Regulation and function of R-loops at repetitive elements

Alice Gambelli [1], Alessandro Ferrando [1], Chiara Boncristiani, Stefan Schoeftner*

Dipartimento di Scienze della Vita, Università degli Studi di Trieste, Via E. Weiss 2, 34127, Trieste, Italy

A B S T R A C T

R-loops are atypical, three-stranded nucleic acid structures that contain a stretch of RNA:DNA hybrids and an unpaired, single stranded DNA loop. R-loops are physiological relevant and can act as regulators of gene expression, chromatin structure, DNA damage repair and DNA replication. However, unscheduled and persistent R-loops are mutagenic and can mediate replication-transcription conflicts, leading to DNA damage and genome instability if left unchecked. Detailed transcriptome analysis unveiled that 85% of the human genome, including repetitive regions, hold transcriptional activity. This anticipates that R-loops management plays a central role for the regulation and integrity of genomes. This function is expected to have a particular relevance for repetitive sequences that make up to 75% of the human genome. Here, we review the impact of R-loops on the function and stability of repetitive regions such as centromeres, telomeres, rDNA arrays, transposable elements and triplet repeat expansions and discuss their relevance for associated pathological conditions.

* Corresponding author.
E-mail address: sschoeftner@units.it (S. Schoeftner).
[1] These authors contributed equally to the manuscript.

1. R-loops form *in cis* and *in trans*

Hybrids between DNA and RNA are typically formed in a transient manner during transcription by RNA polymerases (RNAPs). Transcription is paralleled by the displacement of the non-template strand, forming a three-stranded nucleic acid structure, commonly called "R-loop" [1]. Negative supercoiling of double-stranded DNA (dsDNA) upstream of the processing RNAP produces an highly energetic conformation that is prone to RNA pairing, favouring R-loop formation, that will subsequently mitigate torsional stress by absorbing the local under-twist and relax nearby domains [2,3]. RNAP movement, RNA binding proteins and the action of topoisomerases assure the efficient resolution of RNA:DNA hybrids leaving a double stranded DNA template that can engage in DNA replication or additional rounds of transcription [4,5].

In the past years an increasing number of loci has been identified that form R-loops to control DNA replication, transcription, translation, chromatin structure, and genomic stability [1,4,6,7] Typically, these structures form *in cis* during transcription [8]; however a panel of RNAs was reported to invade into complementary DNA regions *in trans* [9–15]. Sequence regions that are prone to form RNA:DNA hybrids are typically G-rich or G/A-rich [16,17]. Consistent with this, repetitive regions featuring GC- and AT-skew, G-richness and tandem repeats represent privileged sites for novel R-loop formation [18–24]. Recent work unveiled that R-loop forming regions are composed of three distinct parts: an R-loop initiation zone (RIZ) that needs G-clusters, a sequence independent linker and R-loop elongation zone (REZ), composed by high density in G bases [25]. Application of these features allowed the development of an R-loop prediction tool that led to the identification of R-loop forming sequences in the HOTTIP lncRNA that control the formation of topologically associated domains (TAD) in acute myeloid leukemia (AML) [26,27]. In addition, the presence of G/C stretches on the displaced non-template strand can result in the formation stacks of G-quartets, so-called G-quadruplexes, in which four guanine bases form Hoogsteen base pairs in a planar ring configuration. These atypical DNA secondary structures are reported to promote R-loop formation and interfere with DNA replication causing genetic or epigenetic instability [28–30].

2. Physiological relevance of programmed R-loops

R-loops have been reported to control central physiological processes across different species [8,18,20,28,31–34]. Classic examples comprise R-loop driven immunoglobulin class switch recombination (CSR) that promotes antibody isotype diversity, as well as the formation of RNA primers for the replication of the bacteriophage T4 DNA, the ColE1 plasmid, or mitochondrial DNA [35–40]. Research in the past years has demonstrated that R-loops act as potent regulators of gene expression [41,42]. R-loops have been found to be enriched at CpG islands at gene promoters of highly transcribed genes, protecting them from the access of DNA methyltransferase enzymes (DNMTs) and supporting the recruitment of ten-eleven translocation (TET) demethylase enzymes, thus promoting a permissive environment for transcriptional activation [18,43]. In addition, antisense RNA mediated formation of R-loops at promoters was shown to recruit transcription factors such as NF-κB or recruiting chromatin modifiers including the mixed-lineage leukemia (MLL) H3K4 methyltransferase complex that regulates GATA3 expression [14,44]. R-loops were found to stall RNAPII at transcriptional termination sites by inducing an unfavourable DNA torsions, resulting in polymerase dissociation [19]. R-loops have also been functionally linked to chromatin compaction. Repeat containing transcripts such as telomeric repeat-containing RNAs (TERRA) or pericentric RNAs that engage in R-loop formation were shown to interact with multiple chromatin modifying activates, such as SUV39H1 or PRC2, mediating chromatin compaction [45–48]. Recently, RNA:DNA hybrids were also shown to guide DNA damage repair at double strand breaks [49].

3. Un-programmed R-loops drive replication stress and genomic instability

The presence of persistent, unscheduled R-loops represent a risk for genome instability [50–52]. R-loops are I) hotspots for mutations driven by reactive oxygen species, nucleases, and other agents that target the displaced single-stranded DNA loop, II) can impede replication fork progression, III) cause transcription-replication conflicts, leading to replication fork stalling or collapse or IV) interfere with DNA repair by physically blocking the access of repair enzymes to DNA lesions [51,53]. In line with this, R-loops that threat genomic stability can be targeted by multiple DNA damage repair pathways. The resolution of R-loops is commonly linked to the activation of classic DNA damage signalling pathways (for review [4,54,55]). R-loops are physiologically sensed by the single strand DNA binding protein RPA32 that after phosphorylation, recruits a panel of sensor proteins including the ATR interacting protein (ATRIP), the heterotrimeric ring-shaped 9-1-1 complex (RAD9-RAD1-HUS1 and the ATR kinase. Complex formation leads to activation of ATR, signal transduction and the activation of DNA damage response that is thought to lead to the recruitment of R-loop resolution machineries. In line with this model, inhibition of ATR or the activation of DDR impairs the recruitment of the RNA helicase DDX19 and Senataxin to R-loops [56,57]. Remarkably, R-loop and ATR act at centromeres of metaphase centromeres, promoting Aurora kinase B-mediated phosphorylation of histone H3 Serine 10 (H3S10P) to promote chromatin condensation and assure faithful chromosome segregation [58,59].

Additional pioneering work showed that un-resolved R-loops can be processed by the Xeroderma pigmentosum group F and G (XPF and XPG) endonucleases that are involved in transcription-coupled nucleotide excision repair (TC-NER) [60]. The activation-induced deaminase (AID) can mediate cytosine to uracil deamination in the displaced single stranded DNA loop that is processed by the uracil DNA glycosylase, leading to abasic sites and DNA break formation, as demonstrated during immunoglobin class switching [35]. Collision between R-loops and the replication fork was shown to represent a major risk to genome stability. Co-directional collisions between replisome and the RNA synthesis machinery and R-loops activate Ataxia-telangiectasia mutated (ATM) and results in R-loop unwinding by the replicative helicase. In contrast, head-on collisions activate Ataxia telangiectasia and Rad3-related (ATR) via the structure specific MUS81 endonuclease, leading to the stalling of the replication fork. This subsequently facilitates replication fork repair by I) reducing the frequency of transcription replication conflicts, II) mediating a G2/M checkpoint arrest and III) replication fork recovery via the homologous recombination pathway [4,32,61]. This pathway is highly similar to break induced

replication (BIR) that was shown to elongate telomeres in an R-loop dependent manner via a recombination dependent pathway in telomere negative cancer cells [62]. R-loop driven DNA damage that cannot be repaired in a correct manner leads to the accumulation of mutations and genomic instability [63]. Accordingly, R-loops have been linked to functional alterations of oncogenes, tumor suppressor genes and DNA repeat instability with particular relevance for centromeres, telomeres, ribosomal DNA (rDNA) arrays, but also transposable elements and triple repeat expansions [50,64,65].

4. R-loop management factors with impact on genome instability

In the past years, targeted experiments and proteomics based approaches revealed a series of proteins that can bind and suppress R-loops, limiting the risk of genomic instability [42]. Here we present a selection of R-loop resolving machineries ranging from ribonucleases to histone chaperons that are relevant for R-loop management at repetitive elements. Ribonuclease H (RNase H) family proteins resolve R-loops by degrading the RNA component of the RNA:DNA hybrid, allowing the re-annealing of complementary DNA strands [66,67]. The ssDNA sensing replication protein A (RPA) was shown to bind to the exposed ssDNA loop and to recruit RNase H1, suppressing un-programmed R-loops in a vast variety of sequence contexts [68–70]. In contrast to the monomeric RNase H1 [71], RNase H2 is present as a trimeric complex formed by RNase H2A, RNase H2B and RNase H2C when resolving R-loops [72–75]. DNA helicases have a fundamental role in resolving R-loops by unwinding the RNA moiety trapped in RNA:DNA hybrids [74]. The human RNA helicase Senataxin (SETX), and its *S. cerevisiae* counterpart Sen1, translocate along single stranded DNA and RNA, have RNA:DNA unwinding activity and support transcriptional termination, DNA damage repair and replication fork progression [49,76–78]. Other helicases such as the DHX9 RNA helicase, the Bloom's syndrome helicase (BLM), Aquarius (AQR), FANCM, and the Werner syndrome helicase (WRN) have been reported to hold unwinding activity towards R-loops, thus suppressing genome instability [60,79–85]. The resolution of negative supercoils by type I Topoisomerases was found to be fundamental for R-loop suppression from bacteria to humans [7,86,87]. Topoisomerase 1 (TOP1) is an evolutionarily conserved factor that resolves negative supercoils behind the progressing RNA polymerases, thus suppressing the formation of R-loops [88–90]. Accordingly, loss or inactivation of TOP1 leads to increase R-loop levels, replication stress and DNA damage [91–93]. In addition to TOP1, increasing evidence indicates that other Topoisomerases such as TOP2 or TOP3B are involved in R-loop resolution [94]. Transcript maturation factors have also been linked to R-loop management. The splicing regulator SRFSF1 was shown to bind to single stranded RNA and suppress R-loop formation during RNAPII dependent transcription [95,96]. The evolutionary conserved THO/TREX complex, involved in transcription, splicing and RNA export, also supresses R-loops to ensure elongation of transcription and antagonize transcription associated recombination [97–102]. Components of the Fanconi anemia (FA) pathway that sense and coordinate the resolution of intra-strand crosslinks play a central role in targeting and resolving R-loops [103–105]. The central tumor suppressor BRCA1 (FANCS) has been found to bind to R-loops, recruiting SETX to resolve RNA:DNA hybrids [106]. BRCA2 (FANCD1) binds the branched structure of R-loops, facilitating the recruitment of RNase H1 and SETX [107]. Loss of FANCI, FANCD2, or FANCA causes a dramatic increase in R-loop levels and subsequent DNA breaks in various cell models cells [104,108]. In particular, the monoubiquitinated FANCI/D2 complex binds to R-loops, representing an important step in R-loop response. FANCM was shown to use its ATPase/translocase activity to unwind R-loops and has a critical role in suppressing recombination events at telomere repeats in telomerase negative cells [109]. FANCR (RAD51) was shown to bind to ssDNA and RNAs to catalyse R-loop formation *in trans* and initiates recombination events downstream of R-loop mediated DNA damage [15,110]. R-loops also represent a challenge to chromatin structure as they cannot be arranged in a nucleosome template [8,111]. Accordingly, histone chaperons play an important role in R-loop resolution and the re-establishment of a functional chromatin template. The histone H3 variant H3.3 is deposited by the Death Domain Associated Protein (DAXX) in collaboration with the helicase activity of its binding partner, the ATP-dependent helicase and SWI/SNF family chromatin remodeler Alpha Thalassemia/Mental Retardation Syndrome X-Linked (ATRX) [112]. The DAXX/ATRXX complex was shown to suppress R-loop formation in a replication independent manner at heterochromatic regions, such as telomeres and (peri) centromeres [113–116].

5. R-loops map to repetitive elements from yeast to man

The human genome contains up to 75% of repetitive DNA elements, predominantly represented by transposable elements, (peri)centromeric repeats, telomere repeats and rDNA arrays [117,118]. The finding that the vast amount of repetitive DNA elements are subject to transcription anticipates an important role for R-loop management in the control of genome stability, transcription and chromatin status in these regions. Importantly, an increasing number of studies indicate a relevance for R-loops for human diseases that are characterized by genetic or epigenetic alterations in repetitive DNA elements [119]. DNA repeats can be divided in two major groups. Variable Number of Tandem Repeats (VNTRs) that contain microsatellite (1-10bp) and minisatellite (10-100bp) repeats and Small-Scale Repetitive Elements (SSREs) ranging from 0,1 to 8 kb. Hallmark examples for VNTRs are SatI, SatII and SatII satellite DNAs, ranging from 5bp (unit length of SatIII repeats) to 25 bp (unit length of SatI repeats), pericentromeric β-*satellite* (68–69 bp) and telomere (6 bp) repeats, but also disease related tri-nucleotide repeat expansions [119,120]. The group of SSREs DNA comprise centromeric α-*satellite* (171bp) repeats, γ-*satellite* (220 bp) repeats, transposons (ranging from 100 to 10.000 bp in length, composing 2,8% of human genome), LTR retroelements (0.2–3 kb unit length) that compose the 8,3% of the human genome, and non-LTR retroelements such as LINEs (6–8 kb unit length) and SINEs (0.1–0.4 kb unit length) elements that together make up approximately 34% of the human genome [121,122]. The indicated types of satellite repeats build eukaryotic centromeres and pericentromeres, but can also be located at subtelomeric sites and at interstitial regions [123,124]. In addition, the human genome contains 200–600 copies of ribosomal DNA (rDNA, 43 kb unit length) and 500 interspersed tRNA genes [119,120]. The development of DNA:RNA immunoprecipitation (DRIP) techniques based on the use of a monoclonal antibody (S9.6) with affinity for RNA:DNA hybrids allows the localization of R-loops in preparations of genomic DNA [125–128]. Cross-reactivity of S9.6 antibodies with dsRNA represents a significant problem in R-loop mapping experiments [129]. The application of additional control conditions such as pre-treatment of biological samples with recombinant RNaseH1 or the use of alternative mapping methods such as ChIP using a catalytically inactive RNaseH1 (D210N), helped to overcome this issue [130]. Single molecule R-loop footprinting (SMRF) based on bisulfite-based mapping of the extruded single strand DNA in R-loop represents an attractive, alternative approach for R-loop mapping [131].

Pioneering RNA:DNA hybrid mapping experiments in yeast using ChIP-seq or DRIP followed by hybridization to tiling microarrays

provided first evidence for R-loop formation at tandem or dispersed repetitive elements, including tRNA genes, retrotransposons and telomere repeats, especially in *Top1* mutant yeast strains that are characterized by increased R-loop expression [132,133]. Genome-wide DRIP-seq in combination with the detection of nascent transcripts by Global Run-On sequencing (GRO-seq) allowed to investigate transcriptional activity and R-loop formation at repetitive elements in different model systems [34]. In the U-2 OS human osteosarcoma cell line, R-loops were found to be enriched at telomeric and centromeric regions as well as in simple/low complexity repeats, slightly enriched in satellite repeats and rDNA, but were absent in retroelements. In contrast, in *A. thaliana,* R-loops were highly enriched at all types of transposable elements. Centromere and simple/low complexity repeats showed modest R-loop enrichment. In human cells and *A. thaliana*, an overlap between R-loops and nascent transcript production in the range from 70% to 60% was found, indicative for prevalent formation of R-loop *in cis* [34]. Data from *D. melanogaster* revealed that the pattern of R-loops varies between different status of differentiation. In Drosophila embryos, transposable elements and satellite repeats showed R-loop enrichment when compared to Schneider 2 (S2) cells [34]. In contrast, S2 cells showed R-loop accumulation at simple/low complexity repeats. In line with this, preventing R-loop degradation by overexpression of a catalytically inactive version of RNase H1 resulted in hatching defects in *Drosophila* embryos [134]. These results imply that R-loop expression at repeat regions depends on genome sequence features but also on species, developmental status and biological context, such as the cancer setting in U-2 OS cells [34]. In the next sections we focus our attention on R-loop formation and resolution at major classes of repetitive elements and their impact on pathology in humans.

6. R-Loops at centromere regions control chromosome stability

Eukaryote centromere function is controlled by centromeric and pericentromeric RNAs that can engage in R-loop formation. Human centromeres consist of 171-bp α-*satellite* repeats that are organized in a back-to-back fashion, forming high-order repeats (HORs) [135]. Each HOR is repeated hundreds-to-thousands of times, producing 2-5 Mb-long arrays with chromosome specific characteristics. The flanking pericentromere region is formed by β, γ, I, II, and III satellites tandem repeats (5-200bp), also containing LINE, SINE and retroelements in a more relaxed arrangement [136]. A subset of α-*satellite* monomers contains CENP-B boxes, that are bound by the CENP-B protein, promoting the recruitment of the histone H3 variant centromeric protein A (CENP-A) that is essential for the formation of the kinetochore as attachment point for the mitotic spindle [136–138]. During eukaryotic mitosis, Aurora kinase A (AURKA) promotes centrosome maturation and spindle assemble, while the Aurora kinase B (AURKB) and C (AURKC) regulate chromosome condensation, attachment to kinetochores and the correct alignment of metaphase chromosomes [139]. A key step during mitosis is Aurora B kinase mediated phosphorylation of histone H3 Serine 10 (H3S10P), facilitating chromosome condensation [140–142]. Human centromeres and pericentromeres were demonstrated to give rise to non-coding RNAs (ncRNAs) in an RNAPII dependent manner throughout the cell cycle. Centromeric cenRNAs contain α-satellite repeats, localize to centromeres *in cis* and control centromere functions. Depletion of centromere transcripts using specific shRNAs lead to a reduction of CENP-A and CENP-C localization of 30% and 40%, respectively, arresting the cell cycle at S and G2 phase. These results suggest that loss of non-coding α-satellite transcripts impairs kinetochore assembly and arrests cells prior to mitosis [143,144]. Inhibition of RNAPII using α-amanitin decreases RNA levels at centromeres, impairing CENP-C localization and kinetochore function, provoking the formation of lagging chromosomes [145]. In HeLa cells, anti-sense RNA mediated knockdown of pericentromeric (periCEN) RNAs leads to a reduced AURKB association at centromeres and failure in chromosomes alignment and segregation. On the molecular level, SatI periCEN RNAs were shown to interact and stimulate AURKB to enhance the activity of the chromosome passenger complex (CPC) that controls chromosome segregation and cytokinesis [146]. In addition to central regulators of centromere function, including CENP-A, CENP-B, CENP-C and the CPC, SatI periCEN RNAs were also shown to interact with RBMX (RNA Binding Motif Protein X-Linked), a regulator of cohesion function, and the RNA helicase DHX38 (DEAH box RNA helicase) that controls AURKB localization to mitotic chromosomes [147,148]. In line with data from vertebrate cells, *Drosophila* RNAPII was found to localize at centromeres during G1 and M phase, producing a ncRNA that localizes CENP-A to centromeres [149]. Moreover, transcripts derived from a X-linked periCEN SatIII RNA repeated array mediate the positioning of CENP-A and C to Drosophila centromeres [150]. The high rate of transcription and repetitive nature identify (peri)centromeres as candidate hot-spots for R-loops. In line with this, human R-loop mapping revealed that up to 50% of (peri)centromeric sequences are enriched for R-loops [34]. Remarkably, forced resolution of R-loops by ectopic overexpression of RNase H1 leads to defects in centromere cohesion and chromosome segregation during mitosis, thus pheno-copying the effects of (peri)CEN RNA or RBMX depletion. Moreover, binding of RBMX to centromere R-loops leads to the recruitment of the AURKB kinase ensuring mitotic fidelity [151].

Centromere DNA in *S. cerevisiae* is specified by a 125-bp CEN DNA that contains three centromere-determining DNA Elements (CDEs): the central element (CDEII) and two conserved flanking motifs, the 8-bp CDEI and the 25-bp CDEIII. CEN transcripts in *S. cerevisiae* are detectable in the absence of RNA degradation by the exosome, giving rise to 1.2 kb RNA species containing pericentromere sequences that are required for correct kinetochore activity [152]. R-loops have not been detected at centromeric (CEN) chromatin in wild-type budding yeast [133]. However, a yeast strain lacking the RNA binding protein *Hrp1*, a component of the THO/TREX complex, accumulates centromeric R-loops, and is defined by kinetochore instability and decreased chromosome biorientation [153]. In mitotic cells, R-loops were found to be increased at CEN regions (CDEI, CDEII, CDEIII and pericentric) when compared with interphase cells, especially in *Hrp1* mutant strains. Deletion of *Hrp1* resulted in an accumulation of transcripts derived from CENs leading to defects in kinetochore function and chromosome mis-segregation [58]. On the mechanistic level, R-loops were shown to stimulate the recruitment of AURKB to centromeric and pericentromeric sequences, facilitating the phosphorylation of H3S10. ChIP and DRIP analysis of the pericentromeric regions in chromosome VI demonstrated that R-loop formation coincides with histone H3S10 phosphorylation, promoting chromatin condensation and mitotic fidelity. Accordingly, overexpressing RNase H1 reduces H3S10 phosphorylation, driving genome instability [58]. In human cancer cells, additional line of evidence highlights the relevance of R-loops and ATR for mitotic fidelity. Centromere R-loops recruit the single stranded DNA binding hetero-trimer RPA and ATR during mitosis [59]. In this context, activation of ATR was shown to stimulate AURKB via Chk1, preventing the formation of lagging chromosomes. Moreover, an $ATR^{-/-}$ cell model presented reduced H3S10 phosphorylation, further demonstrating the requirement of ATR for full activation of AURKB at centromeres (Fig. 1). This regulatory pathway suggests that R-loops play an integral part in mitotic control and must by subjected to tight regulation to avoid genome instability. Importantly, a BRCA1-SETX complex was

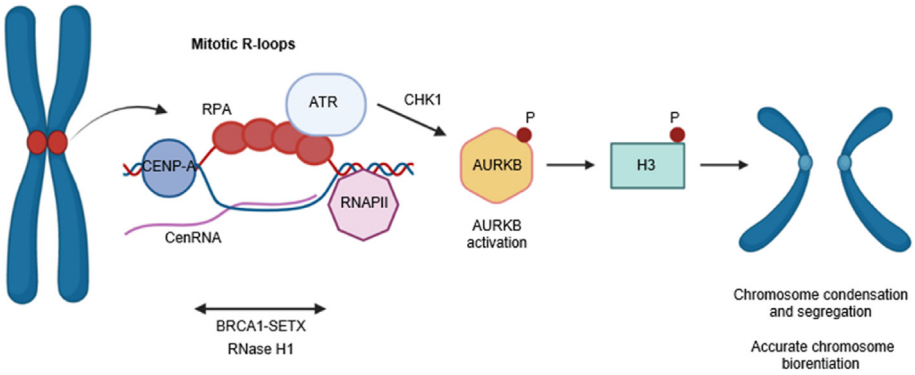

Fig. 1. R-loops control chromosome segregation. Transcription of cenRNAs result in R-loop formation at centromeres. The exposed single stranded DNA loop is bound by RPA and ATR. ATR localization is dependent on AURKA and CENPF (not shown). Activated ATR mediates AURKB activation via CHK1. AURKB is a key player in mitosis and mediates H3Ser10 phosphorylation and the accurate segregation during mitosis. RNase H1 and BRCA1/SETX have a critical role in controlling R-loops at centromeres to assure correct R-loop abundance to assure mitotic fidelity. Figures was generated using BioRender software.

shown to localize to *α-satellite* repeats and ectopic RNase H1 impaired this localization pattern in human cancer cells [106]. Exacerbated centromeric R-loops in BRCA1-SEXT loss-of-function conditions resulted in reduced AURKB kinase localization and H3S10 phosphorylation at centromeres, decreased chromosome cohesion during prometaphase, increased breakage at centromeres, and Rad52 accumulation, indicating a local DNA damage response (Fig. 1) [154,155]. Together this indicates that R-loops at (peri) centromeric regions have conserved role in controlling faithful chromosome segregation during mitosis.

6.1. Connecting (peri)centromeric R-loops with human disease

In tumors, deregulated transcription of satellite-rich (peri) centromeric regions can promote RNA:DNA hybrid formation, contributing to genome instability [156,157]. Interestingly, ectopic expression of human or mouse satellite RNAs using a lentiviral expression vector triggered R-loop mediated replication stress and genomic instability that was able to induce breast cancer formation in experimental mice [157]. This indicates that interfering with the correct expression of pericentric repeats can have oncogenic function. This is in line with the increased expression of pericentric repeats in human cancer tissues [158]. Altered expression of (peri) centric repeats has also been linked with genetic disease. In particular, immunodeficiency, centromeric instability, and Facial abnormalities (ICF) syndrome type I patients carry mutations in the de-novo DNA methyltransferase DNMT3b, ZBTB24, CDCA7 or HELLS genes and show increased R-loop levels at (peri)centric, telomeric and rDNA repeats [159,160]. In patient derived cells, R-loop formation at (peri)centric repeats leads to cleavage by XFG and XPF, DNA damage and the activation of the NHEJ pathway [160,161].

7. R-loops at telomere repeats

Telomeres are heterochromatic nucleoprotein structures that ensure genome stability by protecting eukaryotic chromosome ends from end-to-end fusions and uncontrolled recombination events [162,163]. Vertebrate telomeres are composed of 5′-TTAGGG-3′ microsatellite tandem repeats that are flanked by upstream located subtelomeric regions that contain different arrangements of heterogeneous sequence repetitions [164]. The telomere 3′overhang is extended by the telomerase reverse transcriptase but can also invade double stranded telomeric repeats forming the T-loop that protects chromosome ends from eliciting a DNA damage response [165]. Telomeres are bound by the Shelterin protein complex that ensures telomere function and telomere length homeostasis [166,167]. Constitutive heterochromatin protects telomeres from un-licensed recombination, rearrangements and controls repeat length [168,169]. In vertebrates, subtelomeric promoters drive the expression of a long, $(UUAGGG)_n$ telomere tandem repeat containing non-coding RNA (TERRA) [170,171]. TERRA was initially described as a chromatin associated RNA that acts as recruitment factor for multiple proteins involved in chromatin regulation, replication, telomere protection and telomere length homeostasis [172–175]. Importantly, the G-rich content of TERRA transcripts promotes the formation of R-loops in *cis* or in *trans* at eukaryotic telomeres [45,176–179]. In line with this, genome wide mapping analysis based on DRIP-seq combined to GRO-seq revealed that telomeres are characterized by a remarkable high R-loop enrichment [34]. R-loops at telomeres are kept under check by redundant pathways to reduce the risk for replication-transcription conflicts, DNA breaks and uncontrolled homologous recombination that may drive genomic instability [178,180]. Remarkably, telomere R-loops were demonstrated to fuel recombinogenic substrates that engage in homologous homology-directed repair (HDR) to maintain telomere repeats. Deletion of RNase H1 or the 5′–3′ riboexonuclease Rat1 in *S. cerevisae* provoked the accumulation of R-loops at short telomeres, triggering the formation of recombinogenic substrates that enable the maintenance of telomeres repeats via HDR, thus counteracting the onset of senescence. Importantly, a similar observation was recently obtained in human cells [176,178,181]. In analogous manner, telomerase negative tumors are characterized by elevated R-loop levels at telomeres to drive the BIR dependent, Alternative Lengthening of Telomeres (ALT) pathway, assuring telomere maintenance and replicative immortality [180,182–185].

In a normal physiological context, multiple mechanisms counteract the deleterious formation of R-loops at telomeres. Eukaryotic RNase H enzymes were demonstrate to act as potent machines that resolve telomere R-loops by digesting the trapped TERRA RNA molecule (Fig. 2a) [176,179]. TopoIIα activity was shown to play a critical role in maintaining telomere integrity by protecting from replicative damage, however a direct link to R-loops remains to be identified [186,187]. A series of RNA binding proteins, originally described as TERRA binding proteins, were shown to act as modulators of telomere R-loops. The RNA binding proteins SFPQ (Splicing Factor Proline And Glutamine Rich) and NONO (Non-POU domain-containing octamer-binding protein) were shown to bind TERRA and suppress R-loop formation at telomeres in ALT and telomerase positive cancer cells. Depletion of SFPQ increases the

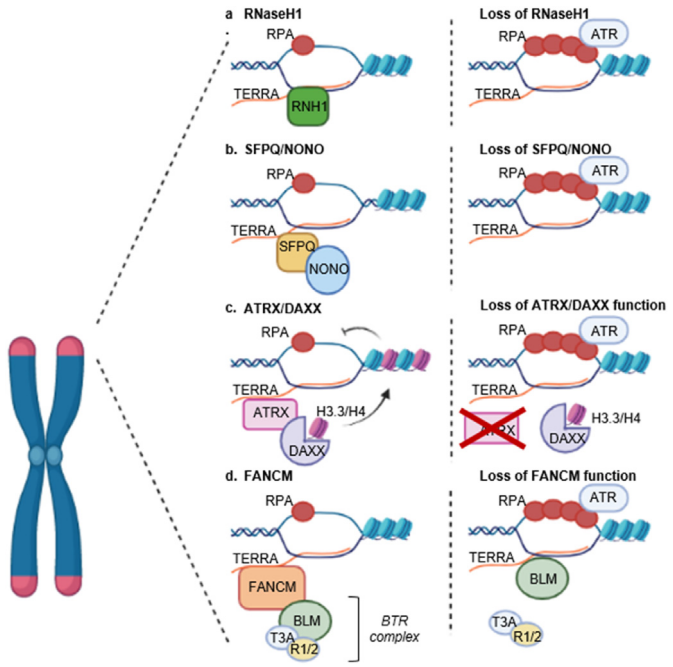

Fig. 2. Different classes of pathways that control the abundance of R-loops at vertebrate telomeres and suppress that activation of ATR. a) RNase H enzymes are general factors that suppress R-loop formation by degrading RNA trapped in R-loop structures (left). In the absence of RNase H activity, R-loops accumulate promoting replication stress and telomere instability. b) NONO and SFPQ form a protein complex and bind TERRA to suppress R-loop formation at telomeres. NONO and SFPQ do not contain enzymatic activity and are thus anticipated to act as a recruitment factor for a yet to define factor (left). Loss of SFPQ and NONO lead to increased R-loop levels, replication stress and activation of the ALT pathway (right). c) ATRX binds TERRA and localizes to telomeres with its binding partner DAXX. ATRX/DAXX heterodimers generate nucleosome arrays containing the histone H3.3 variant in a replication independent manner. This assures the maintenance of local H3K9 methylation (left). Absence of ATRX in ALT cells impairs chromatin structure and allows the accumulation of R-loops at telomeres, promoting ALT. Reconstitution of ATRX expression reverses the ALT phenotype. d) FANCM helicase, binds to R-loops at telomeres recruiting the BTR complex (BLM, in green, TOP3A, in light blue and RMI1/2 in yellow) resulting in DNA:RNA hybrid unwinding (left). In cells depleted for FANCM function, R-loops increase at telomeres, resulting uncontrolled BLM activity that exacerbates the telomere instability in ALT cells, eventually resulting cell death (right). Figures was generated using BioRender software.

number of APBs (ALT-associated Promyelocytic Leukemia nuclear bodies) as well as related recombination events at telomeres in ALT$^+$ cells, while depletion of NONO increased R-loop dependent leading strand telomere fragility [188] (Fig. 2b). However, how NONO and SFPQ alter R-loops abundance is currently not known. Translocated in liposarcoma/fused in sarcoma (TLS/FUS), a multitasking DNA/RNA binding protein implicated in cancer and neurodegenerative diseases was described to regulate human telomeres by binding telomere G-quadruplex DNA and TERRA and appears to be involved R-loop management [189,190]. The TAR DNA-binding protein 43(TDP-43) was shown to interact with TERRA to prevent R-loop mediated replication stress; however, a direct link to telomere function is currently missing [191,192]. The ATP-dependent DNA helicase ATRX (α-thalassemia/mental retardation syndrome X-linked) is reported to bind TERRA and G-quadruplexes, and pairs with the death-domain-associated protein (DAXX) histone chaperone to suppress R-loops at telomeres by deposition of the histone H3.3 variant [112,185,193]. Importantly, frequent mutations of ATRX in telomerase negative tumors were shown to drive the activation of the ALT pathway in human cancer [185]. Moreover, ATRX was shown to be linked to different mechanisms of telomere maintenance, including cell cycle regulation of TERRA and the removal of RPA from telomeres following replication to avoid recombinogenic structures [194] (Fig. 2c). Pioneering studies shed important light into the controlled formation of R-loops and first steps towards homologous recombination at telomeres. The recombinase protein RAD51 has been demonstrated to show increased binding specificity to TERRA when compared to the corresponding ssDNA sequence. RAD51 appears to use this binding specificity to mediate the insertion of TERRA into double-stranded telomere repeat containing DNA, forming R-loops structures, preferentially at short telomeres [15]. This pathway is of particular relevance for telomere maintenance in ALT tumor cells and permit the formation of telomere R-loops in a controlled manner *in trans* using the nucleoplasmatic pool of TERRA molecules [15]. Two other studies demonstrated that the RAD51 associated protein 1 (RAD51AP1) has a central role in promoting the ALT pathway by mediating the R-D-loop switch to promote recombination and generating a chromatin environment that limits the risk of transcription-replication conflicts during recombination events at telomeres [195,196].

Recently, a growing number of studies highlight the role of components of the Fanconi Anemia (FA) pathway in controlling R-loops at telomeres. FA proteins coordinate DNA interstrand cross-link (ICL) repair and rescue of stressed replication forks [197]. BRCA1 was found to directly interact with TERRA and support R-loop resolution, presumably in collaboration with SETX and the 5′-3′ exoribonuclase XRN2, in order to prevent replication stress and telomere aberrations [198–201]. In addition, BRCA2 deletion was shown to cause TERRA hyperexpression and development of an ALT phenotype in telomerase positive colon cancer cells, anticipating a role for BRCA2 in R-loop homeostasis at telomeres [202]. In line with the fact that BRCA1 and BRCA2 represent central components of the Fanconi Anemia (FA) pathway, additional FA proteins were shown to be liked to R-loop homeostasis at telomeres in vertebrate cells. FANCM, FANCA, and FANCL allows FANCD2 recruitment at telomeres, regulating its monoubiquitination status, thus resulting in R-loops suppression [82,203–207]. Additionally, the FANCD2-FANCI complex was also found to bind and resolve R-loops [208]. FANCM appears to have particular relevance for R-loop resolution at ALT telomeres. Loss of FANCM substantially increases R-loop levels and exacerbates downstream features of the ALT pathway, leading to telomere dysfunction, DNA damage and eventually cell death [82,109,209]. Mechanistically, FANCM directly unwinds R-loops at telomeres, mediating the correct recruitment of the BTR complex consisting of BLM, TOP3A and RMI. In FANCM loss of function cells the BTR complex is disassembled and telomeric R-loop are only bound by BLM, that is not able to properly solve them(Fig. 2d) [209].

7.1. Connecting telomere R-loops with human diseases

In telomerase negative tumors, best exemplified by osteosarcoma or pancreatic neuroendocrine tumors (PanNet), DAXX, ATRX and H3.3 are frequently mutated, resulting in the activation of the ALT pathway [116,210–214]. Similar features can be observed in adult lower-grade gliomas, paediatric glioblastoma multiforme, paediatric adrenocortical carcinoma, and neuroblastoma [210]. DAXX and ATRX function as histone chaperon complex and have a major role in depositing H3.3 at telomeres and other repeat elements, ensuring a repressive chromatin environment enriched for H3K9me3 [215]. Accordingly, loss of function of ATRX, DAXX or H3.3 lead to reduced H3K9me3, providing a rational for increased R-loops levels at telomeres [116,216,217]. This results in the production of recombinogenic telomeres that engage in BIR to perform ALT, compensating the absence of telomerase [114,210]. Notably, re-expression of wild-type ATRX or DAXX, suppresses telomere R-

loops and the ALT phenotype, demonstrating a direct connection to R-loop homeostasis [218,219]. Considering the high R-loops levels in ALT tumors, targeting components of central R-loop resolution machinery such as FA pathway or generating a complete null situation of ATRX/DAXX function may exacerbate genomic instability, thus representing a promising therapeutic strategy. In addition to telomerase negative tumors, a cancer relevant connection to telomere R-loops was reported for lung adenocarcinoma. BRCA1 can directly interact with TERRA RNA; a mutant BRCA1 version (R506S) incapable to interact with TERRA, increases telomere R-loops formation, provoking replication stress and alterations at telomeres [199]. In accordance with centromere instability in ICF syndrome type I, patient derived cells also display severe subtelomeric hypomethylation, that allows the expression of elevated levels of TERRA transcripts from DNA methylation sensitive promoters. This results in R-loop formation and accelerated telomere shortening that leads to premature senescence in ICF1 fibroblasts [220–222].

8. A role for R-loops at rDNA arrays and ribosome biogenesis

In humans, rDNA arrays on acrocentric chromosomes 13, 14, 15, 21 and 22 localize to nucleoli giving rise to a RNAPI transcribed 47S precursor that is subsequently processed to the 18S and 28S subunits [223]. In contrast, 5S rDNA arrays that are located on chromosome 1 encode the RNAPIII transcribed 5S rRNA that is transported to the nucleolus for ribosome assembly. In humans, 47S precursor rRNAs are transcribed from approximately 200-600 rDNA repeats that are organized in a head-to-tail organization, separated by 31 kb intergenic spacer regions (IGS) [224,225]. Nucleoli are nuclear membrane-less structures that form via liquid-liquid phase separation at the exit of mitosis and host the process of rRNA biogenesis [226]. They are composed of the granular component (GC), the dense fibrillar component (DFC), and the fibrillar center (FC) [227]. rRNA transcription is driven by RNAPI and the upstream binding factor (UBF), concentrated at the interface between the FC and DFC [228]. Precursor transcripts are then released into the DFC to allow rRNA processing and assembly [229,230].

The high GC-content and repetitive nature that is paired with high transcription rate, render rDNA arrays prone to form R-loops and G-quadruplex structures that can ultimately lead to replication stress and double-stranded breaks (DSBs). In this context, rDNA copies can engage in HDR provoking rDNA copy number alterations [231]. Remarkably, loss of rDNA copies was linked to cellular aging and senescence in yeast [232]. Genomics analyses revealed that rDNA is enriched for R-loops, however only 25% rDNA R-loops have been reported to form by transcription *in cis* [34]. Unresolved R-loops at rDNA repeats drive replication stress and DNA damage leading to a nucleolar-DNA damage response (n-DDR) that inhibits RNAPI transcription activity and drives the formation of nucleolar ATR foci [233,234]. This pathway is evidenced under conditions of hypo-osmotic stress that triggers an accumulation of R-loops at rDNA arrays, leading to RPA recruitment, ATR/ATM activation, DNA damage signalling and repression of nucleolar transcription. Notably, altered rRNA levels were rescued by inhibiting ATR or ATM, or by RPA depletion, promoting R-loop accumulation, highlighting the role of RPA and DNA damage signalling kinases in sensing rDNA R-loops [235]. In line with this, absence of RPA impairs the recruitment of RNase H1 and SETX to rDNA, causing R-loop accumulation [236]. Persistent rDNA damage signalling, has been linked with cancer, viral infection and aging, underling the relevance of rDNA array stability [237–239]. Multiple factors, such as ribonucleases, topoisomerases, helicases, chromatin and transcriptional regulators have been shown to balance R-loops levels at rDNA arrays. In yeast, loss of Top1 leads to an accumulation of R-loops at rDNA arrays, resulting in the accumulation of unprocessed rRNA intermediates and a block of transcriptional activity (Fig. 3) [240]. Accordingly, treatment of vertebrate cells with camptothecin, a topoisomerase I inhibitor, leads to DNA breaks which present Top1 trapped to DNA. These Topo1-linked DNA breaks were found concentrated in highly transcribed rDNA regions, indicating nucleolar localization of Top1 during ongoing rRNA synthesis [241]. Loss of Top1 blocks pre-rRNA synthesis and is linked to concomitant increase in R-loop formation. This effect was exacerbated in Top1/RNase H1 double mutant, concomitant with increased ncRNAs from IGSs [240]. The PHD Finger Protein 6 (PHF6) is reported to act as a repressor of ribosomal RNA transcription by limiting UBF function. Loss of PHF6 increases rRNA transcription, R-loop formation and appearance of DNA damage markers. These phenotypes were rescued by ectopic expression of RNase H1 or knock-down of UBF, underlining that rRNA transcription rate impacts on local R-loop formation and rDNA array stability [242]. RNA helicases such as SETX and the DEAD-box helicase 21 (DDX21) were shown to resolve R-loops at rDNA arrays. SETX was found highly enriched at human IGSs. Upon loss of SETX function, proper nucleoli formation was disrupted, suggesting that physiological levels of R-loops have an impact on nucleolus biogenesis [243]. Nucleolar Sirtuin 7 (SIRT7), a NAD-dependent protein-lysine deacetylase and the DEAD-box helicase DDX21 work together to resolve R loops at rDNA arrays. Previous studies have shown that DDX21 and the NAD-dependent protein-lysine deacetylase 7 (SIRT7) are associated with RNAPI to control rDNA transcription [244,245]. On the molecular level, SIRT7 was shown to deacetylate DDX21, thus increasing its unwinding activity towards R-loops [246]. In the absence of this step, R-loops accumulate, transcription by RNAPI is reduced and DNA damage is accumulating at rDNA genes. Remarkably, ectopic expression of RNase H1 partially rescued rDNA defects, indicating an involvement of R-loops in observed rDNA array defects [246] (Fig. 3).

The DEAD-Box Helicase 47 (DDX47) was shown to directly bind and resolve R-loops in rDNA arrays, impacting on rRNA transcription and alterations in nucleoli area size [247]. DDX47 overexpression was able to partially suppress R-loop accumulation induced by loss of other general suppressors of R-loops, such as DDX23, SETX and FANCD2 [247]. Interestingly, a recent study reports that DDX47 is recruited by FANCD2 to resolve R-loops during mild replication stress [248]. Additional evidence for an involvement of the FA pathway at R-loops in rDNA arrays comes from a study focusing on BRCA1. In particular, BRCA1 was shown to promote antisense-rRNA (as-rRNA) production and to promote the annealing of as-rRNA to rRNA in order to minimize R-loop formation. Accordingly, loss of BRCA1 impaired dsRNA formation and triggered R-loop accumulation, DNA breaks and impaired rRNA processing [249]. In line with the role of ATRX and DAXX in suppressing R-loops at telomeres and other repeats, loss of any of these factors resulted in reduced histone H3.3 deposition at rDNA arrays and R-loop accumulation. Importantly, under these conditions, DNA damage and loss of rDNA repeats were observed, underlining the relevance of chromatin structures for R-loop suppression [250]. R-loops at rDNA arrays were recently shown to have a relevant role in the response to physiological stress. In particular, RNAPI dependent transcription of IGS regions was shown to produce sense intergenic ncRNAs (sincRNAs) that can compromise nucleolar condensates and reduce rRNA expression, producing defects in ribosome biogenesis [243]. Recently IGS have been demonstrated to be also transcribed by RNAPII, producing antisense intergenic ncRNAs (asincRNAs) that engage in R-loop formation [243]. Under normal condition, R-loop formation involving asincRNAs results in reduced IGS sense transcription by RNAPI protecting from sincRNA expression. RNase H1 driven removal of asincRNA containing R-

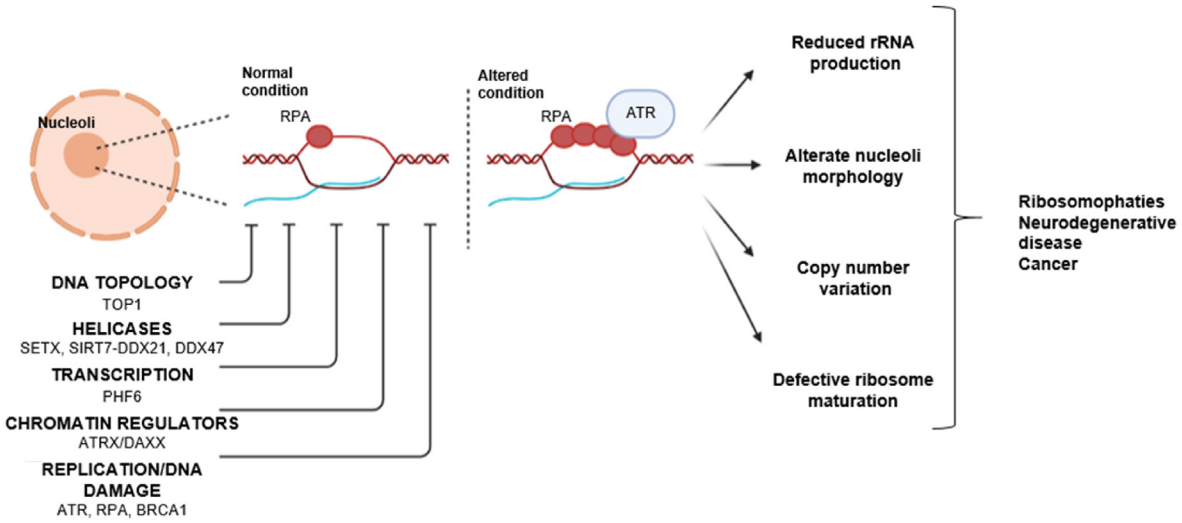

Fig. 3. Multiple pathways control R-loop abundance at rDNA arrays in vertebrates. Mechanisms centring on helicases, rDNA transcription, chromatin structure, the signalling of replication stress and DNA damage keep R-loop levels at rDNA arrays in check. Defects in R-loop surveillance leads to increased DNA:RNA hybrid load that provokes to replication stress. These effects are connected with to an impaired production of rRNAs, defective ribosome maturation, an alteration of nucleolus morphology and rDNA copy number alterations. These features are represented by classic ribosomopathies but are also characteristic for neurodegenerative diseases and cancer. Figure was generated using BioRender software.

loops triggers sincRNA expression and impaired ribosome biogenesis. Thus, RNAPII driven R-loop formation at IGS elements ensures ribosome biogenesis by blocking sincRNA production.

8.1. R-loops in the context of ribosomopathies

Alterations of ribosomal proteins, rDNA processing or ribosome assembly have been linked to diseases that can be classified into congenital or somatic ribosomopathies [251]. Remarkably, altered R-loop regulation in rDNA arrays can be linked to altered ribosome function in a set of human diseases. Bloom's syndrome (BS) is an autosomal recessive disorder that is characterized by mutations in the BLM helicase, leading to persistent pre- and postnatal growth retardation associated with facial erythema, photosensitivity, immunodeficiency and dramatically increased risk of the development of a wide spectrum of cancers at early age [85,252]. BLM suppresses R-loop formation in yeast and vertebrate cells and was shown to interact with multiple factors involved in DNA damage repair [85,253–255]. Remarkably, BS patients were reported to display an 100-fold increase in rDNA repeats coupled with defective DNA repair [256]. In addition, mutations in ataxia-telangiectasia (ATM) have been linked with rDNA array instability in affected patients [256,257]. rDNA instability leads to reduced rDNA transcription, poor ribosome biogenesis and, potentially, defective protein synthesis [256–258]. Interestingly, rDNA repeat number alterations have been linked with other neurological disorders, such as dementia with Lewy bodies, Alzheimer's disease and schizophrenia [259,260]. Members of the FA pathway with known R-loop management function are also connected with altered ribosome biogenesis [261]. Increased instability of rDNA arrays is also a common feature in human cancer. Mutations in the R-loop suppressor ATRX in human sarcoma cells is linked a reduction in rDNA copy number [250]. Alterations in rDNA copy number loss were found to be a general event in human cancer that is connected with hyperactivity of the mTOR pathway that has a tumor promoting effect [262–264]. Together this suggest that R-loops can lead to the alteration of rDNA array organization or ribosome biogenesis and have a relevant contribution to the pathology of disease.

9. R-Loops at transposable elements

Transposable elements (TEs) cover about half of the mammalian genome and have the ability to change their position within a genome [265]. Class I TEs, also known as retrotransposons, move via an RNA intermediate, while class II TEs, also known as DNA transposons, move directly as DNA sequences [266]. TE class I (retrotransposon) contain two major type of interspersed element: LINE (Long Interspersed Nuclear Elements) and SINE (Short Interspersed Nuclear Elements) that account for 20% and 13% of the human genome, respectively [267]. To ensure genomic stability, the mobility of transposable elements is suppressed by RNA interference (RNAi), the Piwi-interacting RNA (piRNA) pathway, and epigenetic silencing [268–270]. LINEs and SINEs activity support genetic diversity and evolution by generating new mutations and rearrangements of the genome. However, in the context of defective TE suppression, transposition events can provoke DNA damage, mutations, and genomic instability and cause disease, as reported for hemophilia A, neurological disorders and cancer [271–274]. On top of this, LINEs and SINEs can give rise to non-coding RNAs (lncRNAs) that interfere with a wide range of cellular processes, including gene regulation and chromatin organization [275].

Recently, transcriptional activation of TEs was found to coincide with R-loop enrichment [34]. In particular, *A. thaliana* and *D. melanogaster* TEs show significant levels of R-loops, and variable pattern related to different developmental status. In humans, TEs are characterized by lower R-loop enrichment [34]. In the genetic context of impaired RNA:DNA hybrid resolution mediated by lacking of RNase H1, RNase H2 or Top1, increased R-loop levels were detected at *S. cerevisiae* TY1 retrotransposons that finally promoted transposition frequency [132]. In *S. pombe,* R-loops in retroelements were linked with chromatin regulation. In particular, RNAs from retroelements located in proximity to centromeres engage in R-loop formation recruiting the RNA-induced transcriptional silencing (RITS) complex, leading to H3K9me3 and formation of constitutive heterochromatin. Ectopic expression of the RNase H1 enzyme leads to loss of heterochromatin formation and de-repression of retroelement transcripts, underlining a link to R-loop homeostasis [276,277].

In maize, the retrotransposon derived CRM1 RNA was shown to regulate centromeres via R-loop formation. In particular, back-splicing results in the formation of a circular CRM1 RNA that binds centromere repeats via R-loops to promote CENP-H3 recruitment and chromatin loop formation, as validated by a 3C mapping assay [278–280].

To this end, direct functional evidence for R-loops related to transposable elements in vertebrates is limited. However, factors involved in R-loop resolution, such as ATRX, DAXX, BRCA1 and FANCD2 were recently demonstrated to control the activity of transposable elements [217,281,282]. Interestingly, in Aicardi-Goutières syndrome (AGS) a link between R-loops and transposable elements was established (see below).

9.1. Connecting R-loops and transposable elements in human disease

A connection between R-loops, transposable elements and disease has been unravelled in the patients with Aicardi-Goutières syndrome (AGS). AGS affects the immune system, brain and skin development and is caused by mutations in TREX1, SAMHD1, ADAR1, or RNase H2 genes. Pathology is commonly linked with the accumulation of R-loops, replication stress and the production of cytoplasmatic DNA, leading to the activation of the cGAS-STING pathway and inflammatory response [32,283]. Involved genes have a common function in suppressing retroelements and R-loops were mapped to LINE-1 and LTR containing sequences in patient derived cell models [284–288]. As expected, cytosolic LINE-1 DNA was found in RNase H2 and TRX loss of function models [289,290]. Of notice, RNase H2 digests the RNA component after reverse transcription of LINE-1 RNA, ensuring retroelement transposition. Thus, an accumulation of LINE-1 RNA:DNA hybrids may act as an activator of the cGAS/STING pathway in patient derived, RNase H2 mutant cells [291,292]. New R-loop mapping studies also reported the accumulation of co-transcriptional R-loops, occurring mainly at short and intronless genes that drive an inflammatory response, suggesting that multiple types of sequences can increase cytoplasmic DNA content via R-loop formation in Aicardi-Goutières syndrome [293].

The activation of retroelements has been reported in a wide spectrum of diseases including neurodegenerative and neoplastic disease [294,295]. R-loop regulators such as BRCA1, FA pathway components, ATRX and DAXX have a clear connection to oncogenic pathways and, at the same time, control retroelement stability [115,296,297]. Investigating R-loop mediated alterations in transposable element activity may provide new insights into genome instability and immunological aspects of these diseases.

10. A role for R-loops in triplet repeat expansions

Separation of DNA strands containing trinucleotide tandem repeats during replication, transcription, recombination, or supercoiling followed by imperfect re-annealing provokes the formation of slipped DNA or hairpin loops in one or both DNA strands [298]. Downstream DNA repair can lead to an increase in triplet repeat numbers that are responsible for the formation of 40 inherited neurological disorders. More recently, also tetra-, penta-, hexa- and even dodeca-nucleotide repeat expansions were linked to genetic diseases [299]. Repeat expansion is dynamic and can increase in successive generation resulting in clinical anticipation of the disease phenotype [299,300]. Triplet repeat expansion were detected in transcribed regions located at I) the 5′ untranslated regions (fragile X syndrome), II) the 3′-UTRs (Huntington's disease); III) introns (Friedreich's ataxia); IV) promoter regions and V) coding region, (Amyotrophic lateral sclerosis/frontotemporal dementia (C9orf72), Huntington disease and Friedreich ataxia Fragile X Syndrome) of involved genes [300]. Importantly, R-loop formation has been observed in CTG, CGG, CCCCGG repeats, but also at GAA repeats in prokaryotic disease model systems [301–304]. This suggests that RNA:DNA hybrid formation at trinucleotide expansions can be linked to the pathomechanism of repeat expansion diseases. In fact, R-loops have been demonstrated to promote the slip out of single stranded DNA that fold in DNA hairpin, G-quadruplex or triple-stranded structures that represent targets for DNA mismatch repair proteins. Inefficient excision of the slip-out DNA and inaccurate pairing with the un-nicked strand followed by repair can lead to an expansion or retraction of local repeat sequences [301,305]. In addition, R-loops were proposed to recruit epigenetic silencing complexes that drive local gene silencing [306].

10.1. Connecting R-loops to triple repeat expansion in neurodegenerative syndromes

Friedreich's ataxia (FRDA) is an autosomal-recessive neurodegenerative and cardiac disorder that is linked with an expansion of $(GAA)_n$ trinucleotides, reaching 44–1700 repeats, localized in the first intron of the frataxin (FXN), leading to gene silencing *in cis*. Reduced frataxin protein levels lead to an accumulation of iron in mitochondria and production of free radicals that drive cell damage and death [307]. Stable R-loops at GAA repeats were reported to promote repeat expansion via break-induced replication but also appear to be involved in the recruitment of epigenetic writers that mediate DNA methylation and H3K9me3 at the FXN locus [308]. Ectopic expression of RNase H1 in human FRDA disease model cells is associated with loss of H3K9me3 and reactivation of FXN expression [302,309,310]. Accordingly, depletion of RNase H1 in HEK293 cells significantly increases R-loop formation in trinucleotide tandem repeats. Together, this suggests that endogenous RNase H1 may antagonize R-loop abundance across the FXN gene to suppress DNA slip-out [302]. In human cell cultures, $(GAA)_n$ expansions directly correlate with passage number and depend on DNA mismatch repair (MMR) and transcription [85,252–254]. Together, these features may also explain the production of a fragile site at the FXN allele that is prone to breakage and rearrangements especially in the context of ATM inhibitors [311]. However, to this end, a direct link between R-loops, and the repeat amplifications remains to be identified.

Fragile X syndrome is a neurodevelopmental disorder characterized by intellectual disability and behavioural symptoms such as hyperactivity and anxiety that - in most cases – is linked with more than 200 CGG triplet repeat expansions in the 5′-UTR of the fragile X mental retardation 1 gene (FMR1) leading to gene silencing *in cis* [312–316]. Reversal of repressive DNA methylation is not sufficient to restore FMR1 expression, anticipating the action of redundant pathways of heterochromatinization [316]. DRIP experiments identified R-loops at $(CGG)_n$ repeat expansions in patient derived cells. Importantly, overexpression of RNase H1 was shown to mediate an up-regulation of FMR1 expression *in vivo* [302,303]. Thus, a current model proposes that R-loops act as primary trigger for repression of expanded FMR1 but also contribute to repeat expansion and the manifestation of the characteristic fragility of the X chromosome in patient cells [291,292,317,318].

Huntington disease (HD) is an autosomal dominant, progressive neurodegenerative disorder characterized by choreiform movements, psychiatric problems and dementia, caused by polyglutamine $(CAG)_n$ repeat expansion in exon 1 of the Huntingtin (HTT) gene [319]. DRIP-seq experiments demonstrated the presence of R-loops at *Htt* alleles in neural stem and progenitor cells, suggesting a potential contribution of R-loops to Huntington's disease pathology [320]. R-loops associated with $(CAG)_n$ repeat expansion were

shown to drive repeat instability *in vitro* and *in vivo*; and loss of RNase H1 or RNase H2 in human cells enhanced transcription mediated genomic instability [301]. R-loops have been further suggested to promote the instability of pathogenic repeat sequences in Huntington's disease [301,321–323].

Amyotrophic lateral sclerosis (ALS) and/or frontotemporal dementia (FTD) is linked to $(GGGGCC)_n$ hexanucleotide repeat expansions (HRE) in the chromosome 9 open reading frame 72 (C9orf72) and represent the most common type of ALS and FTD [324]. The clinical symptoms comprise the motoneurodegeneration of frontal lobe, amyotrophy and degeneration of motoneurons producing muscle denervation. Pathogenesis is mediated by a combination of loss of function of the C9orf72 gene products and toxicity of the HRE containing transcripts that can aggregate in RNA foci. RNA G-quadruplex were shown to form RNA foci that sequester the nucleolar protein nucleolin, causing the depletion of the protein from the nucleolus resulting nucleolar stress [325]. Importantly, C9orf72 HRE regions hold intrinsic potential to form R-loops *in vitro* which in turn promotes the formation of stable secondary structures, including hairpins and G-quadruplexes at the unpaired DNA strand [326]. *In vitro* experiments showed that R-loops promote repeat associated instability leading to HRE contraction or expansion [327]. Remarkably, C9orf72 RNA clusters were found to be enriched for the splicing factor SRSF1 that was previously reported to participate in R-loop resolution. This suggests that SRSF1 sequestration might lead to an increase of R-loop abundance at $(GGGGCC)_n$ repeat expansions, thus further promoting repeat instability [328]. Classically, trinucleotide repeats have been linked to neurodegenerative diseases. Recent genome analyses revealed that microsatellite sequence expansion, including trinucleotide motifs also occur frequently in human cancer [329]. However, to this end a possible link to R-loops remains to be identified.

11. Conclusions

R-loop structures were first described in 1987 and linked to the transcriptional activation of the ColE1 origin [330]. Research performed in the past 10 years revealed a wide spectrum of biological functions linked to R-loops in the context of cell physiology. Multiple factors that control the delicate balance between beneficial and detrimental effects of DNA:DNA hybrids were identified. This is of particular relevance for repetitive regions that can engage into uncontrolled homologous recombination after replication stress induced DNA damage. Telomere and (peri)centromere repeats represent hallmark models that revealed the action of multiple layers of R-loop promoting or resolving machineries. Correct levels of R-loops support centromere function and chromosome segregation; aberrant (peri)centric R-loops levels impair mitotic fidelity. On the other hand, telomerase negative cancer cells select for mutations in genes encoding R-loop resolving machineries thus allowing the generation of replication stress and the formation of recombinogenic substrates that fuel the ALT pathway. Exacerbation of this process leads to uncontrolled recombination events that eventually lead to cell death. In addition to these functionally well dissected models for R-loops function, an increasing number of studies link RNA:DNA hybrids to ribosome biogenesis, suppression of transposable element and triplet repeat expansion. Genome wide mapping studies revealed that R-loops represent frequent structures in eukaryotes that cover up to 5% of the human genome [8,20,33]. This suggests that the rate of R-loop production, R-loop detection and resolution needs to be controlled in a precise and eventually sequence specific manner. However, to this end, information on these aspects is very limited. In the past, an increasing number of proteins with and without enzymatic function have been found to regulate R-loops in an apparently redundant manner. For many of these players their precise function during R-loop recognition, R-loop resolution or R-loop formation it is not clear. In addition, pathways that connect these machineries to cell physiology are poorly understood. It is actually interesting to speculate that the activity and eventual site-specific recruitment of R-loop processing machineries may be controlled by defined signaling pathways. This may help highly-proliferative and transcriptional active cancer cells to avoid excessive R-loop mediated replication stress and DNA damage or facilitate the efficient resolution of R-loops in cancer cells that are resistant to drugs that induce R-loop formation. A critical step in future R-loop research will be the development of efficient technologies that allow R-loop mapping in large cohorts of patients with diseases connected to altered R-loop metabolism such as neoplastic disease, ICF, Aicardi-Goutières syndrome and neurodegenerative diseases. This will be instrumental to finally obtain solid, scientific evidence for the relevance of RNA:DNA hybrids in the pathogenesis of disease. Together is will pave the way towards therapeutic approaches targeting R-loops in human disease.

Author contributions

Conceptualization and writing, A.G; A.F; C.B; S.S; production of figures, A.G.; review and editing, S.S. All authors have read and agreed to the published version of the manuscript.

Funding

S.S. is supported by the Fondazione AIRC per la ricerca sul cancro ETS - Investigator Grant 18381.

Declaration of competing interest

The authors declare no conflict of interest. The funders have no role in the design of the study, in the writing, or in the decision to publish the work.

Acknowledgements

We apologize to colleagues whose work could not be included due to space constrains. Figures were generated using BioRender software.

References

[1] D.F. Allison, G.G. Wang, R-loops: formation, function, and relevance to cell stress, Cell Stress 3 (2019) 38–46.
[2] R. Stolz, et al., Interplay between DNA sequence and negative superhelicity drives R-loop structures, Proc. Natl. Acad. Sci. 116 (2019) 6260–6269.
[3] F. Chedin, C.J. Benham, Emerging roles for R-loop structures in the management of topological stress, J. Biol. Chem. 295 (2020) 4684–4695.
[4] E. Petermann, L. Lan, L. Zou, Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids, Nat. Rev. Mol. Cell Biol. 23 (2022) 521–540.
[5] M. Drolet, X. Bi, L.F. Liu, Hypernegative supercoiling of the DNA template during transcription elongation in vitro, J. Biol. Chem. 269 (1994) 2068–2074.
[6] L. Costantino, D. Koshland, The yin and yang of R-loop biology, Curr. Opin. Cell Biol. 34 (2015) 39–45.
[7] M. Drolet, J. Brochu, R-loop-dependent replication and genomic instability in bacteria, DNA Repair 84 (2019) 102693.
[8] L.A. Sanz, et al., Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals, Mol. Cell 63 (2016) 167–178.
[9] K.D. Westover, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: separation of RNA from DNA by RNA polymerase II, Science 303 (2004) 1014–1016.
[10] L. Wahba, S.K. Gore, D. Koshland, The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome

instability, Elife 2 (2013) e00505.
[11] S.C. Cloutier, et al., Regulated Formation of lncRNA-DNA hybrids enables faster transcriptional induction and environmental adaptation, Mol. Cell 61 (2016) 393–404.
[12] F. Ariel, et al., R-loop mediated trans action of the APOLO long noncoding RNA, Mol. Cell 77 (2020) 1055–1065.e4.
[13] F.C. Beckedorff, et al., The intronic long noncoding RNA ANRASSF1 recruits PRC2 to the RASSF1A promoter, reducing the expression of RASSF1A and increasing cell proliferation, PLoS Genet. 9 (2013) e1003705.
[14] R. Boque-Sastre, et al., Head-to-head antisense transcription and R-loop formation promotes transcriptional activation, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 5785–5790.
[15] M. Feretzaki, et al., RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops, Nature 587 (2020) 303–308.
[16] J.L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, G-quadruplexes: The beginning and end of UTRs, Nucleic Acids Res. 36 (2008) 6260–6268.
[17] L. Ratmeyer, R. Vinayak, Y.Y. Zhong, G. Zon, W.D. Wilson, Sequence specific thermodynamic and structural properties for DNA. Cntdot. RNA duplexes, Biochemistry 33 (1994) 5298–5304.
[18] P.A. Ginno, P.L. Lott, H.C. Christensen, I. Korf, F. Chédin, R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters, Mol. Cell (2012), https://doi.org/10.1016/j.molcel.2012.01.017.
[19] P.A. Ginno, Y.W. Lim, P.L. Lott, I. Korf, F. Chédin, GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination, Genome Res. 23 (2013) 1590–1600.
[20] L. Wahba, L. Costantino, F.J. Tan, A. Zimmer, D. Koshland, S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation, Genes Dev. 30 (2016) 1327–1338.
[21] A. De Magis, 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 (2019) 816–825.
[22] K. Rippe, B. Luke, TERRA and the state of the telomere, Nat. Struct. Mol. Biol. 22 (2015) 853–858.
[23] Y. Liu, et al., Genome-wide mapping reveals R-loops associated with centromeric repeats in maize, Genome Res. 31 (2021) 1409–1418.
[24] M. Groh, M.M.P. Lufino, R. Wade-Martins, N. Gromak, R-Loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome, PLoS Genet. 10 (2014).
[25] D. Roy, M.R.G. Lieber, Clustering is important for the initiation of transcription-induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter, Mol. Cell Biol. 29 (2009) 3124–3133.
[26] P. Jenjaroenpun, T. Wongsurawat, S.P. Yenamandra, V.A. Kuznetsov, QmRLFS-finder: a model, web server and stand-alone tool for prediction and analysis of R-loop forming sequences, Nucleic Acids Res. 43 (2015) W527–W534.
[27] H. Luo, et al., HOTTIP-dependent R-loop formation regulates CTCF boundary activity and TAD integrity in leukemia, Mol. Cell 82 (2022) 833–851.e11.
[28] F. Chédin, Nascent connections: R-loops and chromatin patterning, Trends Genet. 32 (2016) 828–838.
[29] G. Miglietta, M. Russo, G. Capranico, G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders, Nucleic Acids Res. 48 (2020) 11942–11957.
[30] D. Yang, G-Quadruplex DNA and RNA, Methods Mol. Biol. 2035 (2019) 1–24.
[31] J. Sollier, K.A. Cimprich, Breaking bad: R-loops and genome integrity, Trends Cell Biol. 25 (2015) 514–522.
[32] M.P. Crossley, M. Bocek, K.A. Cimprich, R-loops as cellular regulators and genomic threats, Mol. Cell 73 (2019) 398–411.
[33] W. Xu, et al., The R-loop is a common chromatin feature of the Arabidopsis genome, Nat. Plants 3 (2017) 704–714.
[34] C. Zeng, M. Onoguchi, M. Hamada, Association analysis of repetitive elements and R-loop formation across species, Mobile DNA 12 (2021) 3.
[35] K. Yu, F. Chedin, C.L. Hsieh, T.E. Wilson, M.R. Lieber, R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells, Nat. Immunol. (2003), https://doi.org/10.1038/ni919.
[36] M. Muramatsu, et al., Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme, Cell 102 (2000) 553–563.
[37] P. Revy, et al., Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2), Cell 102 (2000) 565–575.
[38] K.N. Kreuzer, J.R. Brister, Initiation of bacteriophage T4 DNA replication and replication fork dynamics: a review in the Virology Journal series on bacteriophage T4 and its relatives, Virol. J. 7 (2010) 358.
[39] T. Itoh, J. Tomizawa, Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H, Proc. Natl. Acad. Sci. U.S.A. 77 (1980) 2450–2454.
[40] J.L.O. Pohjoismäki, et al., Mammalian mitochondrial DNA replication intermediates are essentially duplex, but contain extensive tracts of RNA/DNA hybrid, J. Mol. Biol. 397 (2010) 11144–11155.
[41] C. Niehrs, B. Luke, Regulatory R-loops as facilitators of gene expression and genome stability, Nat. Rev. Mol. Cell Biol. 21 (2020) 167–178.
[42] J. Santos-Pereira, A. Aguilera, R loops : New modulators of genome dynamics and function, Nat. Rev. Genet. 16 (2015) 583–597.
[43] K. Arab, et al., GADD45A binds R-loops and recruits TET1 to CpG island promoters, Nat. Genet. 51 (2019) 217–223.
[44] H.R. Gibbons, et al., Divergent lncRNA GATA3-AS1 regulates GATA3 transcription in T-helper 2 cells, Front. Immunol. 9 (2018) 2512.
[45] R.V. Fernandes, M. Feretzaki, J. Lingner, The makings of TERRA R-loops at chromosome ends, Cell Cycle 20 (2021) 1745–1759.
[46] A. Porro, et al., Functional characterization of the TERRA transcriptome at damaged telomeres, Nat. Commun. 5 (2014) 5379.
[47] J.J. Montero, et al., TERRA recruitment of polycomb to telomeres is essential for histone trymethylation marks at telomeric heterochromatin, Nat. Commun. 9 (2018) 1548.
[48] O.V. Camacho, et al., Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation, Elife 6 (2017) 1–29.
[49] S. Cohen, et al., Senataxin resolves RNA: DNA hybrids forming at DNA double-strand breaks to prevent translocations, Nat. Commun. 9 (2018) 533.
[50] J.P. Wells, J. White, P.C.R. Stirling, Loops and their composite cancer connections, Trends in Cancer 5 (2019).
[51] T. García-Muse, A. Aguilera, Transcription-replication conflicts: how they occur and how they are resolved, Nat. Rev. Mol. Cell Biol. 17 (2016) 553–563.
[52] H. Gaillard, T. García-Muse, A. Aguilera, Replication stress and cancer, Nat. Rev. Cancer 15 (2015) 276–289.
[53] S. Kemiha, J. Poli, Y.-L. Lin, A. Lengronne, P. Pasero, Toxic R-loops: cause or consequence of replication stress? DNA Repair 107 (2021) 103199.
[54] C. Rinaldi, P. Pizzul, M.P. Longhese, D. Bonetti, Sensing R-loop-associated DNA damage to safeguard genome stability, Front. Cell Dev. Biol. 8 (2020) 618157.
[55] J.R. Brickner, J.L. Garzon, K.A. Cimprich, Walking a tightrope: the complex balancing act of R-loops in genome stability, Mol. Cell 82 (2022) 2267–2297.
[56] D. Hodroj, et al., An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism, EMBO J. 36 (2017) 1182–1198.
[57] Ö. Yüce, S. C. Senataxin West, Defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response, Mol. Cell Biol. 33 (2013) 406–417.
[58] M. Castellano-Pozo, et al., R loops are linked to histone H3 S10 phosphorylation and chromatin condensation, Mol. Cell (2013), https://doi.org/10.1016/j.molcel.2013.10.006.
[59] L. Kabeche, H.D. Nguyen, R. Buisson, L. Zou, A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation, Science 359 (2018) 108–114.
[60] J. Sollier, et al., Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability, Mol. Cell 56 (2014) 777–785.
[61] C. Rinaldi, P. Pizzul, M.P. Longhese, D. Bonetti, Sensing R-loop-associated DNA damage to safeguard genome stability, Front. Cell Dev. Biol. 8 (2021).
[62] N. Chappidi, et al., Fork cleavage-religation cycle and active transcription mediate replication restart after fork stalling at Co-transcriptional R-loops, Mol. Cell 77 (2020) 528–541.e8.
[63] P.A. Jeggo, L.H. Pearl, A.M. Carr, DNA repair, genome stability and cancer: a historical perspective, Nat. Rev. Cancer 16 (2016) 35–42.
[64] Y. Yao, W. Dai, Genomic instability and cancer, J. Carcinog. Mutagen. 5 (2014).
[65] I. Kovalchuk, Chapter 25 - human diseases associated with genome instability☆☆Original article in the 1st edition was prepared by bruno césar feltes, joice de Faria poloni, kendi nishino miyamoto, and diego bonatto, in: Second E., in: I. Kovalchuk, O.B.T.-G.S. Kovalchuk (Eds.), Translational Epigenetics, vol. 26, Academic Press, 2021, pp. 479–493.
[66] S.M. Cerritelli, R.J. Crouch, The enzymes in eukaryotes, FEBS J. 276 (2009) 1494–1505.
[67] J.D. Amon, D. Koshland, RNase H enables efficient repair of R-loop induced DNA damage, Elife 5 (2016) e20533.
[68] H.D. Nguyen, et al., Functions of replication protein A as a sensor of R loops and a regulator of RNaseH1, Mol. Cell 65 (2017) 832–847.e4.
[69] S. Liu, et al., Distinct roles for DNA-PK, ATM and ATR in RPA phosphorylation and checkpoint activation in response to replication stress, Nucleic Acids Res. 40 (2012) 10780–10794.
[70] A. Maréchal, L. Zou, DNA damage sensing by the ATM and ATR kinases, Cold Spring Harbor Perspect. Biol. 5 (2013) a012716.
[71] A. Lockhart, et al., RNase H1 and H2 are differentially regulated to process RNA-DNA hybrids, Cell Rep. 29 (2019) 2890–2900.e5.
[72] N.M. Shaban, S. Harvey, F.W. Perrino, T. Hollis, The structure of the mammalian RNase H2 complex provides insight into RNA·DNA hybrid processing to prevent immune dysfunction, J. Biol. Chem. 285 (2010) 3617–3624.
[73] D.A. Cornelio, H.N.C. Sedam, J.A. Ferrarezi, N.M.V. Sampaio, J.L. Argueso, Both R-loop removal and ribonucleotide excision repair activities of RNase H2 contribute substantially to chromosome stability, DNA Repair 52 (2017) 110–114.
[74] E. Jankowsky, RNA helicases at work: binding and rearranging, Trends Biochem. Sci. 36 (2011) 19–29.
[75] K. Skourti-Stathaki, N.J. Proudfoot, N. Gromak, Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote xrn2-dependent termination, Mol. Cell 42 (2011) 794–805.
[76] H.E. Mischo, et al., Yeast Sen1 helicase protects the genome from transcription-associated instability, Mol. Cell 41 (2011) 21–32.
[77] M. Groh, L.O. Albulescu, A. Cristini, N. Gromak, Senataxin: genome guardian at the interface of transcription and neurodegeneration, J. Mol. Biol. 429 (2017) 3181–3195.

[78] A. Cristini, M. Groh, M.S. Kristiansen, N. Gromak, RNA/DNA hybrid interactome identifies DXH9 as a molecular player in transcriptional termination and R-loop-associated DNA damage, Cell Rep. 23 (2018) 1891–1905.
[79] A. Cristini, M. Groh, M.S. Kristiansen, N. Gromak, RNA/DNA hybrid interactome identifies DXH9 as a molecular player in transcriptional termination and R-loop-associated DNA damage, Cell Rep. 23 (2018) 1891–1905.
[80] E.Y.C. Chang, et al., RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instabil, J. Cell Biol. 216 (2017) 3991–4005.
[81] V. Marabitti, et al., R-Loop-Associated genomic instability and implication of WRN and WRNIP1, Int. J. Mol. Sci. 23 (2022).
[82] X. Pan, et al., FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops, Sci. Rep. 9 (2019) 19110.
[83] P. Chakraborty, F. Grosse, Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes, DNA Repair 10 (2011) 654–665.
[84] A. Akay, et al., The helicase aquarius/EMB-4 is required to overcome intronic barriers to allow nuclear RNAi pathways to heritably silence transcription, Dev. Cell 42 (2017) 241–255.
[85] E.Y.-C. Chang, et al., RECQ-like helicases Sgs1 and BLM regulate R-loop--associated genome instability, J. Cell Biol. 216 (2017) 3991–4005.
[86] D. Sutormin, et al., Interaction between transcribing RNA polymerase and topoisomerase I prevents R-loop formation in E. coli, Nat. Commun. 13 (2022) 4524.
[87] M. Drolet, H.Y. Wu, L.F. Liu, Roles of DNA topoisomerases in transcription, Adv. Pharmacol. 29A (1994) 135–146.
[88] M. Drolet, et al., Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I, Proc. Natl. Acad. Sci. 92 (1995) 3526–3530.
[89] A. El Hage, S.L. French, A.L. Beyer, D. Tollervey, Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis, Genes Dev. 24 (2010) 1546–1558.
[90] J.A. Smolka, L.A. Sanz, S.R. Hartono, F. Chédin, Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids, J. Cell Biol. 220 (2021) e202004079.
[91] S. Tuduri, et al., Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription, Nat. Cell Biol. 11 (2009) 1315–1324.
[92] A. Ray Chaudhuri, et al., Topoisomerase I poisoning results in PARP-mediated replication fork reversal, Nat. Struct. Mol. Biol. 19 (2012) 417–423.
[93] J. Marinello, et al., Dynamic effects of topoisomerase i inhibition on R-loops and short transcripts at active promoters, PLoS One 11 (2016) 1–18.
[94] P.S. Patel, R. Krishnan, R. Hakem, Emerging roles of DNA topoisomerases in the regulation of R-loops, Mutat. Res. Toxicol. Environ. Mutagen (2022) 876–877, 503450.
[95] O. Anczuków, et al., SRSF1-Regulated alternative splicing in breast cancer, Mol. Cell 60 (2015) 105–117.
[96] S. Paz, A. Ritchie, C. Mauer, M. Caputi, The RNA binding protein SRSF1 is a master switch of gene expression and regulation in the immune system, Cytokine Growth Factor Rev. 57 (2021) 19–26.
[97] J. Katahira, mRNA export and the TREX complex, Biochim. Biophys. Acta - Gene Regul. Mech. 1819 (2012) 507–513.
[98] G.A. Khan, et al., The transcription and export complex THO/TREX contributes to transcription termination in plants, PLoS Genet. 16 (2020) e1008732.
[99] M.S. Domínguez-Sánchez, S. Barroso, B. Gómez-González, R. Luna, A. Aguilera, Genome instability and transcription elongation impairment in human cells depleted of THO/TREX, PLoS Genet. 7 (2011) 1002386.
[100] S. Chávez, et al., A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae, EMBO J. 19 (2000) 5824–5834.
[101] K. Stäßer, et al., TREX is a conserved complex coupling transcription with messenger RNA export, Nature 417 (2002) 304–308.
[102] J. Katahira, Y. Yoneda, Roles of the TREX complex in nuclear export of mRNA, RNA Biol. 6 (2009) 149–152.
[103] Y. Okamoto, J. Hejna, M. Takata, Regulation of R-loops and genome instability in Fanconi anemia, J. Biochem. 165 (2019) 465–470.
[104] M.L. García-Rubio, et al., The fanconi anemia pathway protects genome integrity from R-loops, PLoS Genet. 11 (2015) 1–17.
[105] A. Rodríguez, A. D'Andrea, Fanconi anemia pathway, Curr. Biol. 27 (2017) R986–R988.
[106] E. Hatchi, et al., BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair, Mol. Cell 57 (2015) 636–647.
[107] V. Bhatia, et al., BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2, Nature 511 (2014) 362–365.
[108] Z. Liang, et al., Binding of FANCI-FANCD2 complex to RNA and R-loops stimulates robust FANCD2 monoubiquitination, Cell Rep. 26 (2019) 564–572.e5.
[109] B. Silva, et al., FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops, Nat. Commun. 10 (2019).
[110] M. Thomas, C. Dubacq, E. Rabut, B.S. Lopez, J. Guirouilh-Barbat, Noncanonical roles of RAD51, Cells 12 (2023).
[111] K. Dunn, J.D. Griffith, The presence of RNA in a double helix inhibits its interaction with histone protein, Nucleic Acids Res. 8 (1980) 555–566.
[112] P.W. Lewis, S.J. Elsaesser, K.-M. Noh, S.C. Stadler, C.D. Allis, Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres, Proc. Natl. Acad. Sci. (2010), https://doi.org/10.1073/pnas.1008850107.
[113] V.M. Morozov, S. Giovinazzi, A.M. Ishov, CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specific chaperone Daxx, Epigenet. Chromatin 10 (2017) 1–18.
[114] D.T. Nguyen, et al., The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats, EMBO Rep. 18 (2017) 914–928.
[115] Q. He, et al., The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation, Cell Stem Cell 17 (2015) 273–286.
[116] F. Li, et al., <scp>ATRX</scp> loss induces telomere dysfunction and necessitates induction of alternative lengthening of telomeres during human cell immortalization, EMBO J. 38 (2019) e96659.
[117] H. Lee, Z. Zhang, H.M. Krause, Long noncoding RNAs and repetitive elements: junk or intimate evolutionary partners? Trends Genet. 35 (2019) 892–902.
[118] T. Liehr, Benign and Pathological Chromosomal Imbalances: Microscopic and Submicroscopic Copy Number Variations (CNVs) in Genetics and Counseling, Academic Press, 2013.
[119] T. Liehr, Repetitive elements in humans, Int. J. Mol. Sci. 22 (2021) 2072.
[120] G.-F. Richard, A. Kerrest, B. Dujon, Comparative genomics and molecular dynamics of DNA repeats in eukaryotes, Microbiol. Mol. Biol. Rev. 72 (2008) 686–727.
[121] N. Bannert, R. Kurth, Retroelements and the human genome: new perspectives on an old relation, Proc. Natl. Acad. Sci. 101 (2004) 14572–14579.
[122] J.N. Wells, C. Feschotte, A field guide to eukaryotic transposable elements, Annu. Rev. Genet. 54 (2020) 539–561.
[123] F. Brändle, B. Frühbauer, M. Jagannathan, Principles and functions of pericentromeric satellite DNA clustering into chromocenters, Semin. Cell Dev. Biol. 128 (2022) 26–39.
[124] L.P. Jenner, V. Peska, J. Fulnečková, E. Sýkorová, Telomeres and their neighbors, Genes 13 (2022).
[125] J.-Y. Chen, et al., Systematic evaluation of different R-loop mapping methods: achieving consensus, resolving discrepancies and uncovering distinct types of RNA:DNA hybrids, bioRxiv 2022 2 (18) (2022) 480986, https://doi.org/10.1101/2022.02.18.480986.
[126] Q. Yan, E.J. Shields, R. Bonasio, K. Sarma, Mapping native R-loops genome-wide using a targeted nuclease approach, Cell Rep. 29 (2019) 1369–1380.e5.
[127] J.-Y. Chen, X. Zhang, X.-D. Fu, L. Chen, R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1, Nat. Protoc. 14 (2019) 1661–1685.
[128] L.A. Sanz, F. Chédin, High-resolution, strand-specific R-loop mapping via S9.6-based DNA–RNA immunoprecipitation and high-throughput sequencing, Nat. Protoc. 14 (2019) 1734–1755.
[129] C. Bou-Nader, A. Bothra, D.N. Garboczi, S.H. Leppla, J. Zhang, Structural basis of R-loop recognition by the S9.6 monoclonal antibody, Nat. Commun. 13 (2022) 1641.
[130] L. Chen, et al., R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters, Mol. Cell 68 (2017) 745–757.e5.
[131] M. Malig, S.R. Hartono, J.M. Giafaglione, L.A. Sanz, F. Chedin, Ultra-deep coverage single-molecule R-loop footprinting reveals principles of R-loop formation, J. Mol. Biol. 432 (2020) 2271–2288.
[132] A. El Hage, S. Webb, A. Kerr, D. Tollervey, Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria, PLoS Genet. 10 (2014) e1004716.
[133] Y.A. Chan, et al., Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip, PLoS Genet. 10 (2014) e1004288.
[134] A. Munden, M.L. Benton, J.A. Capra, J. Nordman, R-loop mapping and characterization during Drosophila embryogenesis reveals developmental plasticity in R-loop signatures, bioRxiv 2021 10 (29) (2021) 465954, https://doi.org/10.1101/2021.10.29.465954.
[135] B.E. Black, D.W. Cleveland, Epigenetic centromere propagation and the nature of CENP-a nucleosomes, Cell 144 (2011) 471–479.
[136] K. Smurova, P. De Wulf, Centromere and pericentromere transcription: roles and regulation … in sickness and in health, Front. Genet. 9 (2018) 674.
[137] V. Régnier, et al., CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1, Mol. Cell Biol. 25 (2005) 3967–3981.
[138] H. Masumoto, M. Nakano, J. Ohzeki, The role of CENP-B and α-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres, Chromosome Res. 12 (2004) 543–556.
[139] V.M. Bolanos-Garcia, Aurora kinases, Int. J. Biochem. Cell Biol. 37 (2005) 1572–1577.
[140] G. Wang, Q. Jiang, C. Zhang, The role of mitotic kinases in coupling the centrosome cycle with the assembly of the mitotic spindle, J. Cell Sci. 127 (2014) 4111–4122.
[141] M. Carmena, S. Ruchaud, W.C. Earnshaw, Making the Auroras glow: regulation of Aurora A and B kinase function by interacting proteins, Curr. Opin. Cell Biol. 21 (2009) 796–805.
[142] A.P. Damodaran, L. Vaufrey, O. Gavard, C. Prigent, Aurora A kinase is a priority pharmaceutical target for the treatment of cancers, Trends Pharmacol. Sci. 38 (2017) 687–700.
[143] S.M. McNulty, L.L. Sullivan, B.A. Sullivan, Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with CENP-A and CENP-C, Dev. Cell 42 (2017) 226–240.e6.

[144] D. Quénet, Y. Dalal, A long non-coding RNA is required for targeting centromeric protein A to the human centromere, Elife 3 (2014) e03254–e03254.
[145] F. Lyn Chan, L.H. Wong, Transcription in the maintenance of centromere chromatin identity, Nucleic Acids Res. 40 (2012) 11178–11188.
[146] T. Ideue, Y. Cho, K. Nishimura, T. Tani, Involvement of satellite I noncoding RNA in regulation of chromosome segregation, Gene Cell. 19 (2014) 528–538.
[147] Y. Cho, T. Ideue, M. Nagayama, N. Araki, T. Tani, RBMX is a component of the centromere noncoding RNP complex involved in cohesion regulation, Gene Cell. 23 (2018) 172–184.
[148] K. Nishimura, et al., DEAH box RNA helicase DHX38 associates with satellite I noncoding RNA involved in chromosome segregation, Gene Cell. 24 (2019) 585–590.
[149] G.O.M. Bobkov, N. Gilbert, P. Heun, Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation, J. Cell Biol. 217 (2018) 1957–1972.
[150] S. Rošić, F. Köhler, S. Erhardt, Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division, J. Cell Biol. 207 (2014) 335–349.
[151] E.C. Moran, et al., Mitotic R-loops direct Aurora B kinase to maintain centromeric cohesion, bioRxiv 2021 1 (14) (2021) 426738, https://doi.org/10.1101/2021.01.14.426738.
[152] J. Houseley, K. Kotovic, A. El Hage, D. Tollervey, Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control, EMBO J. 26 (2007) 4996–5006.
[153] P.K. Mishra, et al., R-loops at centromeric chromatin contribute to defects in kinetochore integrity and chromosomal instability in budding yeast, Mol. Biol. Cell 32 (2021) 74–89.
[154] Q. Zhu, et al., BRCA1 tumour suppression occurs via heterochromatin-mediated silencing, Nature 477 (2011) 179–184.
[155] A. Di Paolo, C. Racca, P. Calsou, F. Larminat, Loss of BRCA1 impairs centromeric cohesion and triggers chromosomal instability, Faseb. J. 28 (2014) 5250–5261.
[156] F. Bersani, et al., Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 15148–15153.
[157] Q. Zhu, et al., Heterochromatin-encoded satellite RNAs induce breast cancer, Mol. Cell 70 (2018) 842–853.e7.
[158] D.T. Ting, et al., Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers, Science 331 (2011) 593–596, 80-.
[159] M. Unoki, et al., CDCA7 and HELLS suppress DNA:RNA hybrid-associated DNA damage at pericentromeric repeats, Sci. Rep. 10 (2020) 17865.
[160] H.-T. Shih, et al., DNMT3b protects centromere integrity by restricting R-loop-mediated DNA damage, Cell Death Dis. 13 (2022) 546.
[161] S. Gopalakrishnan, B.A. Sullivan, S. Trazzi, G. Della Valle, K.D. Robertson, DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions, Hum. Mol. Genet. 18 (2009) 3178–3193.
[162] T. De Lange, et al., Structure and Variability of Human Chromosome Ends, MOLECULAR AND CELLULAR BIOLOGY, 1990.
[163] M.L. DuBois, Z.W. Haimberger, M.W. McIntosh, D.E. Gottschling, A quantitative assay for telomere protection in Saccharomyces cerevisiae, Genetics 161 (2002) 995–1013.
[164] J. Flint, et al., Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains, Hum. Mol. Genet. 6 (1997) 1305–1314.
[165] M.T. Hemann, C.W. Greider, G-strand overhangs on telomeres in telomerase-deficient mouse cells, Nucleic Acids Res. 27 (1999).
[166] De Lange, T. Shelterin, The protein complex that shapes and safeguards human telomeres, Genes Dev. 19 (2005) 2100–2110.
[167] K. Cleal, K. Norris, D. Baird, Telomere length dynamics and the evolution of cancer genome architecture, Int. J. Mol. Sci. 19 (2018) 482.
[168] S. Schoeftner, M.A. Blasco, Chromatin regulation and non-coding RNAs at mammalian telomeres, Semin. Cell Dev. Biol. 21 (2010) 186–193.
[169] M. Achrem, I. Szućko, A. Kalinka, The epigenetic regulation of centromeres and telomeres in plants and animals, Comp. Cytogenet. 14 (2020) 265.
[170] S. Schoeftner, B. Ma, Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II, Nat. Cell Biol. 10 (2008) 228–236.
[171] C.M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto, J. Lingner, Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends, Science 318 (2007) 798–801.
[172] A. Diman, A. Decottignies, Genomic origin and nuclear localization of TERRA telomeric repeat-containing RNA: from Darkness to Dawn, FEBS J. 285 (2018) 1389–1398.
[173] E. Cusanelli, P. Chartrand, Telomeric repeat-containing RNA TERRA: a non-coding RNA connecting telomere biology to genome integrity, Front. Genet. 6 (2015) 143.
[174] N. Bettin, C. Oss Pegorar, E. Cusanelli, The emerging roles of TERRA in telomere maintenance and genome stability, Cells 8 (2019) 246.
[175] A. Chebly, et al., Telomeric repeat-containing RNA (TERRA): a review of the literature and first assessment in cutaneous T-cell lymphomas, Genes 13 (2022) 539.
[176] B. Balk, et al., Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence, Nat. Struct. Mol. Biol. 20 (2013) 1199–1206.
[177] V. Pfeiffer, J. Crittin, L. Grolimund, J. Lingner, The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening, EMBO J. 32 (2013) 2861–2871.
[178] M. Feretzaki, et al., RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops, Nature 587 (2020) 303–308.
[179] R. Arora, et al., RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells, Nat. Commun. 5 (2014) 1–11.
[180] Y. Gong, Y. Liu, R-loops at chromosome ends: from formation, regulation, and cellular consequence, Cancers 15 (2023) 2178.
[181] D. Hanahan, R.A. Weinberg, Hallmarks of Cancer : an Organizing Principle for Cancer Medicine, 2015 (Chapter 2) :.
[182] G. Chin, C.S. Lansing, *The Biological Sciences Collaboratory*. (Pacific Northwest National Lab.(PNNL), 2004. Richland, WA (United States).
[183] M. De Vitis, F. Berardinelli, A. Sgura, Molecular Sciences Telomere Length Maintenance in Cancer: at the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT), 2018, https://doi.org/10.3390/ijms19020606.
[184] R.J. O'Sullivan, J. Telomeres Karlseder, Protecting chromosomes against genome instability, Nat. Rev. Mol. Cell Biol. 11 (2010) 171–181.
[185] J.P. Amorim, G. Santos, J. Vinagre, P. Soares, The role of ATRX in the alternative lengthening of telomeres (ALT) phenotype, Genes 7 (2016).
[186] M.S. d'Alcontres, J.A. Palacios, D. Mejias, M.A. Blasco, TopoIIα prevents telomere fragility and formation of ultra thin DNA bridges during mitosis through TRF1-dependent binding to telomeres, Cell Cycle 13 (2014) 1463–1481.
[187] J. Ye, et al., TRF2 and apollo cooperate with topoisomerase 2alpha to protect human telomeres from replicative damage, Cell 142 (2010) 230–242.
[188] E. Petti, et al., SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability, Nat. Commun. 10 (2019).
[189] K. Takahama, et al., Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS, Chem. Biol. 20 (2013) 341–350.
[190] V.F. Thompson, et al., A role for phase separation to prevent R-loops during transcription, bioRxiv (2022) 2008–2022.
[191] A. Ishiguro, A. Katayama, A. Ishihama, Different recognition modes of G-quadruplex RNA between two ALS/FTLD-linked proteins TDP-43 and FUS, FEBS Lett. 595 (2021) 310–323.
[192] M. Wood, et al., TDP-43 dysfunction results in R-loop accumulation and DNA replication defects, J. Cell Sci. 133 (2020).
[193] M.A. Dyer, Z.A. Qadeer, D. Valle-Garcia, E.A.T.R.X. Bernstein, DAXX, Mechanisms and mutations, Cold Spring Harb. Perspect. Med. 7 (2017) a026567.
[194] H.P. Chu, et al., TERRA RNA Antagonizes ATRX and Protects Telomeres, Cell, 2017, https://doi.org/10.1016/j.cell.2017.06.017.
[195] T. Yadav, et al., TERRA and RAD51AP1 promote alternative lengthening of telomeres through an R- to D-loop switch, Mol. Cell 82 (2022) 3985–4000.e4.
[196] N. Kaminski, et al., RAD51AP1 regulates ALT-HDR through chromatin-directed homeostasis of TERRA, Mol. Cell 82 (2022) 4001–4017.e7.
[197] A. Helbling-Leclerc, C. Garcin, F. Rosselli, Beyond DNA repair and chromosome instability—fanconi anaemia as a cellular senescence-associated syndrome, Cell Death Differ. 28 (2021) 1159–1173.
[198] S. West, N. Gromak, N.J. Proudfoot, Human 5′→ 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites, Nature 432 (2004) 522–525.
[199] J. Vohhodina, et al., BRCA1 binds TERRA RNA and suppresses R-Loop-based telomeric DNA damage, Nat. Commun. 12 (2021) 3542.
[200] A. De Amicis, et al., Role of senataxin in DNA damage and telomeric stability, DNA Repair 10 (2011) 199–209.
[201] K. Skourti-Stathaki, N.J. Proudfoot, N. Gromak, Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote xrn2-dependent termination, Mol. Cell 42 (2011) 794–805.
[202] L. Pompili, C. Maresca, A. Dello Stritto, A. Biroccio, E. Salvati, BRCA2 deletion induces alternative lengthening of telomeres in telomerase positive colon cancer cells, Genes 10 (2019) 697.
[203] D. Li, K. Hou, K. Zhang, S. Jia, Regulation of replication stress in alternative lengthening of telomeres by fanconi anaemia protein, Genes 13 (2022) 180.
[204] K. Sato, P. Knipscheer, Rewiring E2 enzymes, Nat. Chem. Biol. 16 (2020) 227–229.
[205] Q. Fan, F. Zhang, B. Barrett, K. Ren, P.R. Andreassen, A Role for Monoubiquitinated FANCD2 at Telomeres in ALT Cells, vol. 37, 2009, pp. 1740–1754.
[206] Z. Yan, et al., A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability, Mol. Cell 37 (2010) 865–878.
[207] T.R. Singh, et al., Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M, Blood 114 (2009) 174–180.
[208] Z. Liang, et al., Binding of FANCI-FANCD2 complex to RNA and R-loops stimulates robust FANCD2 monoubiquitination, Cell Rep. 26 (2019) 564–572.e5.
[209] R. Lu, et al., The FANCM-BLM-TOP3A-RMI complex suppresses alternative lengthening of telomeres (ALT), Nat. Commun. 10 (2019) 2252.
[210] C.M. Heaphy, et al., Altered telomeres in tumors with ATRX and DAXX mutations, Science 333 (2011) 425.
[211] J. Schwartzentruber, et al., Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma, Nature 482 (2012) 226–231.

[212] Y. Hu, et al., Switch telomerase to ALT mechanism by inducing telomeric DNA damages and dysfunction of ATRX and DAXX, Sci. Rep. 6 (2016) 1–10.
[213] C.M. Heaphy, et al., Altered telomeres in tumors with ATRX and DAXX mutations, Science 333 (2011) 425.
[214] E. Mason-Osann, et al., Identification of a novel gene fusion in ALT positive osteosarcoma, Oncotarget 9 (2018) 32868–32880.
[215] C. Navarro, J. Lyu, A.-M. Katsori, R. Caridha, S.J. Elsässer, An embryonic stem cell-specific heterochromatin state promotes core histone exchange in the absence of DNA accessibility, Nat. Commun. 11 (2020) 5095.
[216] D. Sadic, et al., Atrx promotes heterochromatin formation at retrotransposons, EMBO Rep. 16 (2015) 836–850.
[217] S.J. Elsässer, K.-M. Noh, N. Diaz, C.D. Allis, L.A. Banaszynski, Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells, Nature 522 (2015) 240–244.
[218] D. Clynes, et al., Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX, Nat. Commun. (2015), https://doi.org/10.1038/ncomms8538.
[219] K.E. Yost, et al., Rapid and reversible suppression of ALT by DAXX in osteosarcoma cells, Sci. Rep. 9 (2019) 1–11.
[220] S. Toubiana, et al., Persistent epigenetic memory impedes rescue of the telomeric phenotype in human ICF iPSCs following DNMT3B correction, Elife 8 (2019) e47859.
[221] S. Yehezkel, Y. Segev, E. Viegas-Péquignot, K. Skorecki, S. Selig, Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions, Hum. Mol. Genet. 17 (2008) 2776–2789.
[222] S. Yehezkel, et al., Characterization and rescue of telomeric abnormalities in ICF syndrome type I fibroblasts, Front. Oncol. 3 (2013).
[223] R.G. Worton, et al., Human ribosomal RNA genes: orientation of the tandem array and conservation of the 5′ end, Science 239 (1988) 64–68.
[224] D.O. Warmerdam, R.M.F. Wolthuis, Keeping ribosomal DNA intact: a repeating challenge, Chromosome Res. 27 (2019) 57–72.
[225] J.G. Gibbons, A.T. Branco, S. Yu, B. Lemos, Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans, Nat. Commun. 5 (2014) 4850.
[226] L. Latonen, Phase-to-phase with nucleoli–stress responses, protein aggregation and novel roles of RNA, Front. Cell. Neurosci. 13 (2019) 151.
[227] D.L.J. Lafontaine, J.A. Riback, R. Bascetin, C.P. Brangwynne, The nucleolus as a multiphase liquid condensate, Nat. Rev. Mol. Cell Biol. 22 (2021) 165–182.
[228] A. Maiser, et al., Super-resolution in situ analysis of active ribosomal DNA chromatin organization in the nucleolus, Sci. Rep. 10 (2020) 7462.
[229] R.-W. Yao, et al., Nascent pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus, Mol. Cell 76 (2019) 767–783.e11.
[230] F.J. Lawrence, B. McStay, D.A. Matthews, Nucleolar protein upstream binding factor is sequestered into adenovirus DNA replication centres during infection without affecting RNA polymerase I location or ablating rRNA synthesis, J. Cell Sci. 119 (2006) 2621–2631.
[231] T. Kobayashi, Strategies to maintain the stability of the ribosomal RNA gene repeats–collaboration of recombination, cohesion, and condensation, Genes Genet. Syst. 81 (2006) 155–161.
[232] T. Kobayashi, Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast, Cell. Mol. Life Sci. 68 (2011) 1395–1403.
[233] L.M. Korsholm, et al., Recent advances in the nucleolar responses to DNA double-strand breaks, Nucleic Acids Res. 48 (2020) 9449–9461.
[234] A. El Hage, S. Webb, A. Kerr, D. Tollervey, Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria, PLoS Genet. 10 (2014).
[235] A.K. Velichko, et al., Hypoosmotic stress induces R loop formation in nucleoli and ATR/ATM-dependent silencing of nucleolar transcription, Nucleic Acids Res. 47 (2019) 6811–6825.
[236] S. Feng, J.L. Manley, Replication protein A associates with nucleolar R loops and regulates rRNA transcription and nucleolar morphology, Genes Dev. 35 (2021) 1579–1594.
[237] M. Derenzini, L. Montanaro, D. Treré, What the nucleolus says to a tumour pathologist, Histopathology 54 (2009) 753–762.
[238] A. Salvetti, A. Greco, Since January 2020 Elsevier Has Created a COVID-19 Resource Centre with Free Information in English and Mandarin on the Novel Coronavirus COVID- 19, The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information, 2020.
[239] A. Buchwalter, M.W. Hetzer, Nucleolar expansion and elevated protein translation in premature aging, Nat. Commun. 8 (2017) 328.
[240] A. El Hage, S.L. French, A.L. Beyer, D. Tollervey, Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis, Genes Dev. (2010), https://doi.org/10.1101/gad.573310.
[241] L.F. Liu, Involvement of DNA Topoisomerase I in Transcription of Human Ribosomal RNA Genes Ic, vol. 85, 1988, pp. 1060–1064.
[242] J. Wang, et al., PHF6 regulates cell cycle progression by suppressing ribosomal RNA synthesis, J. Biol. Chem. 288 (2013) 3174–3183.
[243] K.J. Abraham, et al., Nucleolar RNA polymerase II drives ribosome biogenesis, Nature 585 (2020) 298–302.
[244] E. Ford, et al., Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription, Genes Dev. 20 (2006) 1075–1080.
[245] E. Calo, et al., RNA helicase DDX21 coordinates transcription and ribosomal RNA processing, Nature 518 (2015) 249–253.
[246] C. Song, A. Hotz-Wagenblatt, R. Voit, I. Grummt, SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability, Genes Dev. 31 (2017) 1370–1381.
[247] T. Sekiguchi, T. Hayano, M. Yanagida, N. Takahashi, T. Nishimoto, NOP132 is required for proper nucleolus localization of DEAD-box RNA helicase DDX47, Nucleic Acids Res. 34 (2006) 4593–4608.
[248] Y. Okamoto, et al., FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress, FEBS J. 286 (2019) 139–150.
[249] C.-W. Chang, et al., The BRCA1 BRCT promotes antisense RNA production and double-stranded RNA formation to suppress ribosomal R-loops, Proc. Natl. Acad. Sci. 119 (2022) e2217542119.
[250] M. Udugama, et al., Ribosomal DNA copy loss and repeat instability in ATRX-mutated cancers, Proc. Natl. Acad. Sci. 115 (2018) 4737–4742.
[251] K.R. Kampen, S.O. Sulima, S. Vereecke, K. De Keersmaecker, Hallmarks of ribosomopathies, Nucleic Acids Res. 48 (2020) 1013–1028.
[252] J. Tan, X. Wang, L. Phoon, H. Yang, L. Lan, Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase, FEBS Lett. 594 (2020) 1359–1367.
[253] V. Gharibyan, H. Youssoufian, Localization of the Bloom syndrome helicase to punctate nuclear structures and the nuclear matrix and regulation during the cell cycle: comparison with the Werner's syndrome helicase, Mol. Carcinog. 26 (1999) 261–273.
[254] V. Yankiwski, R.A. Marciniak, L. Guarente, N.F. Neff, Nuclear structure in normal and Bloom syndrome cells, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5214–5219.
[255] S. Zhong, et al., A role for PML and the nuclear body in genomic stability, Oncogene 18 (1999) 7941–7947.
[256] M.W. Killen, D.M. Stults, N. Adachi, L. Hanakahi, A.J. Pierce, Loss of Bloom syndrome protein destabilizes human gene cluster architecture, Hum. Mol. Genet. 18 (2009) 3417–3428.
[257] D.M. Stults, M.W. Killen, H.H. Pierce, A.J. Pierce, Genomic architecture and inheritance of human ribosomal RNA gene clusters, Genome Res. 18 (2008) 13–18.
[258] P.M. Grierson, et al., BLM helicase facilitates RNA polymerase I-mediated ribosomal RNA transcription, Hum. Mol. Genet. 21 (2012) 1172–1183.
[259] J. Hallgren, M. Pietrzak, G. Rempala, P.T. Nelson, M. Hetman, Neurodegeneration-associated instability of ribosomal DNA, Biochim. Biophys. Acta, Mol. Basis Dis. 1842 (2014) 860–868.
[260] I.V. Chestkov, et al., Abundance of ribosomal RNA gene copies in the genomes of schizophrenia patients, Schizophr. Res. 197 (2018) 305–314.
[261] A. Gueiderikh, F. Maczkowiak-Chartois, F. Rosselli, A new frontier in Fanconi anemia: from DNA repair to ribosome biogenesis, Blood Rev. 52 (2022) 100904.
[262] B. Xu, et al., Ribosomal DNA copy number loss and sequence variation in cancer, PLoS Genet. 13 (2017) e1006771.
[263] M. Wang, B. Lemos, Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation, PLoS Genet. 13 (2017) e1006994.
[264] E. Smirnov, N. Chmúrčiaková, D. Cmarko, Human rDNA and cancer, Cells 10 (2021).
[265] C.R.L. Huang, K.H. Burns, J.D. Boeke, Active transposition in genomes, Annu. Rev. Genet. 46 (2012) 651–675.
[266] C. Feschotte, E.J. Pritham, DNA transposons and the evolution of eukaryotic genomes, Annu. Rev. Genet. 41 (2007) 331–368.
[267] T. Wicker, et al., A unified classification system for eukaryotic transposable elements, Nat. Rev. Genet. 8 (2007) 973–982.
[268] M. Ravi, S. Ramanathan, K. Krishna, Factors, mechanisms and implications of chromatin condensation and chromosomal structural maintenance through the cell cycle, J. Cell. Physiol. 235 (2020) 758–775.
[269] K. Saito, The epigenetic regulation of transposable elements by PIWI-interacting RNAs in *Drosophila*, Genes Genet. Syst. 88 (2013) 9–17.
[270] D. Holoch, D. Moazed, RNA-mediated epigenetic regulation of gene expression, Nat. Rev. Genet. 16 (2015) 71–84.
[271] H.H. Kazazian Jr., Moran, J. V. Mobile DNA in health and disease, N. Engl. J. Med. 377 (2017) 361–370.
[272] H.H. Kazazian, et al., Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man, Nature 332 (1988) 164–166.
[273] K.H. Burns, Transposable elements in cancer, Nat. Rev. Cancer 17 (2017) 415–424.
[274] F. Yu, N. Zingler, G. Schumann, W.H. Strätling, Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription, Nucleic Acids Res. 29 (2001) 4493–4501.
[275] V. Fort, G. Khelifi, S.M.I. Hussein, Long non-coding RNAs and transposable elements: a functional relationship, Biochim. Biophys. Acta Mol. Cell Res. 1868 (2021) 118837.
[276] M. Nakama, K. Kawakami, T. Kajitani, T. Urano, Y. Murakami, DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation, Gene Cell. (2012), https://doi.org/10.1111/j.1365-2443.2012.01583.x.
[277] Q. Al-Hadid, Y. Yang, R-loop: an emerging regulator of chromatin dynamics, Acta Biochim. Biophys. Sin. 48 (2016) 623–631.
[278] J.A. Birchler, F. Han, Maize centromeres: structure, function, epigenetics, Annu. Rev. Genet. 43 (2009) 287–303.

[279] Y. Liu, et al., Back-spliced RNA from retrotransposon binds to centromere and regulates centromeric chromatin loops in maize, PLoS Biol. 18 (2020) e3000582–e3000582.
[280] X. Li, L. Yang, L.-L. Chen, The biogenesis, functions, and challenges of circular RNAs, Mol. Cell 71 (2018) 428–442.
[281] M. Alkailani, et al., A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, NAR cancer 3 (2021) zcaa040.
[282] Y. Nie, et al., FANCD2 is required for the repression of germline transposable elements, Reproduction 159 (2020) 659–668.
[283] Y.J. Crow, N. Manel, Aicardi–Goutières syndrome and the type I interferonopathies, Nat. Rev. Immunol. 15 (2015) 429–440.
[284] H.E. Volkman, D.B. Stetson, The enemy within: endogenous retroelements and autoimmune disease, Nat. Immunol. 15 (2014) 415–422.
[285] K. Zhao, et al., Modulation of LINE-1 and alu/SVA retrotransposition by aicardi-goutières syndrome-related SAMHD1, Cell Rep. 4 (2013) 1108–1115.
[286] Y.W. Lim, L.A. Sanz, X. Xu, S.R. Hartono, F. Chédin, Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome, Elife 4 (2015) e08007.
[287] C.A. Thomas, et al., Modeling of TREX1-dependent autoimmune disease using human stem cells highlights L1 accumulation as a source of neuroinflammation, Cell Stem Cell 21 (2017) 319–331.e8.
[288] E. Orecchini, et al., ADAR1 restricts LINE-1 retrotransposition, Nucleic Acids Res. 45 (2017) 155–168.
[289] V. Pokatayev, et al., RNase H2 catalytic core Aicardi-Goutières syndrome-related mutant invokes cGAS-STING innate immune-sensing pathway in mice, J. Exp. Med. 213 (2016) 329–336.
[290] P. Li, et al., Aicardi-Goutières syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion, Nucleic Acids Res. 45 (2017) 4619–4631.
[291] K. Bartsch, et al., Absence of RNase H2 triggers generation of immunogenic micronuclei removed by autophagy, Hum. Mol. Genet. 26 (2017) 3960–3972.
[292] M. Benitez-Guijarro, et al., RNase H2, mutated in Aicardi-Goutières syndrome, promotes LINE-1 retrotransposition, EMBO J. 37 (2018).
[293] A. Cristini, et al., RNase H2, mutated in Aicardi-Goutières syndrome, resolves co-transcriptional R-loops to prevent DNA breaks and inflammation, Nat. Commun. 13 (2022) 2961.
[294] A. Saleh, A. Macia, A.R. Muotri, Transposable elements, inflammation, and neurological disease, Front. Neurol. 10 (2019) 894.
[295] C.A. Ishak, M. Classon, D.D. De Carvalho, Deregulation of retroelements as an emerging therapeutic opportunity in cancer, Trends in Cancer 4 (2018) 583–597.
[296] D. Filipponi, J. Muller, A. Emelyanov, D.V. Bulavin, Wip1 controls global heterochromatin silencing via ATM/BRCA1-dependent DNA methylation, Cancer Cell 24 (2013) 528–541.
[297] P. Mita, et al., BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells, Nat. Struct. Mol. Biol. 27 (2020) 179–191.
[298] M.M. Axford, et al., Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues, PLoS Genet. 9 (2013) e1003866.
[299] H. Paulson, Repeat expansion diseases, Handb. Clin. Neurol. 147 (2018) 105–123.
[300] S.M. Mirkin, Expandable DNA repeats and human disease, Nature 447 (2007) 932–940.
[301] Y. Lin, S.Y.R. Dent, J.H. Wilson, R.D. Wells, M. Napierala, R loops stimulate genetic instability of CTG·CAG repeats, Proc. Natl. Acad. Sci. 107 (2010) 692–697.
[302] M. Groh, M.M.P. Lufino, R. Wade-Martins, N. Gromak, R-Loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome, PLoS Genet. 10 (2014) e1004318.
[303] E.W. Loomis, L.A. Sanz, F. Chédin, P.J. Hagerman, Transcription-associated R-loop formation across the human FMR1 CGG-repeat region, PLoS Genet. 10 (2014) e1004294.
[304] X.A. Su, C.H. Freudenreich, Cytosine deamination and base excision repair cause R-loop–induced CAG repeat fragility and instability in Saccharomyces cerevisiae, Proc. Natl. Acad. Sci. 114 (2017) E8392–E8401.
[305] Y. Lin, V. Dion, J.H. Wilson, Transcription promotes contraction of CAG repeat tracts in human cells, Nat. Struct. Mol. Biol. 13 (2006) 179–180.
[306] S. Nageshwaran, R. Festenstein, Epigenetics and triplet-repeat neurological diseases, Front. Neurol. 6 (2015) 262.
[307] M.B. Delatycki, R. Williamson, S.M. Forrest, Friedreich ataxia: an overview, J. Med. Genet. 37 (2000) 1–8.
[308] A.J. Neil, M.U. Liang, A.N. Khristich, K.A. Shah, S.M. Mirkin, RNA–DNA hybrids promote the expansion of Friedreich's ataxia (GAA)n repeats via break-induced replication, Nucleic Acids Res. 46 (2018) 3487–3497.
[309] S. Gomez-Sebastian, A. Gimenez-Cassina, J. Diaz-Nido, F. Lim, R. Infectious Wade-Martins, Delivery and expression of a 135 kb human FRDA genomic DNA locus complements Friedreich's ataxia deficiency in human cells, Mol. Ther. 15 (2007) 248–254.
[310] E. Soragni, et al., Long intronic GAA*TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia, Nucleic Acids Res. 36 (2008) 6056–6065.
[311] D. Kumari, B. Hayward, A.J. Nakamura, W.M. Bonner, K. Usdin, Evidence for chromosome fragility at the frataxin locus in Friedreich ataxia, Mutat. Res. Mol. Mech. Mutagen. 781 (2015) 14–21.
[312] D.L. Nelson, H.T. Orr, S.T. Warren, The unstable repeats–three evolving faces of neurological disease, Neuron 77 (2013) 825–843.
[313] L.W. Wang, E. Berry-Kravis, R.J. Hagerman, X. Fragile, Leading the way for targeted treatments in autism, Neurother. J. Am. Soc. Exp. Neurother. 7 (2010) 264–274.
[314] M. Pieretti, et al., Absence of expression of the FMR-1 gene in fragile X syndrome, Cell 66 (1991) 817–822.
[315] A.J. Verkerk, et al., Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome, Cell 65 (1991) 905–914.
[316] B. Coffee, F. Zhang, S. Ceman, S.T. Warren, D. Reines, Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome, Am. J. Hum. Genet. 71 (2002) 923–932.
[317] A. Chakraborty, et al., Replication stress induces global chromosome breakage in the fragile X genome, Cell Rep. 32 (2020) 108179.
[318] D. Colak, et al., Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome, Science 343 (2014) 1002–1005, 80-.
[319] K. Sathasivam, et al., Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease, Proc. Natl. Acad. Sci. 110 (2013) 2366–2370.
[320] S. Thongthip, A. Carlson, M.P. Crossley, B. Schwer, Relationships between genome-wide R-loop distribution and classes of recurrent DNA breaks in neural stem/progenitor cells, Sci. Rep. 12 (2022) 13373.
[321] K. Reddy, et al., Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats, Nucleic Acids Res. 39 (2011) 1749–1762.
[322] E.I. McIvor, U. Polak, M. Napierala, New insights into repeat instability: role of RNA•DNA hybrids, RNA Biol. 7 (2010) 551–558.
[323] P. Richard, J.L.R. Manley, Loops and links to human disease, J. Mol. Biol. 429 (2017) 3168–3180.
[324] A.J. Hannan, Tandem repeats mediating genetic plasticity in health and disease, Nat. Rev. Genet. 19 (2018) 286–298.
[325] O. Aladesuyi Arogundade, et al., Antisense RNA foci are associated with nucleoli and TDP-43 mislocalization in C9orf72-ALS/FTD: a quantitative study, Acta Neuropathol. 137 (2019) 527–530.
[326] J. Wang, A.R. Haeusler, E.A.J. Simko, Emerging role of RNA•DNA hybrids in C9orf72-linked neurodegeneration, Cell Cycle 14 (2015) 526–532.
[327] K. Reddy, et al., Processing of double-R-loops in (CAG)·(CTG) and C9orf72 (GGGGCC)·(GGCCCC) repeats causes instability, Nucleic Acids Res. 42 (2014) 10473–10487.
[328] G.M. Hautbergue, et al., SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits, Nat. Commun. 8 (2017) 16063.
[329] G.S. Erwin, et al., Recurrent repeat expansions in human cancer genomes, Nature 613 (2023) 96–102.
[330] H. Masukata, S. Dasgupta, J. Tomizawa, Transcriptional activation of ColE1 DNA synthesis by displacement of the nontranscribed strand, Cell 51 (1987) 1123–1130.