-RAD52-POLD3 pathway suppresses ROS-induced telomeric DNA breaks. *Nucleic Acids Res* **48**, 1285-1300, doi:10.1093/nar/gkz1114 (2020).
151 Hatchi, E. *et al.* BRCA1 and RNAi factors promote repair mediated by small RNAs and PALB2-RAD52. *Nature* **591**, 665-670, doi:10.1038/s41586-020-03150-2 (2021).

152 Ortega, P., Merida-Cerro, J. A., Rondon, A. G., Gomez-Gonzalez, B. & Aguilera, A. DNA-RNA hybrids at DSBs interfere with repair by homologous recombination. *Elife* **10**, e69881 (2021).
153 Zhao, H., Zhu, M., Limbo, O. & Russell, P. RNase H eliminates R-loops that disrupt DNA replication but is nonessential for efficient DSB repair. *EMBO Rep* **19**, e45335 (2018).
154 Sessa, G. *et al.* BRCA2 promotes DNA-RNA hybrid resolution by DDX5 helicase at DNA breaks to facilitate their repairdouble dagger. *EMBO J* **40**, e106018, doi:10.15252/embj.2020106018 (2021).
155 McBride, M. J. *et al.* The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma. *Cancer Cell* **33**, 1128-1141 e1127, doi:10.1016/j.ccell.2018.05.002 (2018).
156 Bauer, M. *et al.* The ALPK1/TIFA/NF-kappaB axis links a bacterial carcinogen to R-loop-induced replication stress. *Nat Commun* **11**, 5117, doi:10.1038/s41467-020-18857-z (2020).
157 Nguyen, H. D. *et al.* Spliceosome Mutations Induce R Loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndromes. *Cancer Res* **78**, 5363-5374, doi:10.1158/0008-5472.CAN-17-3970 (2018).
158 Herold, S. *et al.* Recruitment of BRCA1 limits MYCN-driven accumulation of stalled RNA polymerase. *Nature* **567**, 545-549, doi:10.1038/s41586-019-1030-9 (2019).
159 Schwab, R. A. *et al.* The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol Cell* **60**, 351-361, doi:10.1016/j.molcel.2015.09.012 (2015).
160 Yeo, A. J. *et al.* R-loops in proliferating cells but not in the brain: implications for AOA2 and other autosomal recessive ataxias. *PLoS One* **9**, e90219, doi:10.1371/journal.pone.0090219 (2014).
161 Lovejoy, C. A. *et al.* Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* **8**, e1002772, doi:10.1371/journal.pgen.1002772 (2012).
162 Goldberg, A. D. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678-691, doi:10.1016/j.cell.2010.01.003 (2010).
163 Clynes, D. *et al.* Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nat Commun* **6**, 7538, doi:10.1038/ncomms8538 (2015).
164 Wang, Y. *et al.* G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat Commun* **10**, 943, doi:10.1038/s41467-019-08905-8 (2019).
165 Zhang, J. M. & Zou, L. Alternative lengthening of telomeres: from molecular mechanisms to therapeutic outlooks. *Cell Biosci* **10**, 30, doi:10.1186/s13578-020-00391-6 (2020).
166 Zhang, J. M., Genois, M. M., Ouyang, J., Lan, L. & Zou, L. Alternative lengthening of telomeres is a self-perpetuating process in ALT-associated PML bodies. *Mol Cell* **81**, 1027-1042 e1024, doi:10.1016/j.molcel.2020.12.030 (2021).
167 Vohhodina, J. *et al.* BRCA1 binds TERRA RNA and suppresses R-Loop-based telomeric DNA damage. *Nat Commun* **12**, 3542, doi:10.1038/s41467-021-23716-6 (2021).
168 Shiromoto, Y., Sakurai, M., Minakuchi, M., Ariyoshi, K. & Nishikura, K. ADAR1 RNA editing enzyme regulates R-loop formation and genome stability at telomeres in cancer cells. *Nat Commun* **12**, 1654, doi:10.1038/s41467-021-21921-x (2021).

169 Ghisays, F. *et al.* RTEL1 influences the abundance and localization of TERRA RNA. *Nat Commun* **12**, 3016, doi:10.1038/s41467-021-23299-2 (2021).
170 Becherel, O. J. *et al.* A new model to study neurodegeneration in ataxia oculomotor apraxia type 2. *Hum Mol Genet* **24**, 5759-5774, doi:10.1093/hmg/ddv296 (2015).
171 Katyal, S. *et al.* Aberrant topoisomerase-1 DNA lesions are pathogenic in neurodegenerative genome instability syndromes. *Nat Neurosci* **17**, 813-821, doi:10.1038/nn.3715 (2014).
172 Grunseich, C. *et al.* Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. *Mol Cell* **69**, 426-437 e427, doi:10.1016/j.molcel.2017.12.030 (2018).
173 Groh, M., Lufino, M. M., Wade-Martins, R. & Gromak, N. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet* **10**, e1004318, doi:10.1371/journal.pgen.1004318 (2014).
174 Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195-200, doi:10.1038/nature13124 (2014).
175 Wood, M. *et al.* TDP-43 dysfunction results in R-loop accumulation and DNA replication defects. *J Cell Sci* **133**, jcs244129 (2020).
176 Walker, C. *et al.* C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nat Neurosci* **20**, 1225-1235, doi:10.1038/nn.4604 (2017).
177 Park, K. *et al.* Aicardi-Goutieres syndrome-associated gene SAMHD1 preserves genome integrity by preventing R-loop formation at transcription-replication conflict regions. *PLoS Genet* **17**, e1009523, doi:10.1371/journal.pgen.1009523 (2021).
178 Mackenzie, K. J. *et al.* Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response. *EMBO J* **35**, 831-844, doi:10.15252/embj.201593339 (2016).
179 Coquel, F. *et al.* SAMHD1 acts at stalled replication forks to prevent interferon induction. *Nature* **557**, 57-61, doi:10.1038/s41586-018-0050-1 (2018).
180 Lee, S. C. *et al.* Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. *Nat Med* **22**, 672-678, doi:10.1038/nm.4097 (2016).
181 Cheruiyot, A. *et al.* Nonsense-mediated RNA Decay Is a Unique Vulnerability of Cancer Cells harboring SF3B1 or U2AF1 mutations. *Cancer Res* **81**, 4499-4513(2021).
182 De Magis, A. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc Natl Acad Sci U S A* **116**, 816-825, doi:10.1073/pnas.1810409116 (2019).
183 Amato, R. *et al.* G-quadruplex Stabilization Fuels the ALT Pathway in ALT-positive Osteosarcoma Cells. *Genes (Basel)* **11**, 304 (2020).
184 Safari, M. *et al.* R-loop-Mediated ssDNA Breaks Accumulate Following Short-Term Exposure to the HDAC Inhibitor Romidepsin. *Mol Cancer Res* **19**, 1361-1374 (2021).
185 Karanam, N. K., Ding, L., Aroumougame, A. & Story, M. D. Tumor treating fields cause replication stress and interfere with DNA replication fork maintenance: Implications for cancer therapy. *Transl Res* **217**, 33-46, doi:10.1016/j.trsl.2019.10.003 (2020).
186 Tumini, E. *et al.* The Antitumor Drugs Trabectedin and Lurbinectedin Induce Transcription-Dependent Replication Stress and Genome Instability. *Mol Cancer Res* **17**, 773-782, doi:10.1158/1541-7786.MCR-18-0575 (2019).
187 Petti, E. *et al.* SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability. *Nat Commun* **10**, 1001, doi:10.1038/s41467-019-08863-1 (2019).

188 Gomez-Gonzalez, B., Felipe-Abrio, I. & Aguilera, A. The S-phase checkpoint is required to respond to R-loops accumulated in THO mutants. *Mol Cell Biol* **29**, 5203-5213, doi:10.1128/MCB.00402-09 (2009).
189 Gomez-Gonzalez, B. *et al.* Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J* **30**, 3106-3119, doi:10.1038/emboj.2011.206 (2011).
190 Nowotny, M. *et al.* Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell* **28**, 264-276, doi:10.1016/j.molcel.2007.08.015 (2007).
191 Rychlik, M. P. *et al.* Crystal structures of RNase H2 in complex with nucleic acid reveal the mechanism of RNA-DNA junction recognition and cleavage. *Mol Cell* **40**, 658-670, doi:10.1016/j.molcel.2010.11.001 (2010).
192 Chang, E. Y. *et al.* RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instability. *J Cell Biol* **216**, 3991-4005, doi:10.1083/jcb.201703168 (2017).
193 Domingues-Silva, B., Silva, B. & Azzalin, C. M. ALTernative Functions for Human FANCM at Telomeres. *Front Mol Biosci* **6**, 84, doi:10.3389/fmolb.2019.00084 (2019).
194 Mersaoui, S. Y. *et al.* Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. *EMBO J* **38**, e100986, doi:10.15252/embj.2018100986 (2019).
195 Garcia-Rubio, M. L. *et al.* The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet* **11**, e1005674, doi:10.1371/journal.pgen.1005674 (2015).
196 Lam, F. C. *et al.* BRD4 prevents the accumulation of R-loops and protects against transcription-replication collision events and DNA damage. *Nat Commun* **11**, 4083, doi:10.1038/s41467-020-17503-y (2020).

Competing interests

The authors declare no competing interests.

Acknowledgements

We apologize to those authors whose work could not be cited due to space constrains. E. P. is supported by Cancer Research UK (C25526/A28275) and Medical Research Council (MR/S021310/1). L. L. is supported by the NIH (GM118833). L.Z. is the James & Patricia Poitras Endow Chair for Cancer Research and support by the NIH (CA263934).

Author contributions

The authors contributed equally to all aspects of the article.

Peer review information

Nature Reviews Molecular Cell Biology thanks the anonymous reviewers for their contribution to the peer review of this work.

Table 1: Proteins involved in suppression and resolution of R-loops and RNA–DNA hybrids

	R-loop suppressors	Description	References
RNA processing factors	SRSF1	Splicing factor; binds the Pol II CTD and nascent RNA transcripts	[79]
	SFPQ, NONO	Involved in RNA splicing; suppress telomeric R-loops	[187]
	SF3B1, U2AF1, SRSF2	Splicing factors; mutated in myelodysplastic syndrome and cancer	[104,157] [20]
	SPT6	Pol II-associated elongation factor recruits the Integrator complex to suppress v generated from long non-coding RNA	[85]
	TFIIS	Recognizes backtracked Pol II and stimulates transcript cleavage	[78]
	THO, TREX and TREX-2 complexes	Promote transcription termination and assembly and nuclear export of messenger ribonucleoproteins	[188] [189]
	RNA exosome	3'-5'exoribonuclease; degrades non-coding RNAs and prevents R-loop accumulation	[86]
Topoisomerase	TOP1	Removes supercoiling during replication and transcription; suppresses R-loop-associated replication stress at transcription termination sites.	[21] [62]
RNase H	RNaseH1	Cleaves RNA in RNA–DNA hybrids	[190]
	RNaseH2	Cleaves RNA in RNA-DNA hybrids; removes ribonucleotides mis-incorporated in DNA	[191] [39]
	SETX	Putative RNA–DNA helicase; suppresses R-loops at transcription termination sites and DSBs.	[14] [88] [45]
	AQR	Putative RNA–DNA helicase; suppresses R-loops in human cells.	[71]
	BLM	Unwinds RNA–DNA hybrids *in vitro*; suppress R-loops in human cells.	[192]
	FANCM	Unwinds telomeric RNA–DNA hybrids; suppresses telomeric R-loops in human cells.	[106] [193]

RNA–DNA helicases	DDX1	Associates with ATAD5 and suppresses R-loops during replication.	111
	DDX5	Associates with BRCA2 and suppresses R-loops at DSBs. Associates with ATAD5 and suppresses R-loops during replication.	194 111 154
	DDX19	Enters the nucleus following DNA damage in an ATR-dependent manner; suppresses R-loops	112
	DDX21	Unwinds RNA–DNA hybrids *in vitro*; functions with SIRT7 to suppress R-loops at specific genes. Associates with ATAD5; suppresses R-loops during replication.	91 111
	DDX39B	Unwinds RNA–DNA hybrids *in vitro*; suppresses co-transcriptional R-loops genome-wide.	90
	DHX9	Binds RNA–DNA hybrids and unwinds R-loops. Associates with ATAD5 and suppresses R-loops during replication. Also shown to promote R-loop formation.	93 80,92 111
DNA repair proteins (non-helicases)	BRCA1	Suppresses R-loops at promoter-proximal Pol II pause sites and transcription termination sites. Recruits SETX to transcription termination sites.	88 97
	BRCA2	Suppresses R-loops at promoter-proximal Pol II pause sites. Interacts with DDX5 and TREX-2.	84 98
	Fanconi anemia factors (FANCD2, FANCI, FANCA)	FANCD2–FANCI bind RNA–DNA hybrids and RNA processing factors suppress R-loops to coordinate replication and transcription	195 159 103
	RPA	Recruits and stimulates RNaseH1.	104
	XPG, XPF	Structure-specific nucleases involved in transcription-coupled nucleotide excision repair (TC-NER), may cleave and remove R-loops during transcription.	71

	MUS81	A structure-specific nuclease that cleaves Holliday junctions and stalled or reversed replication forks, may process R-loops upon collisions of R-loops and replication forks.	61 115
	CtIP	May process R-loops at active genes.	105
	MRE11	Suppresses R-loops upon collisions of R-loops and replication forks, promotes functions of FA proteins.	113
	APTX	Suppresses R-loops through its function in single-strand break repair.	160
	TDP1	Suppresses R-loops through its function in single-strand break repair.	160
	SAMHD1	Deoxynucleoside triphosphohydrolase (dNTPase) and 3′-5′ -exoribonuclease. Does not resolve R-loops directly but can recruit other factors such as MRE11.	177
DNA-damage checkpoint proteins	ATR, Chk1	May suppress R-loops by preventing collisions between R-loops and replication forks.	61 114
	ATM, Chk2	May suppress R-loops by promoting the repair of DSBs at replication forks collapsed at R-loops.	114
Chromatin modulators	SWI/SNF	ATP-dependent chromatin-remodeling complexes, may suppress R-loops during transcription–replication conflicts.	100
	INO80	Part of the ATP-dependent chromatin-remodeling complex, may reorganize chromatin to resolve R-loops.	118
	FACT	Histone chaperone, may reorganize chromatin at sites of R-loop-replication collisions.	119
	PCR1	Promotes repressive chromatin, suppresses R-loops by decreasing transcription.	121
	SIN3A	A histone deacetylase that interacts with the THO complex.	83
	KAT8	A histone acetyltransferase that functions with BRD2 and BRD4 to suppress R-loops.	122

	BRD2	A reader of histone Lys acetylation; recruits TOP1.	[123]
	BRD4	A reader of histone Lys acetylation; suppresses R-loops at specific genes by preventing Pol II pausing.	[196] [125]

APTX, aprataxin; AQR, RNA helicase aquarius; ATAD5, ATPase family AAA domain-containing protein 5; ATM, ataxia-telangiectasia mutated; ATR, ATM and Rad3-related; BLM, Bloom syndrome protein; BRCA1, breast cancer type 1 susceptibility protein; BRD2, bromodomain-containing protein 2; Chk1, checkpoint kinase 1; CtIP, DNA endonuclease RBBP8; DDX1, ATP-dependent RNA helicase DDX1; DHX9, ATP-dependent RNA helicase A; DSBs, DNA double-stranded breaks; FA, Fanconi anemia; FANCD2, Fanconi anemia group D2 protein; KAT8, Histone acetyltransferase KAT8; Lys, lysine; MRE11, double-strand break repair protein MRE11; MUS81, crossover junction endonuclease MUS81; NONO, Non-POU domain-containing octamer-binding protein; Pol II CTD, RNA polymerase II C-terminal domain; PCR1, Polycomb repressor complex 1; RNaseH1, ribonuclease H1; RPA, replication protein A; SAMHD1, deoxynucleoside triphosphate triphosphohydrolase SAMHD1; SETX, senataxin; SIRT7, NAD-dependent protein deacetylase sirtuin-7; SFPQ, splicing factor, proline- and glutamine-rich; SF3B1, splicing factor 3B subunit 1; SPT6, transcription elongation factor SPT6; SRSF1, serine/arginine-rich splicing factor 1; TFIIS, transcription elongation factor A protein 1; TDP1, tyrosyl-DNA phosphodiesterase 1; TOP1, topoisomerase 1; TREX-2; three prime repair exonuclease 2; U2AF1, splicing factor U2AF 35 kDa subunit; XPG, Xeroderma Pigmentosum complementation group G.

Figure legends

Fig. 1. Sources of R-loops and other RNA–DNA hybrids (I).

- **a.** *Co-transcriptional R-loops.* Left: Formation of R-loops at RNA polymerase II (Pol II) pause sites near transcription start sites (TSSs), at sequences with high GC skew in gene bodies, and at transcription termination sites (TTSs). Right: The pausing of Pol II at TTSs induces antisense transcription, which recruits Argonaute-2 (AGO2) and the histone Lys methyltransferase G9a (also known as EHMT2)[15]. G9a di-methylates histone H3 Lys9 (H3K9me2) and promotes the recruitment of heterochromatin protein 1 (HP1) and the formation of repressive chromatin.
- **b.** *Telomeric and centromeric R-loops.* Top: The long non-coding RNA telomeric repeat-containing RNA (TERRA) forms R-loops at telomeres. TERRA interacts with telomere repeat binding factor 2 (TRF2) and ATRX, thereby contributing to telomere maintenance. Telomeric R-loops also induce replication stress, which promotes alternative lengthening of telomeres (ALT). Bottom: RNA transcripts from centromeric repeats form R-loops in *cis*. Centromeric RNAs bind histone H3-like centromeric protein A (CENPA), thereby promoting the localization of CENPA and centromere protein C (CENPC) to centromeres. In mitosis, centromeric R-loops recruit and activate the kinase ATR, which activates CHK1 and aurora kinase B (AurB) at centromeres and enables accurate chromosome segregation.

Fig. 2. Sources of R-loops and other RNA–DNA hybrids (II).

a. *R-loops generated during replication.* Top: The DNA polymerase α (Polα)-primase complex synthesizes RNA primers in Okazaki fragments on the lagging strand, thereby forming RNA–DNA hybrids. The RNA primers are subsequently removed by flap endonuclease 1 (FEN1) and the DNA is ligated by DNA ligase 1 (LIG1). Ribonucleotides (NTP) misincorporation during DNA replication can also generate RNA–DNA hybrids. Ribonucleases H2 (RNaseH2)-mediated ribonucleotide excision repair (RER) removes misincorporated ribonucleotides. Bottom: When a leading-strand DNA polymerase is stalled, DNA-directed primase/polymerase protein (PrimPol) re-primes replication ahead of the stalled polymerase.

b. *DNA damage-induced hybrids.* Left: RNA polymerase II (Pol II) is recruited to DNA double-stranded breaks (DSBs) by the MRE11–RAD50–NBS1 (MRN) complex, enabling Pol II to carry out *de novo* RNA synthesis. Pol II-displaced DNA flaps are bound by replication protein A (RPA). Right: DSBs in transcriptionally-active regions induce Pol II pausing, allowing RNA transcripts to form RNA–DNA hybrids with help from RAD51 associated protein 1 (RAD51AP1) and USP1 associated factor 1 (UAF1)[138]. Pol II may pause before or after DSBs, and the RNA–DNA hybrids at DSBs may exist in R-loops or on ssDNA ssDNA overhangs.

Fig. 3. Effects of R-loops and RNA–DNA hybrids on genomic stability.

a. *Co-directional and head-on collisions of R-loops and replication forks.* Top: Co-directional collisions of replication forks with R-loops allow the replicative CDC45–MCM–GINS (CMG) helicase to unwind RNA–DNA hybrids, leading to dissociation of RNA and polymerase II (Pol II) and reduction of R-loop levels. Bottom: Head-on collisions of replication forks and R-loops force fork reversal. The processing of reversed forks by the structure-specific endonuclease MUS81 promotes activation of the kinase ATR, which in turn limits the activity of MUS81 and suppresses DSB formation.

b. *DNA supercoiling-induced R-loops and collisions.* DNA supercoiling ahead of Pol II promote head-on collisions of R-loops and replication forks at transcription termination sites (TTSs). Topoisomerase 1 (TOP1) resolves the supercoils and alleviates replication fork collapse and consequently the formation of DNA double-stranded breaks (DSBs) at TTSs. In the absence of TOP1, replication protein A (RPA), which binds single-strand DNA (ssDNA) at TTSs becomes increasingly phosphorylated by ATR, which may facilitate the repair of DSBs.

c. *R-loop-associated chromatin compaction.* R-loops are associated with gene-repressive histone modifications such as histone H3 Lys9 di-methylation (H3K9me2) and histone H3 Ser10 phosphorylation (H3S10p) and with compacted chromatin. H3K9me2 recruits heterochromatin protein 1 (HP1) to form heterochromatin. The compact chromatin induced by R-loops interferes with DNA replication and increases genomic instability.

d. *Single-strand DNA damage in R-loops.* The displaced ssDNA in R-loops is a substrate of activation induced cytidine deaminase (AID) and APOBEC deaminases. The uracil resulting from cytosine deamination is recognized by uracil-DNA glycosylase (UNG) who generates abasic sites, which are subsequently converted to SSBs through base excision repair. Both abasic sites and SSBs interfere with replication forks, leading to stalling of DNA polymerases and fork collapse.

Fig. 4. Mechanisms of R-loop suppression and resolution.

a. *R-loop suppression during transcription.* Ribonuclease H1 (RNaseH1 or RNH1) and RNaseH2 (RNH2) cleave the RNA in RNA–DNA hybrids, and RNA–DNA helicases such as senataxin (SETX) unwind R-loops. R-loop suppressors include chromatin modifiers that repress transcription, such as the histone H2A ubiquitylating (Ub) Polycomb repressive complex 1 (PRC1) and the histone deacetylase complex SIN3A. Lys acetyltransferase 8 (KAT8) can also suppress R-loops, by acetylating (Ac) histone H4 and thus recruiting bromodomain-containing protein 2 (BRD2) and BRD4; in turn, BRD2 recruits topoisomerase 1 (TOP1), which resolves DNA supercoiling and helps suppress R-loops. Prevention of Pol II pausing and backtracking by transcription regulators such as BRD4 counteracts R-loop formation. RNA splicing factors, such as serine and arginine rich splicing factor 1 (SRSF1), and factors involved in the assembly of messenger ribonucleoproteins (mRNPs) for mRNA export, such as the THO complex (THO), also prevent RNA–DNA hybrid formation.

b. *R-loop bypass or resolution during DNA replication.* Bypass R-loops by replication forks is aided by chromatin remodelers such as the complex facilitates chromatin transactions (FACT), the INO80 complex and the SWI/SNF complex, which act in concert with Fanconi anemia group D2 protein (FANCD2) and the helicase senataxin (SETX). The activation of the kinase ATR by reversed replication forks following collision with R-loops prevents excessive fork cleavage by the structure-specific endonuclease MUS81 and promotes nuclear import of the helicase DEAD-box19 (DDX19) to remove R-loops. Replisomes might also resolve R-loops through unwinding by the replicative helicase minichromosome maintenance (MCM), or bypass R-loops through repriming by Polα-primase (not shown) and DNA-directed primase/polymerase protein (PrimPol), for example at triplex structures formed by GAA repeats. RNaseH2 (RNH2) interacts with PCNA and can remove R-loops at replication forks.

c. *Repair and restart of DNA replication forks.* Replication fork reversal at R-loops can lead to fork processing by MUS81, followed by fork repair and restart involving the recombination protein RAD52 and DNA ligase IV (LIG4), and resumption of transcription activity. Fork repair might be aided by chromatin modifiers such as PRC1 and KAT8.
H2A, histone H2A; H4, histone H4

Fig. 5. Roles of R-loops and RNA–DNA hybrids in DNA repair.

a. *Recruitment of DNA repair proteins.* Left: At sites of double-stranded DNA breaks (DSBs), damage-induced long non-coding RNAs (dilncRNAs) are transcribed and converted into double-strand RNAs, which are processed into small RNAs (DDRNAs or diRNAs) by the ribonuclease III Dicer pathway. diRNAs interact with Argonaute-2 (AGO2) and the homologous recombination (HR) factor RAD51 to promote the localization of RAD51 at DSBs. The RNA–DNA hybrids formed by dilncRNAs are also recognized by breast cancer type 1 susceptibility protein (BRCA1), thereby promoting the recruitment of BRCA2 and RAD51. Right: Reactive oxygen species (ROS) generate single-stranded DNA breaks (SSBs) and DSBs in transcribed regions, thereby inducing R-loop formation. The R-loops are recognized by Cockayne syndrome B (CSB) and RAD52, enabling RAD51 recruitment and DSB repair. The enzyme tRNA aspartic acid methyltransferase 1 (TRDMT1) modifies

the RNA in R-loops with 5-methylcytosine (m^5C), thereby promoting RAD52 and RAD51 recruitment.

b. *Single-stranded DNA and DNA end resection.* Left: In cells in G2 phase, DSB-induced R-loops are recognized by RAD52. RAD52 recruits xeroderma pigmentosum group G (XPG) to cleave R-loops, generating ssDNA that is bound by replication protein A (RPA), thereby enhancing HR activity. Right: RNA polymerase III (Pol III) is recruited to DSBs by the MRN complex and CTBP-interacting protein (CtIP). Pol III-mediated RNA synthesis promotes DNA end resection by unwinding double-stranded DNA (ssDNA). The ssDNA flap displaced by Pol III is bound by RPA and removed by DNA replication helicase/nuclease 2 (DNA2) and exonuclease 1 (EXO1).

c. *RNA-templated repair.* When DSBs form in transcribed regions, the preexisting RNA transcripts can hybridize with ssDNA overhangs by RAD52 thereby enabling DNA polymerase θ (POLθ) to extend the ssDNA using RNA as template. The extended ssDNA from one DSB end can captures the ssDNA from the other DSB end, allowing DSB repair by gap filling.

d. *DR-loops.* Preexisting RNA transcripts at DSBs can invade homologous DNA with the help of RAD51 associated protein 1 (RAD51AP1) and USP1 associated factor 1 (UAF1). The formation of R-loops in homologous DNA promotes ssDNA invasion and extension, giving rise to DR-loops. The RNA–DNA hybrids in DR-loops must be removed before the completion of HR.

e. *Repair of single-stranded DNA breaks.* When SSBs form at the transcription termination sites, a complex of BRCA1 and RNAi factors processes local RNA transcripts into single-stranded DNA-damage-associated small RNAs (sdRNAs), which interact with RAD52 and partner and localizer of BRCA2 (PALB2) to localize them at SSB sites to repair the lesions.

Fig. 6. R-loops in pathology.

a. *Gain of R-loops through increased RNA synthesis.* Left: Oncogenes such as EWS-FLI1 (fusion of EWS RNA binding protein 1 (EWSR1) and Friend leukemia integration 1 transcription factor (FLI1)), $HRAS^{V12}$ and the estrogen receptor (ER) can promote R-loop formation by increasing RNA synthesis. Right:

b. *Gain of R-loops through loss of RNA processing or DNA repair.* Suspected oncogenic mutations in splicing factors such as U2 small nuclear RNA auxiliary factor 1 (U2AF1) and serine/arginine-rich splicing factor 2 (SRSF2) may promote R-loop formation by increasing RNA polymerase II (Pol II) pausing owing to defects in co-transcriptional splicing. Loss of BRCA1 or BRCA2 also leads to increased Pol II pausing, and loss of BRCA2 additionally results in defective mRNP export by the TREX-2 complex, leading to R-loop accumulation. Loss of BRCA1 or Fanconi anemia group M (FANCM) may increase R-loop levels through loss of DNA repair.

c. *Telomeric R-loops in cancer cells.* Loss of ATRX is associated with increased telomeric R-loops in cancer cells with alternative lengthening of telomeres (ALT) activity. FANCM and BRCA1 suppress the levels of telomeric R-loops, thereby keeping telomere replication stress and ALT activity at tolerable levels.

d. *Excessive R-loop formation in neurodegenerative disorders.* In cerebellar ataxias caused by loss of DNA repair proteins such as ataxia telangiectasia mutated (ATM), tyrosyl-DNA phosphodiesterase 1 (TDP1) or senataxin (SETX), R-loop accumulation could exacerbate

cytotoxic DNA damage and inhibit Pol II, leading to cell death. In amyotrophic lateral sclerosis, frontotemporal dementia and Friedreich ataxia, R-loops accumulating owing to senataxin mutations or guanine (G)-rich repeat expansions could inhibit Pol II, leading to reduced expression of essential genes and cell death.

Fig. 7. Targeting R-loop-associated vulnerabilities in cancer therapy.

- **a.** *Inhibition of ATR.* In myelodysplastic syndrome and acute myeloid leukemia cells, mutant splicing factors such as $U2AF1^{S34F}$ increase R-loop accumulation and activate ATR. ATR provides a protective feedback to suppress R-loop-induced double-strand DNA breaks (DSBs). ATR inhibitior (ATRi) disrupts the ATR-mediated feedback loop, thereby increasing DSBs and promoting cell death.
- **b.** *Inhibition of nonsense-mediated mRNA decay.* Splicing factor 3B subunit 1 (SF3B1) and nonsense mediated decay (NMD) associated PI3K related kinase (SMG1) prevent aberrant R-loop accumulation during transcription. In MDS and AML cells carrying spliceosome mutations, SF3B1 inhibitor (SF3B1i) and SMG1 inhibitor (SMG1i) further increase R-loop levels, exacerbating R-loop-associated DNA damage and driving cell death.
- **c.** *Inhibition of DNA repair.* The Ewing's sarcoma oncogene EWS-FLI increases R-loop levels and the interaction between elongating Pol II and BRCA1. The binding of BRCA1 to Pol II compromises the localization of BRCA1 to DSBs, thereby reducing homologous recombination (HR) activity and rendering Ewing's sarcoma cells sensitive to poly(ADP-ribose) polymerase inhibitors (PARPi).
- **d.** *Stabilization of G-quadruplexes.* Cancer cells with defects in R-loop suppressors such as BRCA2, are prone to form G-quadruplexes (G4s) in sequences with high GC skew owing to the displacement of G-rich single-strand DNA in R-loops. G4-stabilizing drugs can increase the levels of G4s and R-loops in cancer cells. Collisions of replication forks with G4-containing R-loops (left) or with G4s alone (right) can lead to DNA damage and cell death.

Glossary

BACKTRACKING
Backward movement of transcribing RNA polymerase, which enables proofreading and regulation of transcript elongation.

BREAK-INDUCED REPLICATION
A process in which one-ended DNA breaks generated at replication forks are extended using homologous DNA as template.

CEREBELLAR ATAXIA
Progressive neurological disorder caused by damage to the cerebellum, characterised by inability to control balance, gait and muscle coordination.

CGAS–STING PATHWAY
Cellular signaling pathway that senses DNA in the cytoplasm as a sign of viral or bacterial infection and activates innate immunity responses.

CLASS SWITCH RECOMBINATION
The process of switching the antibody type produced by mature B cells, during which the immunoglubulin gene is subject to transcription-dependent DNA damage followed by repair-mediated rearrangements.

DONOR DNA
The DNA sequence used as template for DNA repair during homologous recombination.

CONTROL REGION OF MTDNA
A non-coding region of the mitochondrial genome that controls RNA and DNA synthesis.

C-RICH STRAND OF TELOMERE DNA
Telomere DNA consists of TTAGGG repeats on the G-rich strand and CCCTAA repeats on the C-rich strand.

G-QUADRUPLEX
A DNA secondary secture formed by guanine (G)-rich sequences through G–G base pairing.

INTEGRATOR COMPLEX
A multisubunit protein complex with RNA endonuclease activity that controls the expression and processing of Pol II transcripts.

RNA EXOSOME
A multisubunit protein complex with 3'-5'exoribonuclease activity that degrades non-coding Pol II transcripts.

Nonsense-mediated mRNA decay
A cellular surveillance mechanism that detects and degrades mRNAs harboring premature termination codons.

REPLICATION FORK REVERSAL
Backward movement of the replication fork during which the nascent newly-synthesized strands dissociate from the template strands and anneal together to form a four-way junction.

REPLICATION STRESS
A plethora of DNA replication impedements that compromise the efficiency or fidelity of DNA synthesis and increase genomic instability.

DNA end resection
A process in which exonucleases cleave one of the DNA strands at DNA double-stranded breaks to generate overhangs.

RNA EDITING
Post-transcriptional enzymatic process that changes RNA nucleotides, for example by deaminating adenosine to inosine.

TRANSCRIPTION-REPLICATION CONFLICTS
The conllisions between transcription and DNA replication complexes and consequences in both transcription and DNA replication.

TRANSLESION SYNTHESIS
A process in which a group of specialized DNA polymerases at or behind replication forks bypass DNA lesions to ensue replication.

TUMOR-TREATING ELECTRIC FIELDS
A non-invasive treatment where alternating electric fields are applied to tumour sites to inhibit cancer cell proliferation

Fig. 1

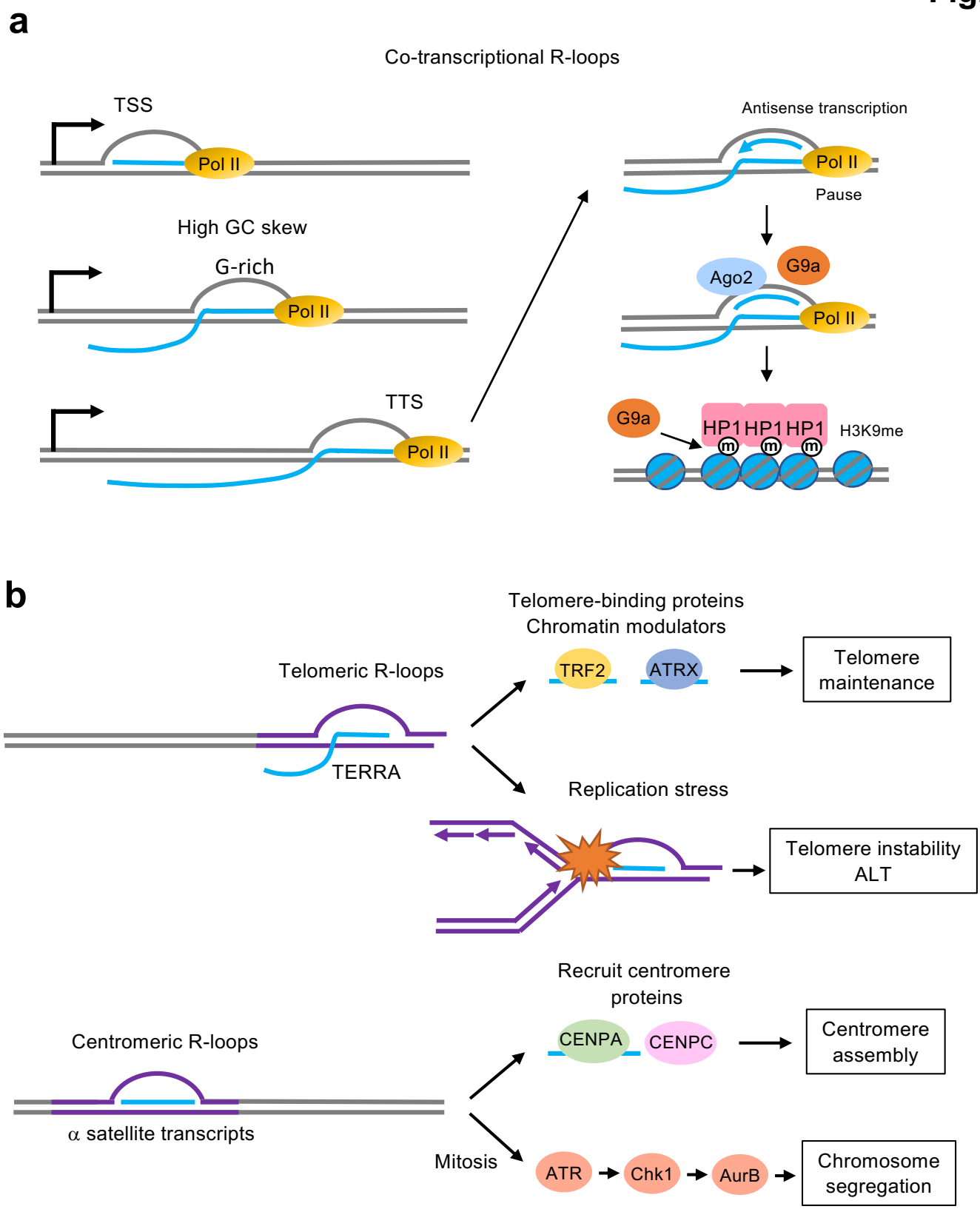

a
Co-transcriptional R-loops
TSS
Pol II
High GC skew
G-rich
Pol II
TTS
Pol II
Antisense transcription
Pol II
Pause
Ago2
G9a
Pol II
G9a
HP1 HP1 HP1
m m m
H3K9me
b
Telomere-binding proteins
Chromatin modulators
Telomeric R-loops
TRF2
ATRX
Telomere maintenance
TERRA
Replication stress
Telomere instability
ALT
Recruit centromere proteins
Centromeric R-loops
CENPA
CENPC
Centromere assembly
α satellite transcripts
Mitosis
ATR
Chk1
AurB
Chromosome segregation

Fig. 2

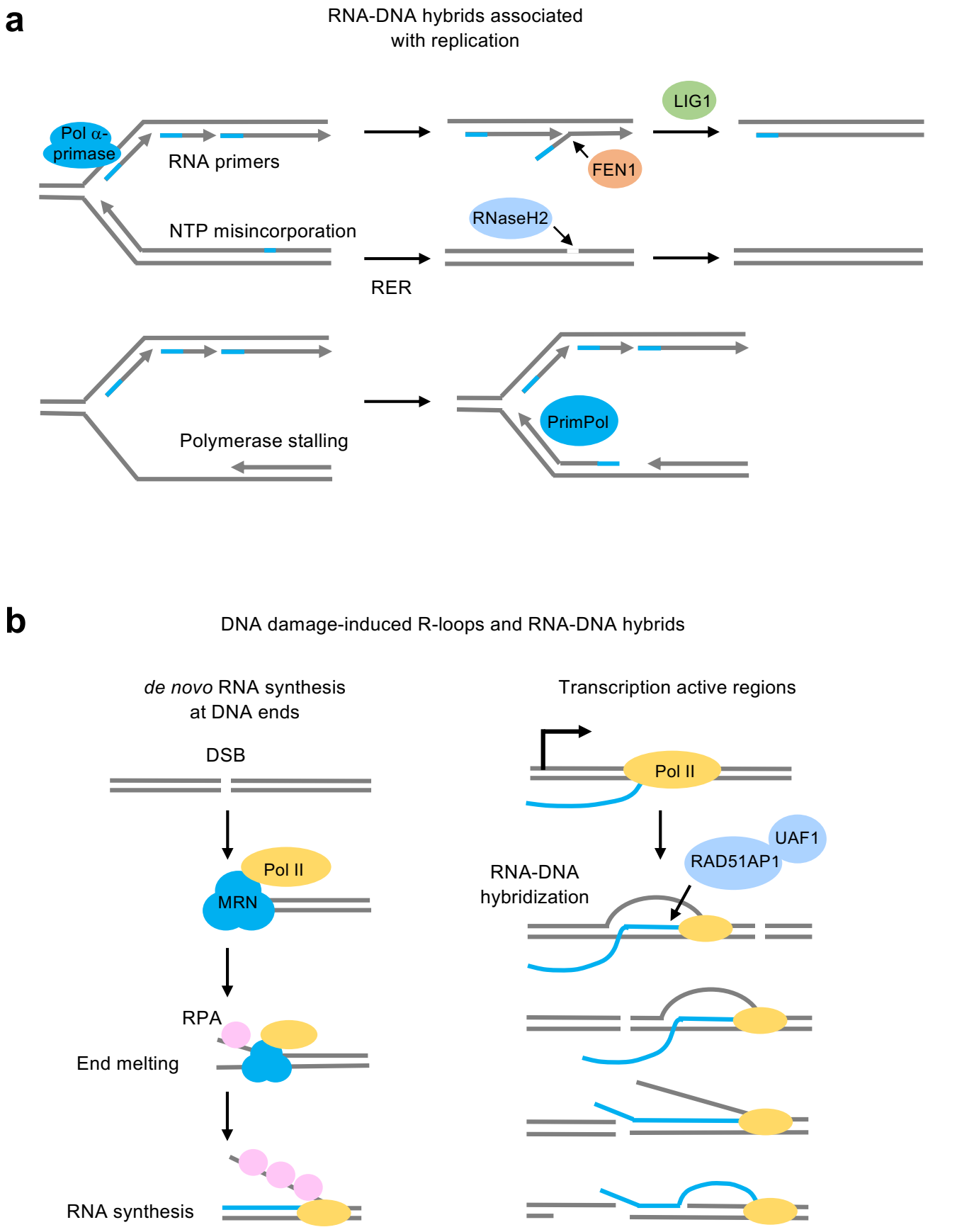

a
RNA-DNA hybrids associated with replication
Pol α-primase
RNA primers
NTP misincorporation
RER
LIG1
FEN1
RNaseH2
Polymerase stalling
PrimPol
b
DNA damage-induced R-loops and RNA-DNA hybrids
de novo RNA synthesis at DNA ends
DSB
Pol II
MRN
RPA
End melting
RNA synthesis
Transcription active regions
Pol II
UAF1
RAD51AP1
RNA-DNA hybridization

Fig. 3

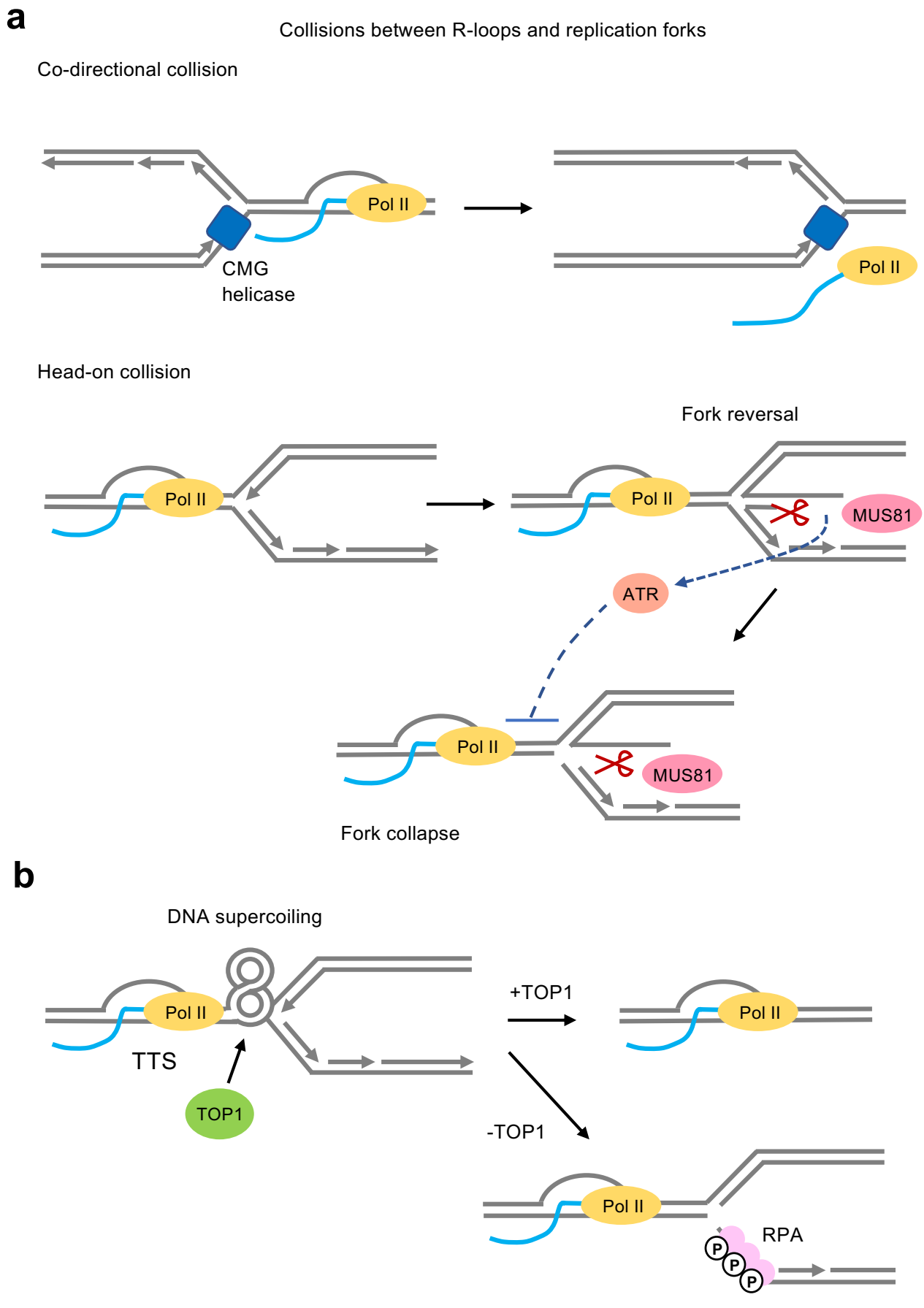
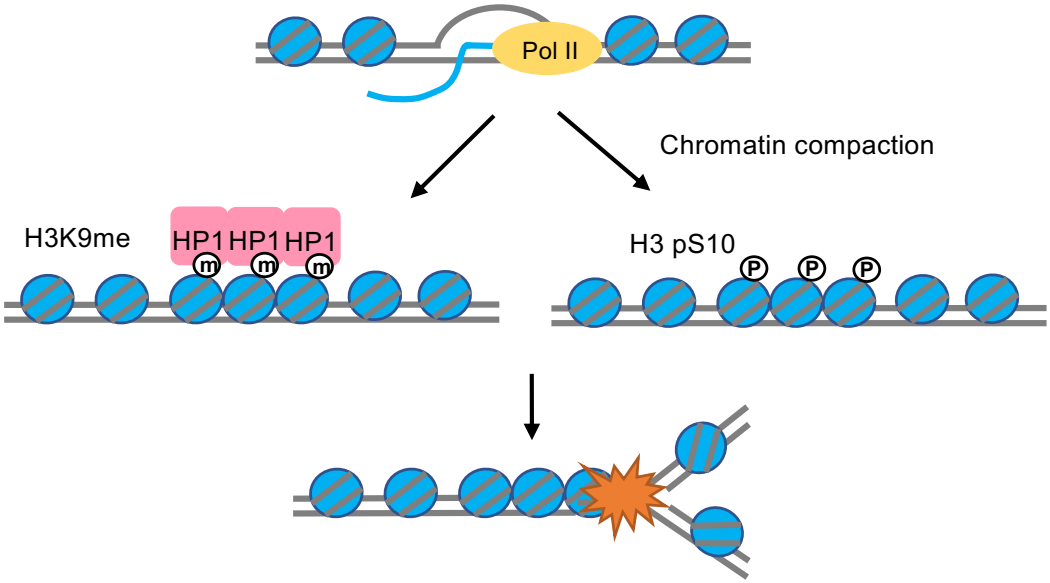

a
Collisions between R-loops and replication forks
Co-directional collision
Pol II
CMG
helicase
Pol II
Head-on collision
Fork reversal
Pol II
Pol II
MUS81
ATR
Pol II
MUS81
Fork collapse
b
DNA supercoiling
Pol II
TTS
TOP1
+TOP1
Pol II
-TOP1
Pol II
RPA
P
P
P

Fig. 3

c


Pol II
Chromatin compaction
H3K9me
HP1 HP1 HP1
m m m
H3 pS10
P P P

d

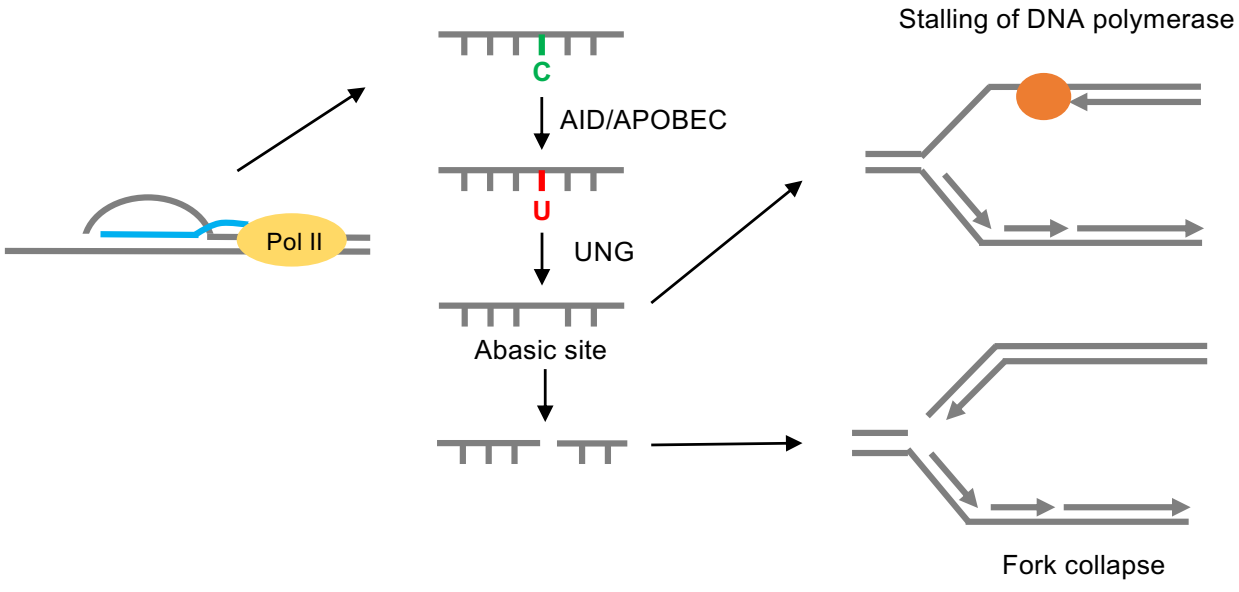

Pol II
C
AID/APOBEC
U
UNG
Abasic site
Stalling of DNA polymerase
Fork collapse

Fig. 4

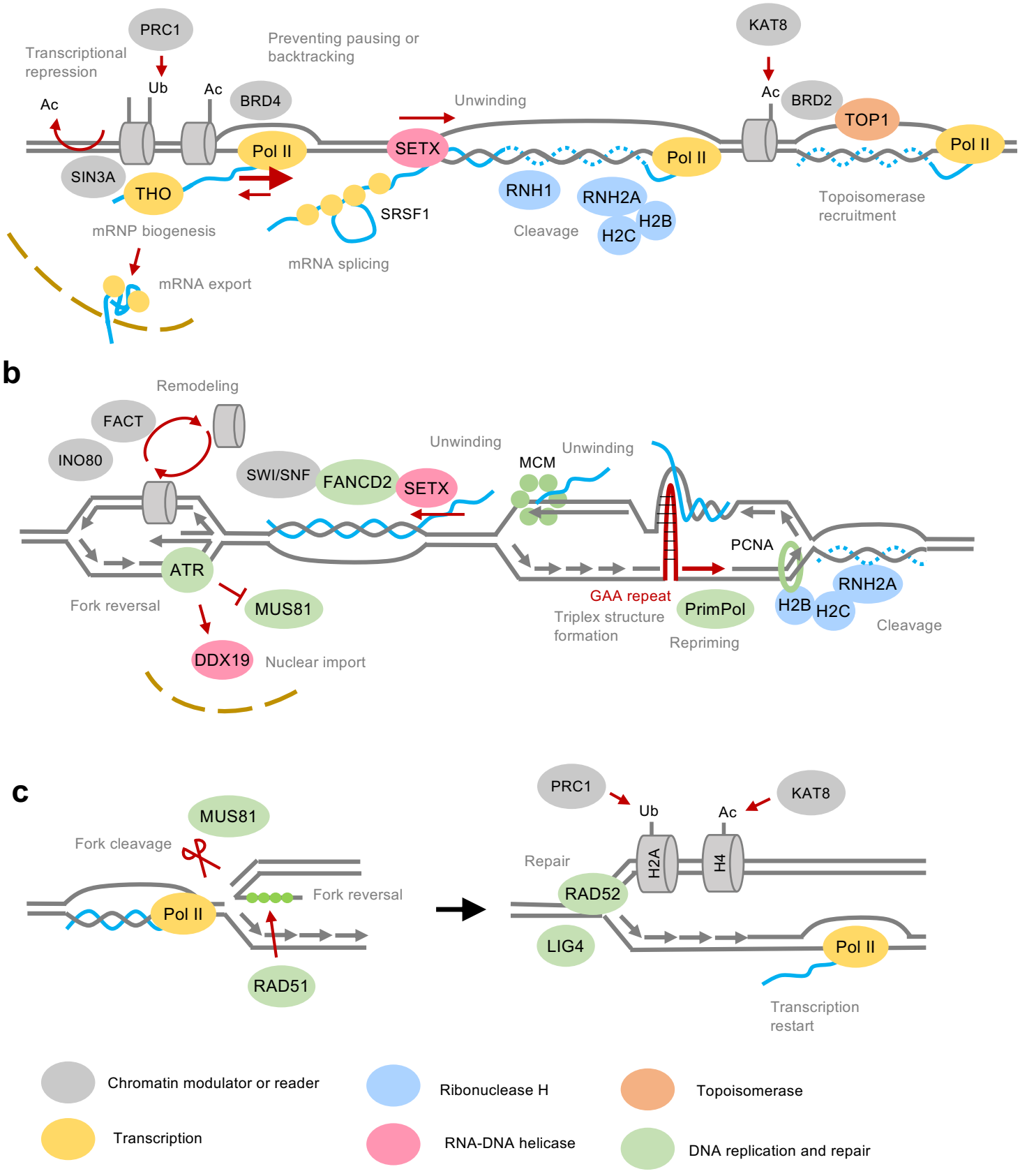

a
PRC1
Transcriptional repression
Ub
Ac
Ac
BRD4
Preventing pausing or backtracking
Unwinding
KAT8
Ac
BRD2
TOP1
Pol II
Pol II
SETX
Pol II
SIN3A
THO
SRSF1
RNH1
RNH2A
H2B
H2C
Topoisomerase recruitment
mRNP biogenesis
mRNA splicing
Cleavage
mRNA export
b
Remodeling
FACT
INO80
SWI/SNF
FANCD2
SETX
Unwinding
MCM
Unwinding
PCNA
ATR
Fork reversal
MUS81
GAA repeat
Triplex structure formation
PrimPol
Repriming
H2B
H2C
RNH2A
Cleavage
DDX19
Nuclear import
c
MUS81
Fork cleavage
Pol II
Fork reversal
RAD51
PRC1
Ub
Ac
KAT8
Repair
H2A
H4
RAD52
LIG4
Pol II
Transcription restart
Chromatin modulator or reader
Transcription
Ribonuclease H
RNA-DNA helicase
Topoisomerase
DNA replication and repair

Fig. 5

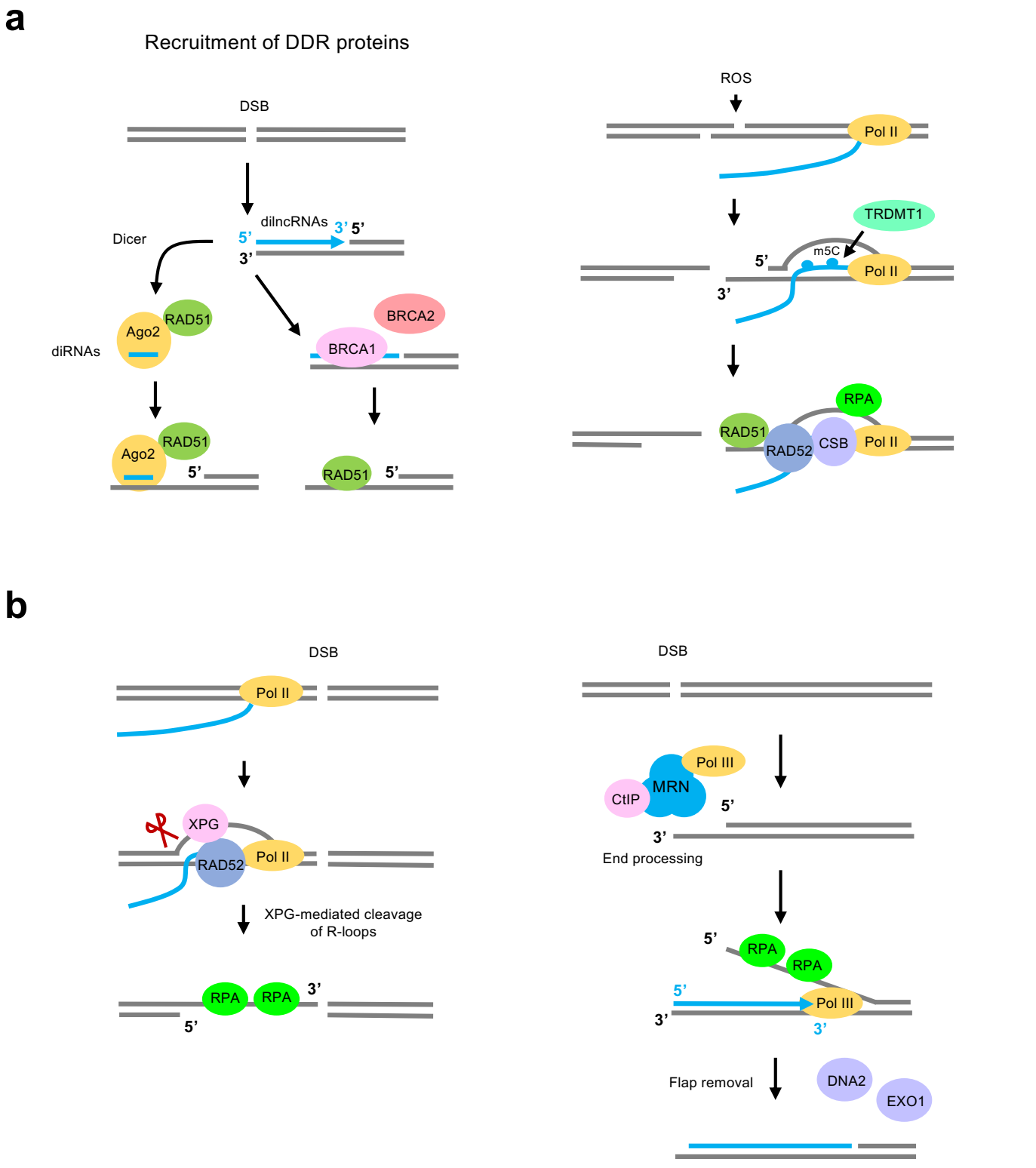
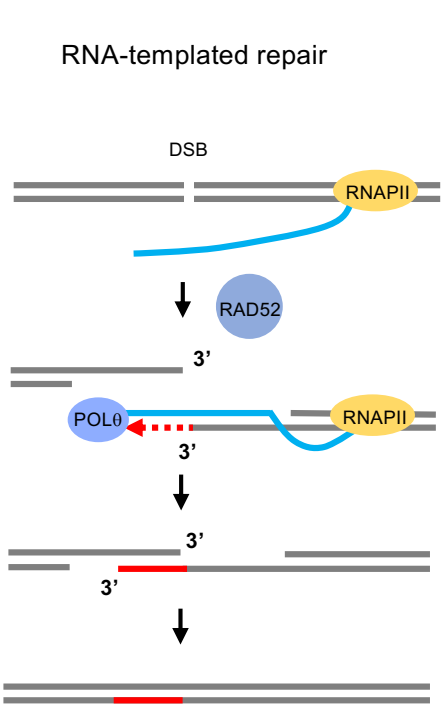
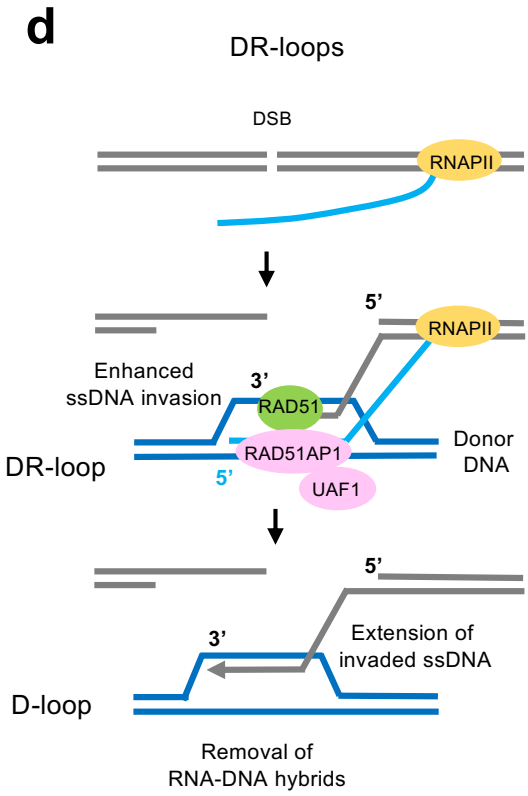

a
Recruitment of DDR proteins
DSB
Dicer
dilncRNAs
5'
3'
3'
5'
diRNAs
Ago2
RAD51
BRCA2
BRCA1
Ago2
RAD51
5'
RAD51
5'
ROS
Pol II
TRDMT1
m5C
5'
Pol II
3'
RPA
RAD51
RAD52
CSB
Pol II
b
DSB
Pol II
XPG
RAD52
Pol II
XPG-mediated cleavage
of R-loops
RPA
RPA
3'
5'
DSB
Pol III
CtIP
MRN
5'
3'
End processing
5'
RPA
RPA
5'
Pol III
3'
3'
Flap removal
DNA2
EXO1

Fig. 5

c


RNA-templated repair
DSB
RNAPII
RAD52
3'
POLθ
RNAPII
3'
3'
3'

d


DR-loops
DSB
RNAPII
5'
RNAPII
Enhanced
ssDNA invasion
3'
RAD51
DR-loop
RAD51AP1
Donor
DNA
5'
UAF1
5'
3'
Extension of
invaded ssDNA
D-loop
Removal of
RNA-DNA hybrids

e

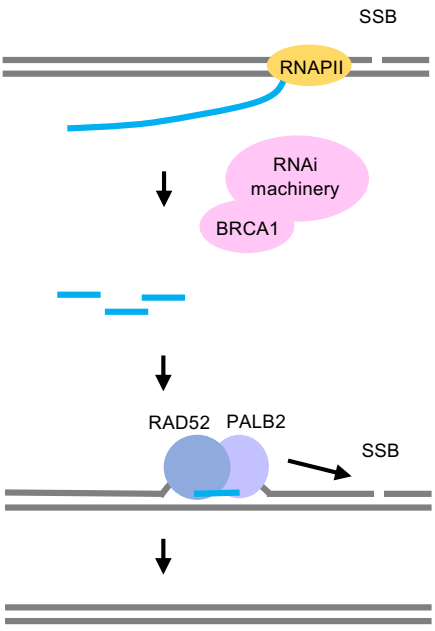

RNA-mediated SSB repair
SSB
RNAPII
RNAi
machinery
BRCA1
RAD52
PALB2
SSB

Fig. 6

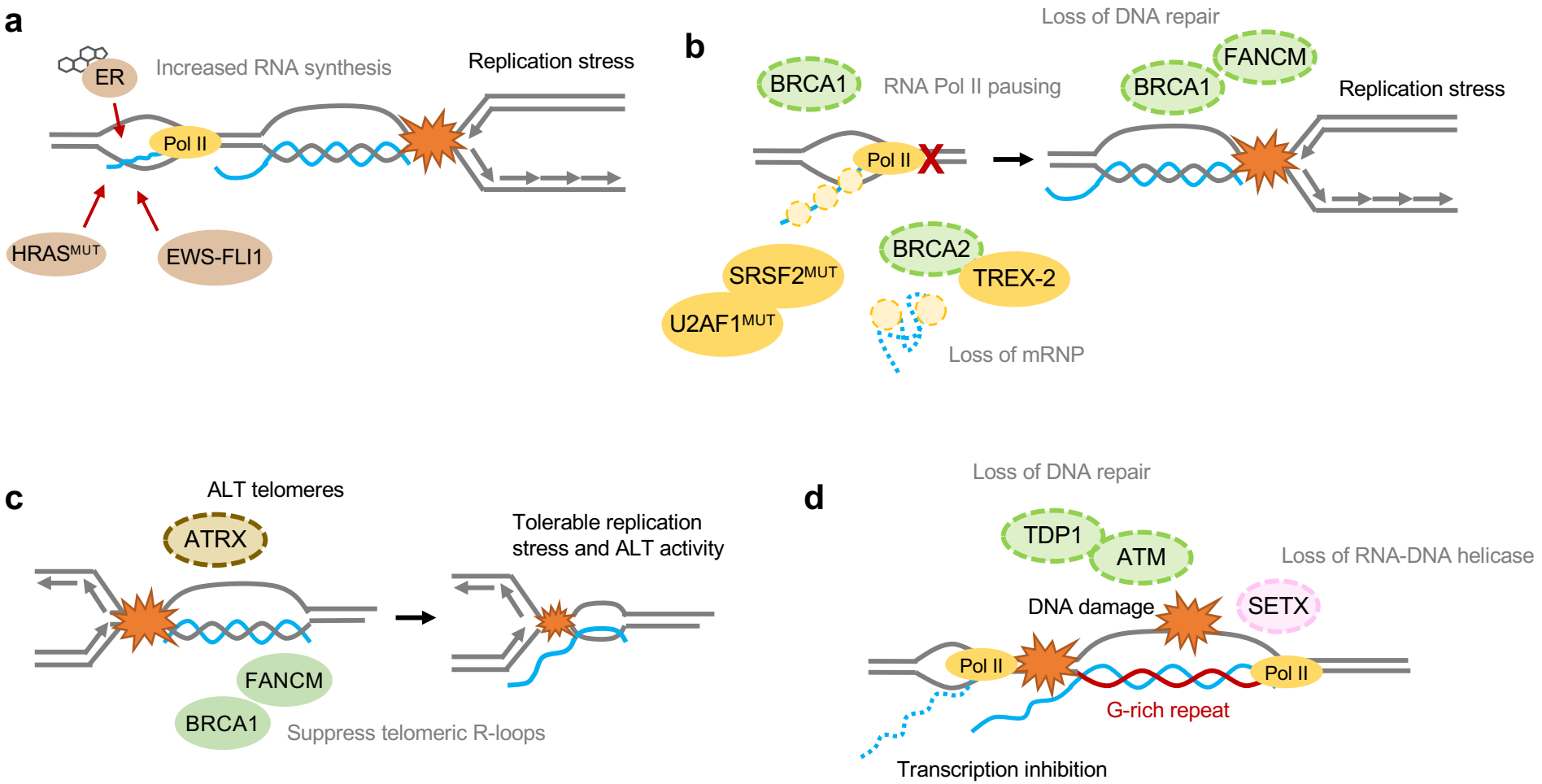

a
ER
Increased RNA synthesis
Replication stress
Pol II
HRASMUT
EWS-FLI1
b
Loss of DNA repair
BRCA1
RNA Pol II pausing
BRCA1
FANCM
Replication stress
Pol II
SRSF2MUT
U2AF1MUT
BRCA2
TREX-2
Loss of mRNP
c
ALT telomeres
ATRX
Tolerable replication stress and ALT activity
FANCM
BRCA1
Suppress telomeric R-loops
d
Loss of DNA repair
TDP1
ATM
Loss of RNA-DNA helicase
DNA damage
SETX
Pol II
Pol II
G-rich repeat
Transcription inhibition

Fig. 7

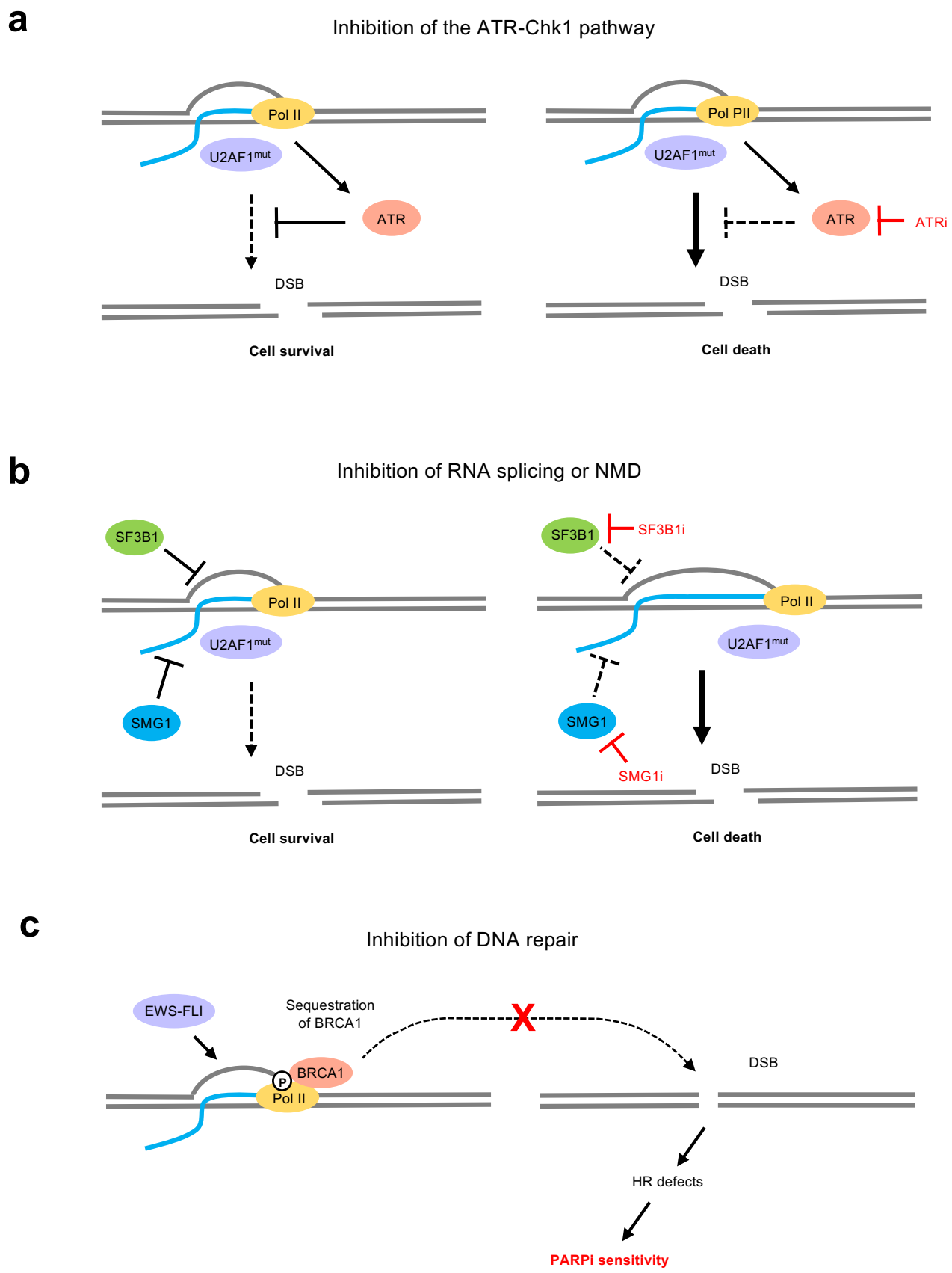
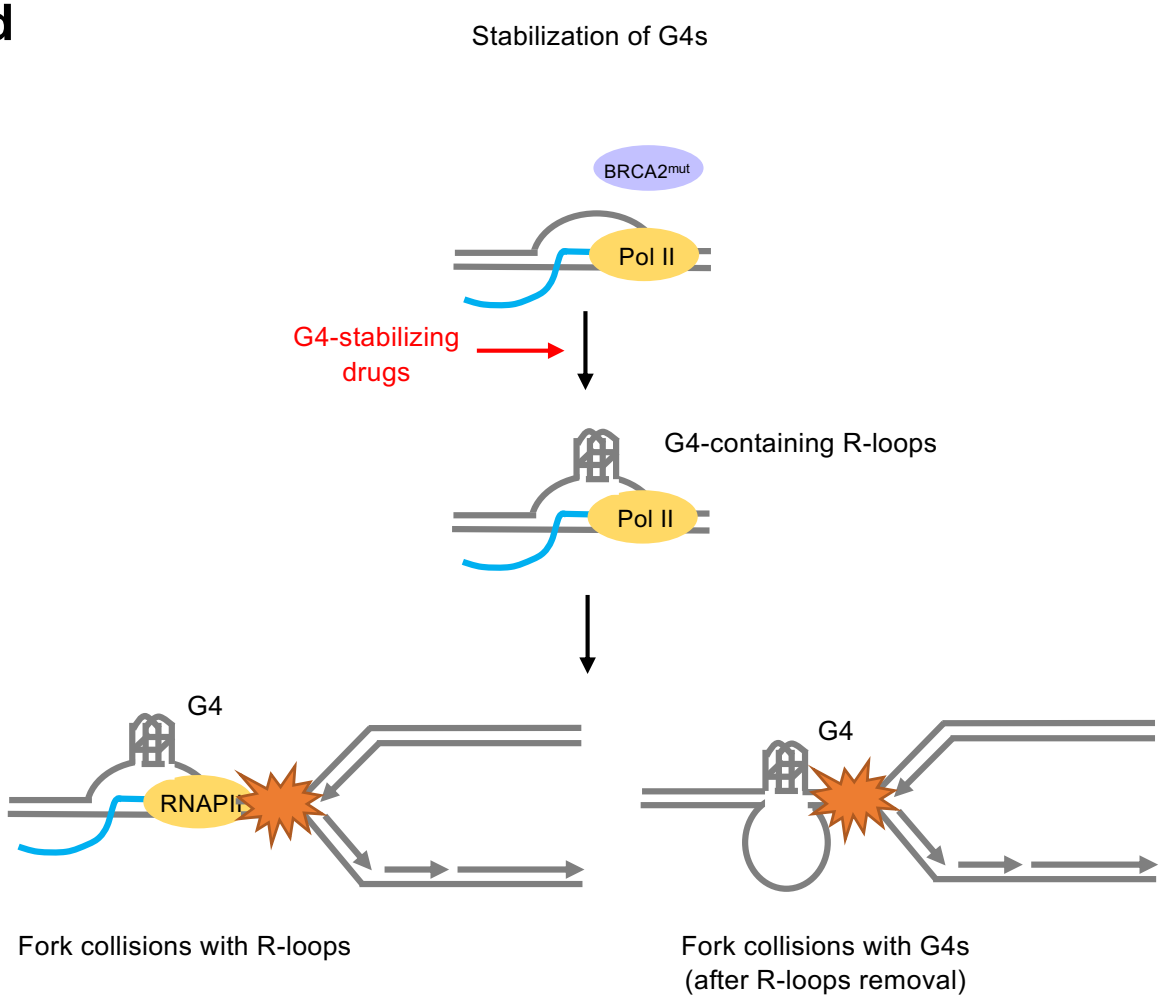

a
Inhibition of the ATR-Chk1 pathway
Pol II
U2AF1mut
ATR
DSB
Cell survival
Pol PII
U2AF1mut
ATR
ATRi
DSB
Cell death
b
Inhibition of RNA splicing or NMD
SF3B1
Pol II
U2AF1mut
SMG1
DSB
Cell survival
SF3B1
SF3B1i
Pol II
U2AF1mut
SMG1
SMG1i
DSB
Cell death
c
Inhibition of DNA repair
EWS-FLI
Sequestration
of BRCA1
BRCA1
P
Pol II
DSB
HR defects
PARPi sensitivity

Fig. 7

d


Stabilization of G4s
BRCA2mut
Pol II
G4-stabilizing drugs
G4-containing R-loops
Pol II
G4
RNAPII
G4
Fork collisions with R-loops
Fork collisions with G4s
(after R-loops removal)