

Review

Pathophysiological Role and Diagnostic Potential of R-Loops in Cancer and Beyond

Essak S. Khan ^{1,2,3,4}  and Sven Danckwardt ^{1,2,3,5,*}

- ¹ Posttranscriptional Gene Regulation, Cancer Research and Experimental Hemostasis, University Medical Center Mainz, 55131 Mainz, Germany
 - ² Institute for Clinical Chemistry and Laboratory Medicine, University Medical Center of the Johannes Gutenberg University, 55131 Mainz, Germany
 - ³ Center for Thrombosis and Hemostasis (CTH), University Medical Center of the Johannes Gutenberg University, 55131 Mainz, Germany
 - ⁴ German Consortium for Translational Cancer Research (DKTK), DKFZ Frankfurt-Mainz, 60590 Frankfurt am Main, Germany
 - ⁵ German Center for Cardiovascular Research (DZHK), Rhine-Main, 55131 Mainz, Germany
- * Correspondence: sven.danckwardt@unimedizin-mainz.de

Abstract: R-loops are DNA–RNA hybrids that play multifunctional roles in gene regulation, including replication, transcription, transcription–replication collision, epigenetics, and preserving the integrity of the genome. The aberrant formation and accumulation of unscheduled R-loops can disrupt gene expression and damage DNA, thereby causing genome instability. Recent links between unscheduled R-loop accumulation and the abundance of proteins that modulate R-loop biogenesis have been associated with numerous human diseases, including various cancers. Although R-loops are not necessarily causative for all disease entities described to date, they can perpetuate and even exacerbate the initially disease-eliciting pathophysiology, making them structures of interest for molecular diagnostics. In this review, we discuss the (patho) physiological role of R-loops in health and disease, their surprising diagnostic potential, and state-of-the-art techniques for their detection.

Keywords: R-loop; co-transcriptional processing; transcription termination; (alternative) polyadenylation; human disease; biomarker; prognostic power



Citation: Khan, E.S.; Danckwardt, S. Pathophysiological Role and Diagnostic Potential of R-Loops in Cancer and Beyond. *Genes* **2022**, *13*, 2181. <https://doi.org/10.3390/genes13122181>

Academic Editor: Martin Dutertre

Received: 7 October 2022

Accepted: 14 November 2022

Published: 22 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

R-loops are three-stranded structures also known as non-B DNA structures, which form when RNA hybridizes with a complementary DNA strand, resulting in DNA–RNA hybrids and a displaced non-template, single-stranded DNA (ssDNA) (Figure 1). While short DNA–RNA hybrids form transiently during transcription and lagging-strand DNA synthesis, R-loops differ from these structures. They span 100–2000 base pairs [1] and are typically located at the 5' end of the elongating RNA polymerase II. R-loops are abundant in the mammalian genome [2]. They are frequently found in GC-rich regions where they play a multifunctional role in modulating diverse aspects of gene regulation including replication, transcription, transcription–replication collision, and epigenetics, as well as maintaining genome integrity [3,4]. Apart from their regulatory function, the accumulation of unscheduled R-loops can have a detrimental effect on genome integrity. They serve as a primer for aberrant replication, act as replication blocks and effectors of transcription stress, and function as a source of DNA damage, thereby disrupting genome integrity [4–8]. Hence, cells have evolved mechanisms that tightly control the formation and resolution of R-loops to ensure timely controlled gene regulation and to avoid undesired mutagenic events [5,9,10]. The abundance of R-loops is known to be regulated directly or indirectly by many proteins, mostly preventing RNA from hybridizing with DNA, thus reducing excessive R-loop accumulation [11]. Among these are proteins required for

efficient transcription initiation [12], transcription elongation [13] and termination [14], polyadenylation [15], RNA splicing [16], and RNA packaging and export [17,18]. While substantial progress has been made to uncover the underlying mechanisms and biological role of R-loop formation (reviewed elsewhere in great detail [4–8]), increasing numbers of reports indicate the detrimental consequences of unscheduled RNA formation. For example, aberrant R-loop formation and accumulation has been associated with several human diseases [3–5,7,19], such as cancer [20–22] and neurological [23,24], hematological [25], and cardiovascular disorders [21,26]. In addition, the initial reports [27,28] indicate a potential role of R-loops as a potent diagnostic biomarker. In this review, we summarize the physiological role and pathophysiological consequences of R-loops and shed light on their diagnostic potential—a rapidly growing subject. We also discuss the current methodologies used to detect R-loops and future perspectives in this area.

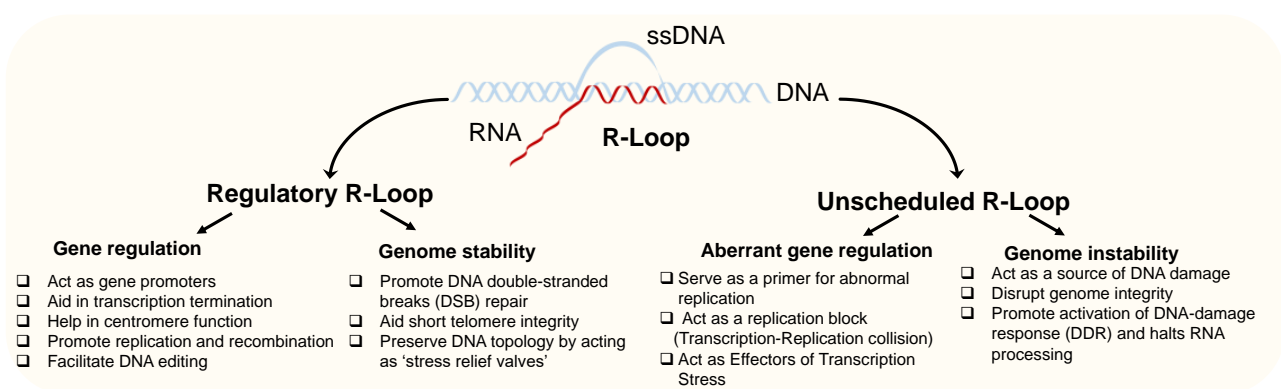


Figure 1. Roles of Regulatory and Unscheduled R-loops. Regulatory R-loops are intermediates required for gene regulation and genome stability (left). R-loops regulate gene activity by modulating transcriptional activity, replication, recombination, centromere function, and DNA editing. R-loops are also involved in stabilizing the genome by promoting the repair of DNA double-strand breaks (DSB) and short telomere structures. Absence of timely removal or prevention of unwanted R-loop accumulation results in the formation of unscheduled R-loops (right). Unscheduled R-loops are a source for aberrant gene regulation (by serving as a primer for aberrant replication, enhancing transcription–replication collision, and acting as an effector of transcription stress) and contribute to the instability of the genome (by inducing DNA damage). In turn, unscheduled R-loop formation promotes the DNA-damage response (DDR) activation and halts RNA processing to allow for genome repair [29].

2. Roles of Regulatory R-Loops in Gene Regulation and Genome Stability

For a long time, R-loops have been considered accidental by-products of transcription and a mere source of genomic instability when not removed correctly [5,9]. However, it has recently become evident that R-loops have various regulatory functions in biological processes. Depending on their role, R-loops are broadly classified as physiological or “regulatory” and pathological or “unscheduled” R-Loops [10] (Figure 1).

Regulatory R-loops are intermediates that modulate gene activity and the organization as well as stability of the genome. R-loops that are found in the promoter regions of genes are also known as *Promoter R-loops*. They actively regulate genes through several mechanisms [30–34]. They ease transcription by preventing the binding of DNA methylation enzymes (DNMT) [30,31]. They facilitate the binding of transcription factors [32] and prevent transcriptional repressors from binding to promote the transcription of genes [33]. In some cases, R-loops formed at the promoter region of transcription factors may also cause their silencing [34].

R-Loops are also enriched at the 3' ends of some genes and aid transcription termination by promoting various mechanisms [9]. They promote the termination of backtracked Pol II [10]. They also trigger antisense transcription to reinforce RNAPII pausing by re-

cruiting the RNA interference machinery [8,35]. R-loops are known to regulate the activity of chromatin remodelers [35–38]. They can induce chromatin de-condensation [36]. On the contrary, they can also promote heterochromatin assembly [35,37] and chromatin compaction [38]. Other than transcription, R-loops also facilitate DNA replication initiation in mitochondrial DNA [5], immunoglobulin (Ig) class switch recombination [39], or CISPR-Cas9 activity to facilitate Cas-9 mediated cleavage [40]. These few examples illustrate that regulatory R-loops have a plethora of functions by modulating diverse aspects including genomic structure, organization, and gene expression.

Apart from gene regulation, R-loops are also involved in maintaining the integrity of the genome. They promote DNA double-stranded breaks' (DSB) repair, the repair of short telomeres, and preserve DNA topology [7,9]. The R-loops at DSBs develop from either de novo transcription from a free 3' end at the break site or from transcription stalling within a gene [9]. Mechanistically, the R-loops formed at the DSBs attract breast cancer susceptibility protein 1 (BRCA1 [41]) and eventually other repair factors (BRCA2 [41]) to repair the DNA. Later, R-loops are removed by senataxin (SETX) [42], DEAD-box helicase 1 (DDX1) [43], ribonuclease (RNase) H1 [44] or RNase H2 [45], or homologous recombination [46] (further details on R-loop-modifying proteins are provided in Table 1 and Section 4). The accumulation of *telomere repeat-containing RNA* (TERRA) R-loops [47–49] at short telomeres [49] promotes DNA repair through RAD51-mediated homology-dependent repair [41,42,50]. Furthermore, R-loops are known to act as topological “stress relief valves” by relieving the higher energy of supercoiled DNA with a surprising capacity for relaxing tension in the genome (e.g., one R-loop can relax ~18 negative supercoils [51]) and can also extend to a variety of lengths beyond conserved sequences rich in CG or purines [52]. Taken together, regulatory R-loops are thus important for central aspects of genomic organization, gene regulation, and the maintenance of genome integrity.

3. Role of Unscheduled R-Loops as a Source of DNA Damage and Genomic Instability

Over the last few years, there has been increasing evidence that R-loops act as a double-edged sword [6–8]. Although R-loops are important regulatory elements for gene expression and genome stability, the aberrant or unscheduled formation of R-loops has been implicated in various diseases [6,7]. Unscheduled R-loop formation and R-loop accumulation can become a source of DNA damage. They are often linked to genome instability by forming weak ssDNA and DSBs [53] as well as potentiating replication fork stalling [54] and transcription–replication collision due to the low level or activity of topoisomerase 1 leading to recombination and genome instability [55,56].

The generation of DNA damage and replication stress triggers the activation of complex surveillance mechanisms, collectively called the “DNA damage response” (DDR), which are crucial to maintaining genome integrity and thus avoiding perturbations of a wide array of biological processes. Recent evidence suggests that unscheduled R-loop formation results in the activation of phosphatidylinositol 3-kinase-related protein kinases, including mammalian ATM (Ataxia–telangiectasia-mutated) and ATR (ATM- and Rad3-related) [57], which are key players governing the DDR. Both kinases are activated by DDRs, but their specificities are different. The dysfunction of R-loop homeostasis results in the activation of ATM/Tel1 and ATR/Mec1 kinase in a distinct manner. ATM/Tel1 is a multifunctional kinase that helps maintain genomic stability through its control of numerous aspects of cellular survival including telomere homeostasis [58,59]. To maintain the integrity of the genome, stalled replication forks are controlled by a checkpoint whose central player is the human kinase ATR/Mec1. It helps stabilize the replisome directly or by activating the checkpoint response to control DNA repair, fork restart, and other mechanisms for cell cycle progression [60]. Here, DNA damage due to unscheduled R-loop formation causes transcription–replication conflicts (TRCs) that trigger the ATR activation of the S-phase checkpoint [60]. Although many functional aspects of the underlying mechanisms remain to be elucidated, unscheduled or pathological R-loops are an important source of DNA damage and genome instability and can, therefore, become pathogenic

(further detail is provided below). Thus, cellular mechanisms must be in place that control the formation and resolution of R-loops.

4. Role of R-Loop-Binding Proteins and (Co-)Transcriptional Mechanisms in R-Loop Formation, Resolution, and Prevention of Aberrant R-Loop Accumulation

Numerous proteins and almost all processes involved in gene regulation control the dynamics of R-loop formation directly or indirectly, thereby preventing excessive R-loop accumulation (Table 1). The latter is likely best explained by the fact that the coordinated co-transcriptional processing and packaging of the nascent transcript into ‘inert’ ribonucleoprotein particles (RNPs) ensures that R-loops occur in a scheduled manner, thus preventing the formation of otherwise deleterious DNA–RNA hybrid structures [61]. Events that perturb the coordinated co-transcriptional processing can result in unscheduled R-loop accumulation, leading to replication-associated DNA damage. Several examples illustrate this principle. These include events and proteins involved in transcription initiation [62], elongation [17,63], RNA splicing [16] and polyadenylation [15], RNA packaging and export [17,18], R-loop processing [64–66], and DNA topology [67]. For example, during transcription initiation, R-loops are formed by capping enzyme–RNA polymerase II complex (CE–Pol II) to facilitate efficient transcription [68]. Here, the capping enzyme is recruited on Pol II, which modulates the displacement of the nascent RNA to form R-loops and thus enables the co-transcriptional capping of the pre-mRNA and elongation [62]. This is followed by the recruitment of the *Facilitates chromatin transcription* (FACT) complex (a chromatin-reorganizing complex that swaps nucleosomes around the RNA polymerase during transcription elongation and replication). It promotes replication by acting as a histone deposition chaperone contributing to nucleosome assembly to regulate untimely R-loop-mediated TRCs [17]. Hence, the depletion of the FACT complex causes excessive R-loop accumulation and impairment in replication and transcription, illustrating a critical function in the resolution of R-loop-mediated TRCs [17]. In addition, after transcription initiation and elongation, mRNA cleavage and polyadenylation (CPA) factors contribute to maintaining genome integrity by suppressing R-loop formation and modulating efficient mRNA cleavage [15,29]. CPA constitute an essential step in preventing replication-stress-associated DNA damage as transcription hinders replication fork progression and stability, and vice versa, compensatory pausing of co-transcriptional CPA emerges as a conserved mechanism of the DNA damage response (DDR) to allow for DNA repair and to avoid the propagation of genomic mutations [29]. Hence, perturbations of CPA-mediated co-transcriptional RNA processing are detrimental and can result in the formation of unscheduled R loops and genome instability [29]. For example, a loss of function of the 3′ end cleavage and polyadenylation factors (including PCF11, CLP1, FIP1L1, CFT2, and WDR33) results in DNA damage [15]. This cause’s excessive unscheduled R-loop accumulation tightly associated with replication-stress-induced DNA damage, the inhibition of transcription-rescued fork speed, origin activation, and alleviated replication catastrophe [69]. However, further factors localized at the RNA 3′ end and cooperating in transcription termination show a similar functional outcome, as demonstrated by the loss of function of Rtt103, the yeast homolog of RPRD1B [15], SETX [70], and XRN2 [71].

Another functional example is the *RNA-processing and export* (THO/TREX) complex, which is also involved in R-loop regulation. THO/TREX, a conserved nuclear complex, functions in messenger ribonucleoprotein (mRNP) biogenesis [72]. It inhibits aberrant R-loop formation and prevents transcription-associated recombination [13,73]. Mutations of *hpr1* (part of the THO complex) result in the genome-wide impairment of replication progression in transcribed genes due to extensive R-loop accumulation at the replication fork [73].

Table 1. Identity and function of R-loop-binding and R-loop-regulating proteins.

Protein	Function
Transcription initiation and capping	
Capping enzyme-Pol II complex [62]	Responsible for transcription initiation by modulating displacement of nascent RNA during transcription, thereby promoting R-loop formation
Transcription elongation	
Facilitates Chromatin Transcription (FACT) complex [17]	Helps in preventing R-loop accumulation-causing TRCs
Transcription termination, cleavage, and polyadenylation	
Cleavage and Polyadenylation (CPA) factors [15,29,69] (PCF11, CLP1, FIP1L1, CFT2, WDR33)	Suppresses R-loop formation and facilitates efficient mRNA cleavage, thereby preventing replication-stress-associated DNA damage
RNA processing and export	
Transcription and export complex (THO/TREX complex; Tho2/THOC2, Hpr1/THOC1, Mft1, Thp2, Sub2/UAP56) [13,73]	Inhibits aberrant R-loop formation and transcription-associated recombination
Splicing	
Serine- And Arginine-Rich Splicing Factor SRSF2 [16]	Prevents the formation of mutagenic R loop structures
RNA-Binding Protein With Serine Rich Domain 1 (RNPS1) [16]	Forms complex with ASF/SF2 to prevent transcriptional R-loops
R-loop degradation	
RNase H1/2 [10,74]	Prevents aberrant R-loop formation by timely removal of these hybrids
R-Loop-processing factors (DNA–RNA helicases)	
Senataxin (SETX) [75]	Binds to replication forks to protect its integrity across RNA-Polymerase-II-transcribed gene and unwinds unnecessary R-loops
Aquarius (AQR) [76]	Prevents R-loop formation by unwinding DNA–RNA hybrids
DExH-Box Helicase 9 (DHX9) [77]	Prevents R-loop formation by melting DNA–RNA hybrid with a 3′–5′ polarity
DExH-Box Helicase 11 (DHX11) [78]	Converts RNA G-Quadruplex structures into R-Loops to promote IgH class switch recombination
Werner Syndrome RecQ-Like Helicase (WRN) [64]	Protects the replication fork by preventing unscheduled R-loop formation
DNA topology	
Topoisomerase I/IIIB [55,67,79]	Involved in maintaining R-loop resolution by interacting with RNA-splicing and DNA-processing factors
DNA repair and genome maintenance	
Ataxia Telangiectasia Mutated (ATM)/Ataxia Telangiectasia And Rad3 Related (ATR) Kinase [57]	DNA-damage response (DDR) kinases that become activated when R-loop-mediated DNA damage occurs
Breast Cancer Type 2 Susceptibility Protein (BRCA2/FANCD1) [41]	Binds to R-loops in response to dsDNA breaks to invite other DNA repair factors

The improper dissolution of co-transcriptionally formed R-loops constitutes potential roadblocks for transcription and enhances transcription-associated recombination

events [10]. The intricate coupling between RNA processing, R-loop formation, and genome integrity also manifests in defects of RNA splicing and RNA degradation. During splicing, RNA *Serine/arginine-rich splicing factors* (SRs) such as the ASF/SF2 protein are recruited to nascent transcripts by RNA polymerase II, thus preventing the formation of mutagenic R loop structures [16,80]. The depletion of ASF/SF2 was shown to result in a single-stranded non-template strand of a transcribed gene due to the formation of aberrant R-loop structures, giving rise to genomic instability [16]. ASF/SF2 depletion-induced genomic instability can be alleviated by the overexpression of the *RNA-binding protein with Serine Rich Domain 1* (RNPS1) that has been suggested to function together with ASF/SF2 to form RNP complexes on nascent transcripts, thereby preventing the formation of transcriptional R-loops [16]. Along with splicing factors, RNA-degrading enzymes such as RNase 1 and 2 are also involved in regulating R-loops [10,74]. They eliminate hybrids created accidentally during replication, thereby suppressing genome instability associated with R-loop formation [15,81,82]. Consequently, the mutation of RNA-degrading enzymes has been shown to increase the formation of hybrids and associated genome instability [18].

Apart from the RNA-processing machinery, DNA–RNA hybrid-processing proteins such as DNA/RNA helicases are also involved in regulating R-loop resolution and processing. Popularly known as R-loop-processing factors, SETX [75], Aquarius (AQR) [76], DEXH-Box Helicase 9 (DHX9) [77], DEXH-Box Helicase 11 (DHX11) [78], and Werner Syndrome RecQ Like Helicase (WRN) [64] are DNA/RNA helicases, which are also involved in relieving replication stress. SETX associates with replication forks to protect its integrity across RNA-Polymerase-II-transcribed genes and unwinds unnecessary R-loops [75]. AQR prevents R-loop formation by constantly unwinding the DNA–RNA hybrid [76]. DHX9, belonging to the SF2 superfamily of nucleic acid-unwinding enzymes, melts DNA–RNA strands with a 3′–5′ polarity, thus contributing to transcriptional activation and thereby maintaining genomic stability [77]. DHX11, an RNA helicase that converts RNA G-Quadruplex structures into R-loops, promotes IgH class switch recombination [78]. WRN, which belongs to the RecQ family of helicases (RecQ helicases are an important family of genome surveillance proteins conserved from bacteria to humans often referred to as ‘guardians of the genome’), is involved in multiple pathways of DNA repair and the maintenance of genome integrity [83]. It protects the replication fork by forming a complex with *Werner helicase-interacting protein 1* (WRNIP1) and preventing unscheduled R-loop formation [64]. While DHX9 behaves similarly to WRN to unwind with a 3′–5′ polarity, DHX9 is considerably faster than WRN in unwinding RNA hybrids [64]. WRN preferably unwinds RNA-containing Okazaki fragment-like substrates whereas DHX9 fails to bind in order to unwind Okazaki fragment-like hybrids, suggesting a role in the lagging strand maturation of DNA replication. The depletion or mutations of all these DNA/RNA helicases cause aberrant, unscheduled R-loop accumulation leading to dsDNA breaks [64–66]. The exposed ssDNA acts as a source of DNA damage [65,66,77], ultimately causing impairment in replication and the transcription-associated recombination of cells.

Finally, R-loops can also be affected by the DNA topology itself, wherein topoisomerase I and IIB are involved in modulating R-loop dynamics. Topoisomerase I (Top1) acts at the interface between DNA replication, transcription, and mRNA maturation. It prevents replication fork collapse by suppressing the formation of R-loops by interacting with ASF/SF2 [55]. The loss of Top1 promotes R-loop formation, especially in the 18S 5′ region of the ribosomal DNA, imposing persistent transcription blocks when RNase H is limited [56]. Topoisomerase IIIB (Top3B) is a component of the *Tudor domain-containing protein 3* (TDRD3) complex, which relaxes negatively supercoiled DNA and reduces transcription-generated R-loops [67,79]. The loss of function of the TDRD3 complex leads to an increased accumulation of R-loops resulting in abnormal chromosomal translocations of genes [84].

Altogether, a plethora of R-loop-binding or R-loop-modulating proteins reflect dedicated functions in RNA and DNA metabolism as well as R-loop organization and processing. However, there are also R-loop-binding proteins that have functions beyond these biological processes [6,28]. As one would conclude from the information above, interfer-

Table 2. Cont.

		R-Loop Binding Proteins									
		AQR	BUB3	EWSR1	RNASEH1	RNASEH2A	RNASEH2C	THOC5	TREX1	DDX5	EXOSC3
RNA processing	RBM5										
	PTBP1										
	DDX39B										
	DDX23										
	U2AF1										
	SCAF1										
	SRSF1										
	SRSF3										
	SRSF4										
	RNPS1										
	SCAF1										
	XRN2										
	DIS3L										
	SKIV2L2										
	DCP2										
	ZFP36										
	KHSRP										
	TARBP2										
	HNRNPH1										
	Epigenetics	CHD1									
PARN											
Others	MAPK9										

5. R-Loops Associated with Human Disease

The integrity of the genome and gene regulation is intensely controlled by scheduled R-loops (Figure 1). In contrast, unscheduled R-loops represent a source of DNA damage and genome instability. Hence, it is not surprising that unscheduled R-loops are increasingly linked to disorders. This includes disease entities where genome instability is an inherent element of the underlying pathophysiology (such as cancer) but also extends to other disorders including neuro-pathologies, where the intrinsic cell/tissue repair capacity is limited and hence perturbations of basic biological mechanisms are more likely to become visible. In the following, we present a few examples that document the stunning spectrum of such disorders associated with R-loops (Table 3).

A. R-loops in Nucleotide Expansion Diseases

About one million short tandem repeats are present in the human genome. These repeats are essential for genome integrity and function [90]. Unwanted expansions in

DNA short tandem repeats give rise to so-called *Nucleotide Expansion Diseases* [91]. DNA repeats vary in size from dodecamers to longer, and the threshold at which these repeats expand to become symptomatic depends on the disease [92]. Over fifty human disorders are known [90]. A significant number of these expansions cause aberrant R-loop formation. This has been linked to common genetic disorders such as amyotrophic lateral sclerosis (ALS), where expanded hexanucleotide GGGGCC repeats are found in C9orf72 and ATXN [93,94], frontotemporal dementia (FTD, with hexanucleotide GGGGCC repeats in C9orf72) [94], polyglutamine-associated ataxias [95], spinocerebellar ataxias (SCAs, with an (CAG)_n nucleotide expansion in ATXN1/2), Huntington's disease, and Friedreich's ataxia (GAA or TTC) [96–99]. R-loop accumulation is also seen in other nucleotide expansion diseases such as myotonic dystrophy (displaying CAG/CTG expansions) [100] and intellectual disability disorders such as Fragile X syndrome [101] (CGG-repeat-containing alleles of the FMR1 gene [102]). These examples suggest that alterations in cis that result in enhanced R-loop formation can be associated with various disorders, preferentially affecting neuronal cells.

Table 3. R-Loop-linked diseases and genes associated with aberrant R-loop accumulation.

Diseases	Genes Associated with R-Loops
Aging [8,103,104]	SETX
Alzheimer's [8,103–105]	SETX, WW domain-containing oxidoreductase
Aicardi-Goutières syndrome (AGS) [106]	TREX1, RNASEH2
AIDS-associated malignancies [107]	TREX complex
Amyotrophic lateral sclerosis (ALS) [30,93,94]	C9orf72 and ATXN2 (GGGCC) _n , SETX
Alternative lengthening of telomere (ALT)-dependent cancers [108]	TERRA complex
Ataxia with oculomotor apraxia (AOA2) [7,109,110]	SETX
Breast cancer [111–114]	BRAC1, BRAC2, Estrogen, SETX
Burkitt's lymphoma [84]	c-MYC, TRD3-TOP3B
Colon cancer [115,116]	VIM
Myotonic dystrophy type 1 (DM1) [100]	DMPK
Embryonal tumors with multilayered rosettes (EMTR) [27]	C19MC
Eosinophilic leukemia [15]	FIP1
Ewing's sarcoma [117]	EWS-FLI, BRCA1
Frontotemporal dementia (FTD) [94]	C9orf7 (GGGCC) _n
Fragile X syndrome type E (FRAXE) [101,102]	FRM2 (CCG) _n
Friedreich ataxia (FRDA) or fragile X syndrome type A (FRAXA) [96–99]	FXN (GAA) _n , FRM1 (CCG) _n
Huntington's disease (HD) [96–99]	HTT (CAG) _n
Infertility [118]	SETX
Multiple myeloma [84,119]	c-MYC, TRD3-TOP3B, IFN
Myelodysplastic syndromes [120]	U2AF1 (S34F), SRSF2
Polyglutamine-associated ataxias [95]	Multifactorial Nucleotide Expansion disorder
Parkinson's disease [8]	SETX
Spinocerebellar ataxias (SCAs) [39]	ATXN1/2 (CAG) _n
Immunodeficiency, centromere instability, and facial anomalies (ICF) syndrome [121]	TERRA
Wiskott–Aldrich syndrome (WAS), X-linked thrombocytopenia (XLT), and X-linked neutropenia [25]	XLT-WAS

B. R-loops in Neuronal Diseases

In addition, to repeat nucleotide expansions, alterations in trans (in R-loop modifying components, such as SETX [35] or the THO/TREX complex [7]) can lead to aberrant R-loop accumulation and result in similar phenotypes. For example, recessive mutations in senataxin (SETX, a protein that protects the replication fork across the RNA-Polymerase-II-transcribed gene and unwinds unnecessary R-loops) cause R loop accumulation [35] in a particular type of ALS (ALS4) and Ataxia with oculomotor apraxia type 2 (AOA2) [7,109,110]. These mutations are also linked to aging and other neurodegenerative disorders, such as Parkinson's and Alzheimer's diseases [8,103,104].

In Aicardi–Goutières syndrome (AGS), mutations in RNASEH2 and TREX1 result in an increase in R-loop accumulation [106]. This demonstrates that RNASEH2 and TREX1, acting as an exonuclease as part of the THO/TREX complex, can have a clinical impact on R-loop resolution [7]. By analogy, the downregulation of the WW domain-containing oxidoreductase (proteins that are responsible for regulating transcription–replication collisions and preventing unwanted R-loop accumulation) is responsible for Alzheimer's disease, showing aberrant R-loop accumulation [105]. These examples highlight the functional importance of the components that are involved in the formation and resolution of R-loops.

C. R-loops in Cancer

Genome instability is a characteristic of most cancers. Hence, it is not surprising that a fraction of cancers including AIDS-associated malignancies [107] and *Alternative lengthening of telomere* (ALT)-dependent cancers [108] are linked to R-loop accumulation in *Telomeric DNA and long noncoding RNA* (TERRA). TERRA is a key mediator of the *Alternative Lengthening of Telomeres* (ALT) pathway due to a dysfunction of RNase H1 [107,108]. In Estrogen (ET)-enriched breast cancer (BC), R-loop accumulation is highly enriched at E2-responsive genomic loci, resulting in E2-dependent R-loop-driven DNA damage [111]. In other BC types, BRCA1/SETX complexes cause R-loop-driven DNA damage [112,113]. The depletion of BRCA2 increases R-loop accumulation [122], which acts as a chief source of replication stress and cancer-associated instability in BC [114]. In multiple myeloma and Burkitt's lymphoma, the depletion of TDRD3 (a protein responsible for gene transcription by interacting with TopIIIB, see above) exhibits elevated R-loop accumulation at the c-MYC locus in B cells, resulting in DNA damage and frequent chromosomal translocations. In eosinophilic leukemia, the truncated CPA component of FIP1L1 (mRNA 3' end-processing factor interacting with PAPOLA and CPSF1) results in R-loop accumulation, thereby causing DNA damage and chromosome breakage [15]. In Ewing Sarcoma (an aggressive pediatric cancer of the bone and soft tissue), alterations of damage-induced transcription by the EWSR1 protein cause increased replication stress due to elevated R-loop accumulation [117]. All these examples highlight that R-loop accumulation acts as a source of DNA damage and causes the dysfunction of gene regulation and genomic instability in various cancers.

D. R-loops in other diseases

Apart from neuronal disorders and various cancers, unscheduled R-loop accumulation is also found in other diseases. In Immunodeficiency, Centromere instability, and Facial anomalies (ICF) syndrome, cells that exhibit short telomeres and elevated TERRA levels are enriched with R-loops at telomeric regions throughout the cell cycle. These telomeric R-loop hybrids are associated with high levels of DNA damage at the chromosome ends [121]. In *Ataxia with oculomotor apraxia* (AOA2) patients (caused due to mutations in SETX, a component with RNA helicase activity responsible for resolving R-loops), R-loop accumulation-induced DNA damage in cells undergoing spermatogenesis has been shown to cause sterility in males [118]. In *Myelodysplastic Syndromes* (MDS), caused by splicing factor mutations in components such as SRSF2 and U2AF1, aberrant R-loop accumulation is linked to the compromised proliferation of bone-marrow-derived blood progenitors, a characteristic feature of MDS [120].

The above-mentioned examples and further disorders [4–8] associated with unscheduled R-loops highlight the importance of mechanisms that prevent the formation and/or

promote the resolution of unscheduled R-loop structures. While R-loops are not necessarily the cause for all the disease entities described so far, they may perpetuate and even aggravate the initially disease-eliciting pathophysiology. In either case, the abundance and association of R-loops with disease make them interesting structures for diagnostics.

6. R-Loops as a Diagnostic Biomarker?

Emerging reports suggest that accumulating unscheduled R-loops can be used for diagnostics and stratifying patients. For example, unscheduled R-loop accumulation is considered a novel molecular defect that is causative of TH1 immunodeficiency and genomic instability in patients with *Wiskott–Aldrich syndrome* (WAS) [113] and WAS-related disorders such as *X-linked thrombocytopenia* (XLT) and *X-linked neutropenia* [123]. In addition, the accumulation of undesired R-loops has been suggested to represent a potential biomarker for determining the prognostic outcomes in the XLT-WAS clinical spectrum [25]. Whole-genome sequencing of 193 primary *Embryonal tumors with multilayered rosettes* (ETMRs) revealed that unscheduled R-loop structures are widespread across these tumors due to a loss of DICER1 function (DICER1 promotes transcription termination at sites of transcription–replication collisions with DNA damage [27]), causing genomic instability. The targeting of R-loops with topoisomerase and poly ADP ribose polymerase inhibitors showed promising treatment strategies for this deadly disease [27]. Aberrant, unscheduled R-loop accumulation in *Uterine fibroid* (UF) patients (benign monoclonal neoplasms of the myometrium, which represent the most frequent non-cutaneous tumors) was demonstrated to specify potentially malignant tumor progression in the dominant UF subtype [124] (Figure 2A). In a study on *Epidermal growth factor receptor variant III* (EGFRvIII)-positive glioblastoma (GBM) patients, increased R-loop accumulation and genome instability caused by replication stress have been suggested to be associated with tumor heterogeneity and allow for the stratification of patients for individualized therapeutic approaches [125] (Figure 2B). In *Multiple Myeloma* (MM), elevated levels of R-loops and the failure to resolve R-loops can cause the sustained activation of a systemic inflammatory response characterized by the interferon (IFN) gene expression signature, which results in a poor prognosis [126]. Along with an elevated expression of para-speckle genes, R-loops have been identified to correlate with MM's progression [119]. Altogether, unscheduled R-loop accumulation could be used to identify distinct biological properties (heterogeneity) in tumors [119,124,125], can guide the monitoring of the severity of disorders [25], and has a strong prognostic potential.

Beyond the identification of unscheduled R-loops, the abundance of R-loop-binding proteins (Table 1) can also be exploited as a potential biomarker. In a study by Boros-Oláh et al. in 2019, a systematic pharmacogenomic analysis was performed to test the drugging and diagnostic potential of an R-loop and its binding proteins in 33 cancer types [28]. R-loop-binding proteins of various categories with defined R-loop functions were selected for detailed analyses in this study (including proteins such as AQR, ATXN1/2, BLM, BRCA1/2, BUB3, BUGZ (ZNF207), CARM1, DDX19A, DHX9, EWSR, FANCD2, FANCM, GADD45A, PIF1, PRMT1, RNASEH1, RNASEH2, RTEL1, SETX, SRSF1, SRSF2, THO/TREX, TOP1, TOP3B, and U2AF1). Among these, the mRNA expression of BUB3, DHX9, PRMT1, THOC4, THOC7, U2AF1, and ZNF207 (BUGZ) were found to increase in several primary tumors compared to healthy tissues, while SRSF1 (ASF/SF2) was downregulated in most cancers (except for acute myeloid leukemia—LAML). Moreover, the expression of some of these R-loop-binding proteins was found to be directly linked to patient survival in various cancers, e.g., a low level of RNASEH2A, THOC6, PRMT1, and PIF1 were associated with prolonged survival in mesothelioma patients (MESO), while a low FANCM mRNA level appeared advantageous for breast cancer survival (Figure 2C). In another study, the depletion of TDP-43 (an R-loop-binding protein involved in RNA processing and with structural resemblance to heterogeneous ribonucleoproteins) in *Amyotrophic Lateral Sclerosis* (ALS) was linked to increased numbers of R-loops and DNA damage, which has been reported as a potential guide for developing ALS therapies [127]. Altogether, these examples highlight that R-loops and their binding proteins could represent potentially helpful biomarkers.

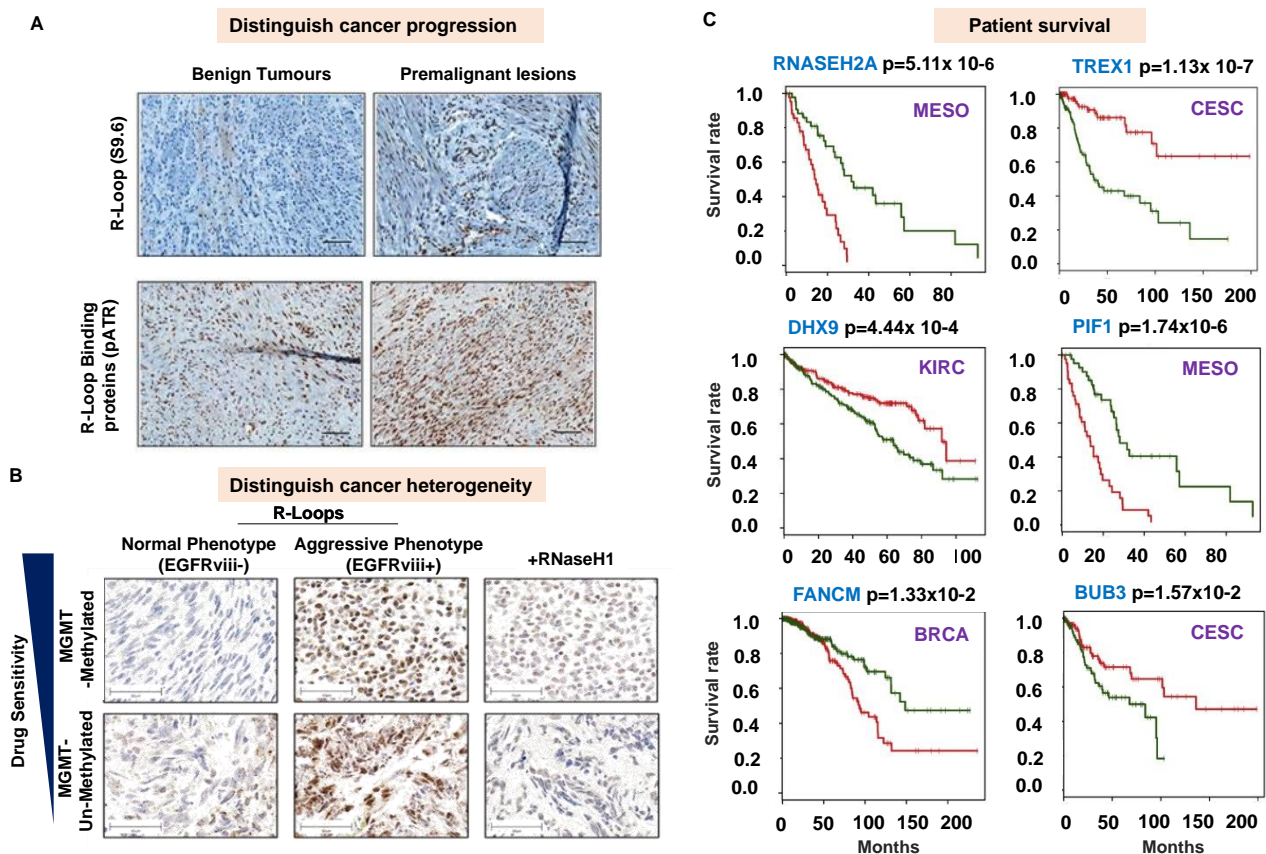


Figure 2. R-loops and R-loop-binding proteins as biomarkers. (A). Benign and Premalignant tumor Uterine fibroids are tissue-specifically characterized by aberrant R-loop accrual and its binding protein. Immunohistochemistry (IHC) staining with (upper panel) R-loop-specific S9.6 and (lower panel) pATR (Ser428)-specific antibodies indicating phosphorylated (activated) replication stress signaling correlative with R loop accumulation in different mutants (Re-adapted from Ref [124]). (B). R loop accumulation used for distinguishing tumor heterogeneity in Glioblastoma (GBM) patients, highlighting distinct biological properties and complementing the readout of established clinical marker for GBM (O6-methylguanine (O6-MeG)-DNA methyltransferase (MGMT)). IHC staining of nonaggressive versus aggressive phenotype GBM. Epidermal growth factor receptor variant III (EGFRvIII) areas are depicted in the left column, and EGFRvIII+ areas are displayed in the middle column. Aggressive phenotype GBM incubated with RNase H1 (shown in right column) before immunohistochemical staining with S9.6 antibody is used as specificity control (removing R-loops) for staining, indicating that increased EGFRvIII expression is associated with an increase in R loop accumulation predisposed to DNA damage and genomic instability (Re-adapted with Creative Commons CC-BY-NC license permissions from [126]). (C). Representative Kaplan–Meier curves showing the overall survival probability in various cancer types (Mesothelioma (MESO), Cervical and endocervical cancers (CESC), Kidney renal clear cell carcinoma (KIRC), and Breast invasive carcinoma (BRCA)) depending on the gene expression of various R-loop-binding proteins (high and low expression shown in red and green, respectively; R-loop genes, cancer IDs, and p -values are indicated (Re-adapted from [28])).

7. Emerging Technologies for Detecting R-Loops

As R-loops and their binding proteins show promising diagnostic potential, it is important to discuss the state-of-the-art technologies that can be used to detect R-loops. In 1980, the first technique for identifying R-loops was established using an antibody (S9.6) recognizing DNA–RNA hybrids, which revolutionized the R-loop field [128]. DNA–RNA Immunoprecipitation (DRIP) using the S9.6 antibody was further developed in vivo to uncover the contribution of R-loops to biological processes [56,70,129–131]. With the advent

of next-generation sequencing, several techniques now exist to map the distribution, size, and dynamic changes of R-loops.

Globally, wet-lab techniques detect the accumulation of R-loops by foot-printing or pull-down assays probed with the R-loop-binding S9.6 antibody (Table 4). DRIP sequencing (DRIP-seq) is a widely used technique for the genome-wide profiling of R-loops. It utilizes the sequence-independent but high structure-specificity and affinity of the S9.6 monoclonal antibody to capture R-loops for the large-scale parallel DNA sequencing of the genomic fragments containing DNA–RNA hybrids [31]. However, this technique has some limitations, including bias and resolution limits because of the fragmentation of the genome using restriction enzymes, limited strand sensitivity, and a decreased feasibility with respect to conducting a quantitative analysis to study the genomic distribution. In addition, it does not discriminate between the R-loop sequence and the surrounding elements. To address these limitations, variants of DRIP have been developed. S1-DRIP-seq uses sonication for fragmentation where the displaced ssDNA is removed prior to sonication using S1 nuclease [132]. Although S1-DRIP generates a readout in high resolution (compared to DRIP), the technique’s reproducibility can be sometimes challenging, which currently limits the ‘clinical’ use of this technique (the S1 nuclease is delicate, and the reaction is difficult to control to obtain reproducible data [133]).

Table 4. Diagnostic tools used to detect R-loops and R-loop-binding proteins.

	Detection Method	Method Name	Processing Method	Advantages	Disadvantages
Wet lab techniques					
Techniques for detecting R loop	S9.6 antibody staining DNA/RNA hybrid	Immunoprecipitation (IP)/ Immunohistochemistry (IHC) [128].	Immunostaining of DNA–RNA hybrid	Good signal, likely useful for the analysis of samples from tissue banks	Limited to the microscopic examination of R-loops
		DRIP [56,70,129–131]	Restriction digestion (RE) of genome followed by IP	Robust signal	Better resolution than IP/IHC but is still low
		DRIP-seq [31]	RE of genome followed by IP and dsDNA sequencing	Robust signal is widely adopted, and is easy to set up	Low resolution, no strand specificity, and cannot be used in situ
		S1-DRIP-seq [132]	Sonication of samples followed by IP and dsDNA sequencing	Higher resolution than DRIP-seq	No strand specificity and cannot be used in situ. S1 nuclease is delicate, and it is difficult to control the reaction, which may make it challenging to reproduce the data in clinical setting
		DRIPc-seq [133]	RE of genome followed by RNA sequencing	Strand-specific, high resolution	Not in situ, requires longer sample preparation, S9.6 may recognize dsRNA
		RDIP-seq [134]	Sonication of genome followed by RNA sequencing		Not in situ, tedious preparation
		Bis-DRIP-seq [135]	RE of genome followed by sequencing dsDNA with bisulfite conversions	Strand-specific, provides additional control to ensure S9.6 signal arises from an R-loop in situ	Requires many replicates and shows R-loop enrichment in promoter regions only
		qDRIP [136]	RE of genome followed by IP of DNA–RNA hybrid and synthetic DNA/RNA hybrid used as internal standards followed by dsDNA sequencing and quantification using internal standards as a reference.	Internal standards help with high-resolution, strand-specific sequencing	Spikes in hybrids shorter than 150 bp are unlikely to be useful for normalization. Additional spike-in may be required

Table 4. Cont.

	Detection Method	Method Name	Processing Method	Advantages	Disadvantages
		SMRF-seq [137]	RE of genome followed by IP of DNA/RNA hybrid and sequencing of dsDNA with bisulfite conversions at single molecule level	Strand-specific, single-molecule resolution, avoids biases inherent to read-count normalization by accurately profiling signals in regions unaffected by transcription inhibition thus providing accurate differential peak calling between conditions	As with any foot-printing method, SMF is agnostic to the distinguishing of DNA-binding proteins creating the footprints.
	Catalytically inactive RNase H	DRIVE-seq [31]	RE of genome followed by targeting catalytically inactive RNaseHs and dsDNA sequencing	Provides independent verification of some DRIP-seq results	Low enrichment, low resolution, reagent not commercially available, no strand specificity, not in situ
		R-ChIP-seq [138]	Sonication followed by targeting catalytically inactive RNaseHs and ssDNA sequencing	Strand specific, in situ capture	Cell line must be engineered to express catalytically inactive RNase H construct, inactive RNase H may alter hybrid dynamics
	RNase H to guide micrococcal nuclease to R-loops	MapR [139]	Antibody-independent R-loop-profiling technique that utilizes RNase H to guide micrococcal nuclease to R-loops, which are subsequently cleaved, released, and identified by sequencing	Heavily based on CUT&RUN, a new and fast method to identify transcription factor binding sites genome-wide	Does not discriminate between the template and non-template strands and, therefore, cannot identify which DNA strand is involved in DNA–RNA hybrid formation.
	Sensor that binds to R loop	R loop CUT&Tag [140]	Combines CUT&Tag and GST-His6-2×HBD (glutathione S-transferase–hexahistidine–2× hybrid-binding domain) tags as an artificial R loop hybrid sensor to specifically recognize the DNA–RNA hybrids.	Sensitive, reproducible and generates good resolution to sense the R-loop instead of capture strategies that largely contribute to disparities in the previous techniques including R-loop Mapping	Current form of R-loop CUT&Tag does not provide strand information about R loops
Techniques for detecting R loop and R loop-binding proteins	Fusion protein that binds to hybrid-binding domain (HBD) of RNaseH1 and an engineered variant of ascorbate peroxidase	RDProx (RNA–DNA Proximity Proteomics) [141]	Provides a snapshot of the R-loop-proximal proteome	In vivo labelling of R-loop-proximal proteins is performed, difficult to solubilize proteins that are amenable to the analysis can be identified, even transient spatiotemporal interactions with low affinity and transient interactions are detected.	Unable to distinguish between direct protein-binding or indirect proteins associated with RNA
	Staining of both DNA–RNA hybrid (S9.6 antibody)+ R-loop-binding proteins (Antibody specific to the desired protein)	IP/IHC (R-loops+ R-Loop-binding proteins) [124]	Immunostaining of DNA–RNA hybrid and their binding proteins	Fast analysis of pathology specimens	Low resolution and limited to routine microscopic analysis

Table 4. Cont.

	Detection Method	Method Name	Processing Method	Advantages	Disadvantages
Bioinformatics tools and databases					
Techniques for detecting R loop		QmRLFS-finder [142]	Identifies three structural features of R loop including a short G-cluster-rich region (R-loop initiation zone or RIZ), a structurally non-specified linker (linker), and long downstream region that has high G-density R-loop elongation zone (or REZ) based on experimental data	User-friendly web server and stand-alone tool for rapid and accurate prediction of RLFSs in DNA or RNA sequences shows strong agreement with existing genes and genome-scale experimentally determined R-loops	Information is limited and an updated version needs to be integrated with the growing experimental data
		R loop tracker [143]			
	Structure-based detection and prediction based on existing wet-lab data	R-loop atlas [144]	About 63 million peaks called from 254 plant species by ssDRIP-seq and deepR-loopPre are available	User-friendly web server for plants species based on experimental data	Limited to plant species only
Techniques for R loops and their binding proteins		R-loop DB [145]	Consists of computationally predicted R-loop-forming sequences (RLFSs) in human genic regions. Using the QmRLFS, the updated version of this database now has an increased number of RLFSs predicted in the human genes and in the genomes of other organisms	Provides comprehensive annotation of Ensembl RLFS-positive genes to study comparative evolution and genome-scale analyses, also R loop-binding proteins	Limited information and an updated version needs to be integrated with the growing experimental data

To address both the strand specificity and resolution of DRIP-RNA-seq [146], DRIPc-seq [133] and RDIP-seq [134] are techniques that have been established wherein the RNA components of R-loops are sequenced instead of the DNA. BisDRIP sequencing (bisDRIP-seq) allows for researchers to distinguish between the loops themselves and the surrounding DNA. It uses bisulfite to selectively convert cytosine residues into uracil residues within genomic DNA regions that contain single-stranded DNA. BisDRIPseq thereby allows for the mapping of R-loops at a near-nucleotide resolution to identify single-stranded regions based on the preferential labelling of one strand of the DNA and the requirement that the labelling is transcription-dependent [135]. SMRF-seq (Single-Molecule R-loop Foot-printing) is a technique consisting of the modified bisulfite-based mapping of the extruded single-strand DNA in an R-loop adapted for single-molecule long-read sequencing [137]. qDRIP is a quantitative differential DNA–RNA immunoprecipitation method combining synthetic DNA–RNA hybrid internal standards with high-resolution, strand-specific sequencing. It avoids biases inherent to read-count normalization by accurately profiling signals in regions unaffected by transcription inhibition. It thereby provides accurate differential peak calling between perturbed versus control conditions to obtain previously unattainable biological insights [136].

Alternative approaches to S9.6-based methods are DRIVE-seq, R-ChIP-seq, and MapR, which take advantage of catalytically inactive RNase H binding to R-loops. DRIVE-seq is conceptually similar to DRIP-seq as it utilizes the pull-down of hybrids through catalytically inactive RNase H instead of S9.6 [31]. R-ChIP is another method, and it can be used for the genome-wide profiling of R-loops and is designed to sequence the 5' end of the template strand DNA, thus making the analysis different from that of typical ChIP-seq [138]. Both bis-DRIP and R-CHIP require an in situ step and show highly concentrated signals at the promoters of genes, but barely any signal at the 3' end of genes. Other variants of DRIP show a strong signal at the promoter's position and an appreciable signal in the gene body and at termination regions [33,110]. It is under debate whether the in situ step-containing technique is less sensitive to other regions or whether other variants of DRIP yield artifactual or non-specific signals [9]. Finally, Map-R is a fast, antibody-independent R-loop-profiling technique that utilizes RNase H to guide micrococcal nuclease to R-loops,

which are subsequently cleaved, released, and identified by sequencing. It provides an output allowing for genome wide coverage with a low level of input material in a fraction of the time and with high sensitivity [139].

In addition, R-loop CUT&Tag sequencing has been established as a sensor-based technique used to overcome large discrepancies in R-loop mapping so as to provide an accurate and comprehensive profile of native R-loops across the genome. Some of its discrepancies arise due to fragmentation bias by restriction enzymes, disparate specificities of RNase H1 and S9.6 to R loops, or differences in sequencing and capture strategies, such as R-loop capture *ex vivo* or *in situ*. Moreover, full-length recombinant catalytically inactive RNase H1 is not very efficient in affinity pulldowns [8,12], which is the principle of DRIP-related R loop-mapping methods (see above). Therefore, R-loop-mapping methods that are independent of S9.6 or catalytically inactive RNase H1 are urgently needed to resolve controversies. R-loop CUT&Tag combines CUT&Tag and GST-His6-2×HBD (glutathione S-transferase–hexahistidine–2× hybrid-binding domain) tags as an artificial R-loop hybrid sensor to specifically recognize DNA–RNA hybrids. It is sensitive and generates good resolution for sensing the R-loop as compared to capture strategies that largely contribute to disparities in the previous techniques [140].

Very few techniques are available for characterizing both R-loops and their binding proteins. Recently, R-loop proximity proteomics has been developed to identify proteins that bind to R-loops and regulate them. RNA–DNA Proximity Proteomics (RDProx) is a technique that enables the mapping of an R-loop using the fusion protein of the hybrid-binding domain (HBD) of RNaseH1 and an engineered variant of ascorbate peroxidase (APEX2) [11]. It allows for the characterization of transient interactions of the proteome with the R-loop in a spatiotemporal manner [141].

As shown in Figure 2A,B, the Immunoprecipitation (IP) and Immunohistochemistry (IHC) of different tissue samples can be used by employing the S9.6 antibody complemented with an R-loop-binding protein antibody in routine diagnostics for preliminary insights.

While several experimental methods are now available to detect, quantify, and study R-loop dynamics, the structural and functional characterization of an R-loop still remains a major challenge for developing new (clinically relevant) therapeutic probes. Along these lines, studies that structurally characterize and target R-loops are currently under development [147] with the help of systematic prediction and detection pipelines. Bioinformatics tools such as QmRLFS-finder, R-loop tracker, and databases such as the R-loop Atlas, and R-loop base are next-generation, *in silico* approaches developed to characterize R-loops. The Quantitative Model of the R-loop-Forming Sequence (RLFS) finder (QmRLFS-finder) [142] is a web-based server that predicts RLFSs based on experimentally supported structural models of RLFSs in RNA/DNA sequences. It demonstrates highly accurate predictions of the RLFSs detected. The R-loop tracker tool is a similar web-based server that focuses on the prediction of R-loops in genomic DNA with an unlimited input size [143]. It allows for the cross-evaluation of *in silico* results with experimental data, if available, and helps correlate these with other genomic features and markers with an enhanced visualization output [143]. In 2017, R-loopDB, a database that contains computationally predicted RLFSs in human genetic regions, was developed. Using QmRLFS, the updated version of this database now has an increased the number of RLFSs predicted in human and other genomes [145]. It also provides a comprehensive annotation of Ensembl RLFS-positive genes for studying comparative evolution and genome-scale analyses in R-loop biology. The R-loop Atlas is a database harboring about 63 million peaks collected from 254 plant species by ssDRIP-seq and deepR-loopPre (a deep-learning tool for predicting locations and profiles of strand-specific R loops; <http://bioinfor.kib.ac.cn/R-loopAtlas/index.html>, accessed on 18 November 2022) [144]. R-loop base is another database that includes a reference set of human R-loop zones for high-confidence R-loop localization and for spotting conserved genomic features that are associated with R-loop formation. The data are cured in a comprehensive manner by integrating knowledge from multi-omics analyses

and literature mining. A list of R-loop-regulatory proteins and their targeted R-loops in multiple species, to date, can be obtained [148].

Overall, an impressive spectrum of techniques has been developed aiming at elucidating the R-loop distribution genome-wide and its function. These techniques are expected to enrich “clinical” diagnostics and drug discovery in the R-loop field. However, from a diagnostic point of view, there is still a scope for further technical improvements. For example, it would be desirable to acquire techniques that can differentiate between scheduled and unscheduled R-loops, techniques that permit the detection of R-loops from circulating cells, or techniques that allow for R-loop detection with a nucleotide resolution from widely available clinical (paraffin-embedded) material. Furthermore, it would be desirable to develop and advance techniques that allow for the (large-scale) parallel profiling of R-loops in a reliable manner, which in turn may help the identification of disease-specific R-loop-prone loci. Such techniques would foster a more comprehensive understanding of the functional importance and diagnostic potential of R-loops in the context of human disorders. They may also yield complementary biological insights and help us to understand whether R-loop perturbations almost exclusively manifest in pathologies affecting cells with a limited regenerative potential (such as neuronal cells). Given the broad functional spectrum of R-loops in genome organization, gene regulation, and genome integrity, it seems possible that perturbations of R-loop biology could also confer detrimental effects during early (human) development. Such effects are not normally searched for, e.g., in human fetal tissue, and hence the further evolution of technologies used to detect R-loops could provide novel insights into the biology and consequences of scheduled and unscheduled R-loops. Finally, generalized guidelines would be necessary for the clinical use of easy and reproducible techniques.

8. Conclusions

Considerable progress has been made in understanding the complexity and regulation of R-loops and their resulting biological functions. Although the perturbations of R-loops have not yet been exhaustively studied beyond the disease entities mentioned in this review, aberrant R-loops appear to trigger or even perpetuate a variety of disorders. R-loops represent a valuable resource that reflects the actual state in real-time and in specific situations, e.g., during development, maturation, and in pathophysiological conditions, making them exciting structures for diagnosis and informing clinical decisions. Burgeoning R-loop-screening technologies are expected to fuel the current and future techniques towards their use in clinical diagnostics and precision medicine.

Author Contributions: E.S.K. and S.D. performed the literature survey and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: ESK thanks SPMD 2021 grant from the German Society of Clinical and Laboratory Medicine and University Medical Center of the Johannes Gutenberg University intra-university research funding 2021. Our work in the Danckwardt lab was kindly supported by the German Research Foundation Priority Program SPP 1935, German Research Foundation grants DA 1189/2-1 and DA 1189/5-1, Germany Research Foundation Major Research Instrumentation Program INST 371/33-1, the German Research Foundation graduate school GRK 1591, Federal Ministry of Education and Research (BMBF01EO1003), the German Society of Clinical and Laboratory Medicine, and the Hella Bühler Award for Cancer Research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We apologize to those whose works were not cited here due to space limitations.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Santos-Pereira, J.M.; Aguilera, A. R loops: New modulators of genome dynamics and function. *Nat. Rev. Genet.* **2015**, *16*, 583–597. [[CrossRef](#)] [[PubMed](#)]
2. Sanz, L.A.; Hartono, S.R.; Lim, Y.W.; Steyaert, S.; Rajpurkar, A.; Ginno, P.A.; Xu, X.; Chédin, F. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol. Cell* **2016**, *63*, 167–178. [[CrossRef](#)]
3. Allison, D.F.; Wang, G.G. R-loops: Formation, function, and relevance to cell stress. *Cell Stress* **2019**, *3*, 38–46. [[CrossRef](#)] [[PubMed](#)]
4. Hegazy, Y.A.; Fernando, C.M.; Tran, E.J. The balancing act of R-loop biology: The good, the bad, and the ugly. *J. Biol. Chem.* **2020**, *295*, 905–913. [[CrossRef](#)]
5. Aguilera, A.; García-Muse, T. R loops: From transcription byproducts to threats to genome stability. *Mol. Cell* **2012**, *46*, 115–124. [[CrossRef](#)] [[PubMed](#)]
6. Groh, M.; Gromak, N. Out of balance: R-loops in human disease. *PLoS Genet.* **2014**, *10*, e1004630. [[CrossRef](#)] [[PubMed](#)]
7. Mackay, R.P.; Xu, Q.; Weinberger, P.M. R-Loop Physiology and Pathology: A Brief Review. *DNA Cell Biol.* **2020**, *39*, 1914–1925. [[CrossRef](#)]
8. Skourti-Stathaki, K.; Proudfoot, N.J. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* **2014**, *28*, 1384–1396. [[CrossRef](#)]
9. Crossley, M.P.; Bocek, M.; Cimprich, K.A. R-Loops as Cellular Regulators and Genomic Threats. *Mol. Cell* **2019**, *73*, 398–411. [[CrossRef](#)] [[PubMed](#)]
10. Niehrs, C.; Luke, B. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 167–178. [[CrossRef](#)]
11. Mosler, T.; Conte, F.; Longo, G.M.C.; Mikicic, I.; Kreim, N.; Möckel, M.M.; Petrosino, G.; Flach, J.; Barau, J.; Luke, B.; et al. R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat. Commun.* **2021**, *12*, 7314. [[CrossRef](#)] [[PubMed](#)]
12. Lam, F.C.; Kong, Y.W.; Huang, Q.; Vu Han, T.L.; Maffa, A.D.; Kasper, E.M.; Yaffe, M.B. BRD4 prevents the accumulation of R-loops and protects against transcription-replication collision events and DNA damage. *Nat. Commun.* **2020**, *11*, 4083. [[CrossRef](#)] [[PubMed](#)]
13. Huertas, P.; Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **2003**, *12*, 711–721. [[CrossRef](#)] [[PubMed](#)]
14. Skourti-Stathaki, K.; Proudfoot, N.J.; Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **2011**, *42*, 794–805. [[CrossRef](#)] [[PubMed](#)]
15. Stirling, P.C.; Chan, Y.A.; Minaker, S.W.; Aristizabal, M.J.; Barrett, I.; Sipahimalani, P.; Kobor, M.S.; Hieter, P. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev.* **2012**, *26*, 163–175. [[CrossRef](#)] [[PubMed](#)]
16. Li, X.; Manley, J.L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **2005**, *122*, 365–378. [[CrossRef](#)] [[PubMed](#)]
17. Herrera-Moyano, E.; Mergui, X.; García-Rubio, M.L.; Barroso, S.; Aguilera, A. The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. *Genes Dev.* **2014**, *28*, 735–748. [[CrossRef](#)] [[PubMed](#)]
18. Wahba, L.; Amon, J.D.; Koshland, D.; Vuica-Ross, M. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* **2011**, *44*, 978–988. [[CrossRef](#)]
19. Richard, P.; Manley, J.L. R Loops and Links to Human Disease. *J. Mol. Biol.* **2017**, *429*, 3168–3180. [[CrossRef](#)]
20. Wells, J.P.; White, J.; Stirling, P.C. R Loops and Their Composite Cancer Connections. *Trends Cancer* **2019**, *5*, 619–631. [[CrossRef](#)]
21. Rakshit, S.; Sunny, J.S.; George, M.; Hanna, L.E.; Sarkar, K. R-loop modulated epigenetic regulation in T helper cells mechanistically associates coronary artery disease and non-small cell lung cancer. *Transl. Oncol.* **2021**, *14*, 101189. [[CrossRef](#)] [[PubMed](#)]
22. Barchi, M.; Bielli, P.; Dolci, S.; Rossi, P.; Grimaldi, P. Non-Coding RNAs and Splicing Activity in Testicular Germ Cell Tumors. *Life* **2021**, *11*, 736. [[CrossRef](#)]
23. Cuartas, J.; Gangwani, L. R-loop Mediated DNA Damage and Impaired DNA Repair in Spinal Muscular Atrophy. *Front. Cell. Neurosci.* **2022**, *16*, 826608. [[CrossRef](#)]
24. Perego, M.G.L.; Taiana, M.; Bresolin, N.; Comi, G.P.; Corti, S. R-Loops in Motor Neuron Diseases. *Mol. Neurobiol.* **2019**, *56*, 2579–2589. [[CrossRef](#)] [[PubMed](#)]
25. Sarkar, K.; Han, S.S.; Wen, K.K.; Ochs, H.D.; Dupré, L.; Seidman, M.M.; Vyas, Y.M. R-loops cause genomic instability in T helper lymphocytes from patients with Wiskott-Aldrich syndrome. *J. Allergy Clin. Immunol.* **2018**, *142*, 219–234. [[CrossRef](#)] [[PubMed](#)]
26. Liu, Z.; Gao, X.; Zhou, Z.; Kang, S.W.; Yang, Y.; Liu, H.; Zhang, C.; Wen, Z.; Rao, X.; Wang, D.; et al. San1 deficiency leads to cardiomyopathy due to excessive R-loop-associated DNA damage and cardiomyocyte hypoplasia. *Biochim. Biophys. Acta Mol. Basis Dis.* **2021**, *1867*, 166237. [[CrossRef](#)] [[PubMed](#)]
27. Lambo, S.; Gröbner, S.N.; Rausch, T.; Waszak, S.M.; Schmidt, C.; Gorthi, A.; Romero, J.C.; Mauermann, M.; Brabetz, S.; Krausert, S.; et al. The molecular landscape of ETMR at diagnosis and relapse. *Nature* **2019**, *576*, 274–280. [[CrossRef](#)] [[PubMed](#)]
28. Boros-Oláh, B.; Dobos, N.; Hornyák, L.; Szabó, Z.; Karányi, Z.; Halmos, G.; Roszik, J.; Székvölgyi, L. Drugging the R-loop interactome: RNA-DNA hybrid binding proteins as targets for cancer therapy. *DNA Repair* **2019**, *84*, 102642. [[CrossRef](#)]
29. Spada, S.; Luke, B.; Danckwardt, S. The Bidirectional Link Between RNA Cleavage and Polyadenylation and Genome Stability: Recent Insights From a Systematic Screen. *Front. Genet.* **2022**, *13*, 854907. [[CrossRef](#)] [[PubMed](#)]

30. Grunseich, C.; Wang, I.X.; Watts, J.A.; Burdick, J.T.; Guber, R.D.; Zhu, Z.; Bruzel, A.; Lanman, T.; Chen, K.; Schindler, A.B.; et al. Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. *Mol. Cell* **2018**, *69*, 426–437.e7. [[CrossRef](#)] [[PubMed](#)]
31. Ginno, P.A.; Lott, P.L.; Christensen, H.C.; Korf, I.; Chédin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **2012**, *45*, 814–825. [[CrossRef](#)] [[PubMed](#)]
32. Ariel, F.; Lucero, L.; Christ, A.; Mammarella, M.F.; Jegu, T.; Veluchamy, A.; Mariappan, K.; Latrasse, D.; Blein, T.; Liu, C.; et al. R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Mol. Cell* **2020**, *77*, 1055–1065.e4. [[CrossRef](#)]
33. Chen, P.B.; Chen, H.V.; Acharya, D.; Rando, O.J.; Fazio, T.G. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.* **2015**, *22*, 999–1007. [[CrossRef](#)]
34. Arab, K.; Karaulanov, E.; Musheev, M.; Trnka, P.; Schäfer, A.; Grummt, I.; Niehrs, C. GADD45A binds R-loops and recruits TET1 to CpG island promoters. *Nat. Genet.* **2019**, *51*, 217–223. [[CrossRef](#)]
35. Skourti-Stathaki, K.; Kamieniarz-Gdula, K.; Proudfoot, N.J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **2014**, *516*, 436–439. [[CrossRef](#)] [[PubMed](#)]
36. Powell, W.T.; Coulson, R.L.; Gonzales, M.L.; Crary, F.K.; Wong, S.S.; Adams, S.; Ach, R.A.; Tsang, P.; Yamada, N.A.; Yasui, D.H.; et al. R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13938–13943. [[CrossRef](#)]
37. Nakama, M.; Kawakami, K.; Kajitani, T.; Urano, T.; Murakami, Y. DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation. *Genes Cells* **2012**, *17*, 218–233. [[CrossRef](#)]
38. Castellano-Pozo, M.; Santos-Pereira, J.M.; Rondón, A.G.; Barroso, S.; Andújar, E.; Pérez-Alegre, M.; García-Muse, T.; Aguilera, A. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **2013**, *52*, 583–590. [[CrossRef](#)]
39. Yu, K.; Chedin, F.; Hsieh, C.L.; Wilson, T.E.; Lieber, M.R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* **2003**, *4*, 442–451. [[CrossRef](#)]
40. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821. [[CrossRef](#)]
41. D’Alessandro, G.; Whelan, D.R.; Howard, S.M.; Vitelli, V.; Renaudin, X.; Adamowicz, M.; Iannelli, F.; Jones-Weinert, C.W.; Lee, M.; Matti, V.; et al. BRCA2 controls DNA: RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nat. Commun.* **2018**, *9*, 5376. [[CrossRef](#)]
42. Cohen, S.; Puget, N.; Lin, Y.L.; Clouaire, T.; Aguirrebengoa, M.; Rocher, V.; Pasero, P.; Canitrot, Y.; Legube, G. Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nat. Commun.* **2018**, *9*, 533. [[CrossRef](#)]
43. Li, L.; Germain, D.R.; Poon, H.Y.; Hildebrandt, M.R.; Monckton, E.A.; McDonald, D.; Hendzel, M.J.; Godbout, R. DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol. Cell. Biol.* **2016**, *36*, 2794–2810. [[CrossRef](#)]
44. Britton, S.; Dernoncourt, E.; Delteil, C.; Froment, C.; Schiltz, O.; Salles, B.; Frit, P.; Calsou, P. DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Res.* **2014**, *42*, 9047–9062. [[CrossRef](#)] [[PubMed](#)]
45. Ohle, C.; Tesorero, R.; Schermann, G.; Dobrev, N.; Sinning, I.; Fischer, T. Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* **2016**, *167*, 1001–1013.e7. [[CrossRef](#)]
46. Yasuhara, T.; Kato, R.; Hagiwara, Y.; Shiotani, B.; Yamauchi, M.; Nakada, S.; Shibata, A.; Miyagawa, K. Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. *Cell* **2018**, *175*, 558–570.e11. [[CrossRef](#)]
47. Balk, B.; Maicher, A.; Dees, M.; Klermund, J.; Luke-Glaser, S.; Bender, K.; Luke, B. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1199–1205. [[CrossRef](#)]
48. Arora, R.; Azzalin, C.M. Telomere elongation chooses TERRA ALternatives. *RNA Biol.* **2015**, *12*, 938–941. [[CrossRef](#)]
49. Pfeiffer, V.; Crittin, J.; Grolimund, L.; Lingner, J. The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. *EMBO J.* **2013**, *32*, 2861–2871. [[CrossRef](#)]
50. Graf, M.; Bonetti, D.; Lockhart, A.; Serhal, K.; Kellner, V.; Maicher, A.; Jolivet, P.; Teixeira, M.T.; Luke, B. Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* **2017**, *170*, 72–85.e14. [[CrossRef](#)]
51. Stolz, R.; Sulthana, S.; Hartono, S.R.; Malig, M.; Benham, C.J.; Chedin, F. Interplay between DNA sequence and negative superhelicity drives R-loop structures. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6260–6269. [[CrossRef](#)]
52. Huang, F.T.; Yu, K.; Balter, B.B.; Selsing, E.; Oruc, Z.; Khamlichi, A.A.; Hsieh, C.L.; Lieber, M.R. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region. *Mol. Cell. Biol.* **2007**, *27*, 5921–5932. [[CrossRef](#)]
53. Domínguez-Sánchez, M.S.; Barroso, S.; Gómez-González, B.; Luna, R.; Aguilera, A. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* **2011**, *7*, e1002386. [[CrossRef](#)] [[PubMed](#)]
54. Kotsantis, P.; Silva, L.M.; Irmscher, S.; Jones, R.M.; Folkes, L.; Gromak, N.; Petermann, E. Increased global transcription activity as a mechanism of replication stress in cancer. *Nat. Commun.* **2016**, *7*, 13087. [[CrossRef](#)]
55. Tuduri, S.; Crabbé, L.; Conti, C.; Tourrière, H.; Holtgreve-Grez, H.; Jauch, A.; Pantescio, V.; de Vos, J.; Thomas, A.; Theillet, C.; et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.* **2009**, *11*, 1315–1324. [[CrossRef](#)]

56. El Hage, A.; French, S.L.; Beyer, A.L.; Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* **2010**, *24*, 1546–1558. [[CrossRef](#)]
57. Blackford, A.N.; Jackson, S.P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* **2017**, *66*, 801–817. [[CrossRef](#)] [[PubMed](#)]
58. Marabitti, V.; Lillo, G.; Malacaria, E.; Palermo, V.; Sanchez, M.; Pichierri, P.; Franchitto, A. ATM pathway activation limits R-loop-associated genomic instability in Werner syndrome cells. *Nucleic Acids Res.* **2019**, *47*, 3485–3502. [[CrossRef](#)]
59. Di Domenico, E.G.; Romano, E.; Del Porto, P.; Ascenzioni, F. Multifunctional role of ATM/Tel1 kinase in genome stability: From the DNA damage response to telomere maintenance. *Biomed. Res. Int.* **2014**, *2014*, 787404. [[CrossRef](#)]
60. Friedel, A.M.; Pike, B.L.; Gasser, S.M. ATR/Mec1: Coordinating fork stability and repair. *Curr. Opin. Cell Biol.* **2009**, *21*, 237–244. [[CrossRef](#)]
61. Danckwardt, S.; Hentze, M.W.; Kulozik, A.E. 3' end mRNA processing: Molecular mechanisms and implications for health and disease. *EMBO J.* **2008**, *27*, 482–498. [[CrossRef](#)] [[PubMed](#)]
62. Kaneko, S.; Chu, C.; Shatkin, A.J.; Manley, J.L. Human capping enzyme promotes formation of transcriptional R loops in vitro. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17620–17625. [[CrossRef](#)]
63. Mason, P.B.; Struhl, K. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell Biol.* **2003**, *23*, 8323–8333. [[CrossRef](#)]
64. Marabitti, V.; Valenzisi, P.; Lillo, G.; Malacaria, E.; Palermo, V.; Pichierri, P.; Franchitto, A. R-Loop-Associated Genomic Instability and Implication of WRN and WRNIP1. *Int. J. Mol. Sci.* **2022**, *23*, 1547. [[CrossRef](#)]
65. Belotserkovskii, B.P.; Tornaletti, S.; D'Souza, A.D.; Hanawalt, P.C. R-loop generation during transcription: Formation, processing and cellular outcomes. *DNA Repair* **2018**, *71*, 69–81. [[CrossRef](#)] [[PubMed](#)]
66. Sollier, J.; Stork, C.T.; García-Rubio, M.L.; Paulsen, R.D.; Aguilera, A.; Cimprich, K.A. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell* **2014**, *56*, 777–785. [[CrossRef](#)]
67. Wilson-Sali, T.; Hsieh, T.S. Preferential cleavage of plasmid-based R-loops and D-loops by Drosophila topoisomerase IIIbeta. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7974–7979. [[CrossRef](#)]
68. Bharati, A.P.; Singh, N.; Kumar, V.; Kashif, M.; Singh, A.K.; Singh, P.; Singh, S.K.; Siddiqi, M.I.; Tripathi, T.; Akhtar, M.S. The mRNA capping enzyme of *Saccharomyces cerevisiae* has dual specificity to interact with CTD of RNA Polymerase II. *Sci. Rep.* **2016**, *6*, 31294. [[CrossRef](#)]
69. Teloni, F.; Michelena, J.; Lezaja, A.; Kilic, S.; Ambrosi, C.; Menon, S.; Dobrovolna, J.; Imhof, R.; Janscak, P.; Baubec, T.; et al. Efficient Pre-mRNA Cleavage Prevents Replication-Stress-Associated Genome Instability. *Mol. Cell* **2019**, *73*, 670–683.e12. [[CrossRef](#)] [[PubMed](#)]
70. Mischo, H.E.; Gómez-González, B.; Grzechnik, P.; Rondón, A.G.; Wei, W.; Steinmetz, L.; Aguilera, A.; Proudfoot, N.J. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell.* **2011**, *41*, 21–32. [[CrossRef](#)]
71. Morales, J.C.; Richard, P.; Patidar, P.L.; Motea, E.A.; Dang, T.T.; Manley, J.L.; Boothman, D.A. XRN2 Links Transcription Termination to DNA Damage and Replication Stress. *PLoS Genet.* **2016**, *12*, e1006107. [[CrossRef](#)]
72. Jimeno, S.; Aguilera, A. The THO complex as a key mRNP biogenesis factor in development and cell differentiation. *J. Biol.* **2010**, *9*, 6. [[CrossRef](#)] [[PubMed](#)]
73. Gómez-González, B.; García-Rubio, M.; Bermejo, R.; Gaillard, H.; Shirahige, K.; Marín, A.; Foiani, M.; Aguilera, A. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J.* **2011**, *30*, 3106–3119. [[CrossRef](#)] [[PubMed](#)]
74. Yang, Z.; Hou, Q.; Cheng, L.; Xu, W.; Hong, Y.; Li, S.; Sun, Q. RNase H1 Cooperates with DNA Gyases to Restrict R-Loops and Maintain Genome Integrity in Arabidopsis Chloroplasts. *Plant Cell* **2017**, *29*, 2478–2497. [[CrossRef](#)] [[PubMed](#)]
75. Alzu, A.; Bermejo, R.; Begnis, M.; Lucca, C.; Piccini, D.; Carotenuto, W.; Saponaro, M.; Brambati, A.; Cocito, A.; Foiani, M.; et al. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* **2012**, *151*, 835–846. [[CrossRef](#)] [[PubMed](#)]
76. Sakasai, R.; Isono, M.; Wakasugi, M.; Hashimoto, M.; Sunatani, Y.; Matsui, T.; Shibata, A.; Matsunaga, T.; Iwabuchi, K. Aquarius is required for proper CtIP expression and homologous recombination repair. *Sci. Rep.* **2017**, *7*, 13808. [[CrossRef](#)] [[PubMed](#)]
77. Chakraborty, P.; Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair* **2011**, *10*, 654–665. [[CrossRef](#)]
78. Ribeiro de Almeida, C.; Dhir, S.; Dhir, A.; Moghaddam, A.E.; Sattentau, Q.; Meinhart, A.; Proudfoot, N.J. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol. Cell* **2018**, *70*, 650–662.e8. [[CrossRef](#)] [[PubMed](#)]
79. Wang, J.C. Cellular roles of DNA topoisomerases: A molecular perspective. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 430–440. [[CrossRef](#)]
80. Zheng, X.; Peng, Q.; Wang, L.; Zhang, X.; Huang, L.; Wang, J.; Qin, Z. Serine/arginine-rich splicing factors: The bridge linking alternative splicing and cancer. *Int. J. Biol. Sci.* **2020**, *16*, 2442–2453. [[CrossRef](#)] [[PubMed](#)]
81. Luna, R.; Gaillard, H.; González-Aguilera, C.; Aguilera, A. Biogenesis of mRNPs: Integrating different processes in the eukaryotic nucleus. *Chromosoma* **2008**, *117*, 319–331. [[CrossRef](#)]
82. Cerritelli, S.M.; Crouch, R.J. Ribonuclease H: The enzymes in eukaryotes. *FEBS J.* **2009**, *276*, 1494–1505. [[CrossRef](#)] [[PubMed](#)]
83. Newman, J.A.; Gileadi, O. RecQ helicases in DNA repair and cancer targets. *Essays Biochem.* **2020**, *64*, 819–830. [[CrossRef](#)]

84. Yang, Y.; McBride, K.M.; Hensley, S.; Lu, Y.; Chedin, F.; Bedford, M.T. Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol. Cell* **2014**, *53*, 484–497. [[CrossRef](#)] [[PubMed](#)]
85. Ogorodnikov, A.; Levin, M.; Tattikota, S.; Tokalov, S.; Hoque, M.; Scherzinger, D.; Marini, F.; Poetsch, A.; Binder, H.; Macher-Göppinger, S.; et al. Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat. Commun.* **2018**, *9*, 5331. [[CrossRef](#)] [[PubMed](#)]
86. Marini, F.; Scherzinger, D.; Danckwardt, S. TREND-DB-a transcriptome-wide atlas of the dynamic landscape of alternative polyadenylation. *Nucleic Acids Res.* **2021**, *49*, D243–D253. [[CrossRef](#)]
87. Nourse, J.; Spada, S.; Danckwardt, S. Emerging Roles of RNA 3'-end Cleavage and Polyadenylation in Pathogenesis, Diagnosis and Therapy of Human Disorders. *Biomolecules* **2020**, *10*, 915. [[CrossRef](#)] [[PubMed](#)]
88. Kilchert, C.; Wittmann, S.; Vasiljeva, L. The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 227–239. [[CrossRef](#)] [[PubMed](#)]
89. Ogorodnikov, A.; Danckwardt, S. TRENDseq-A highly multiplexed high throughput RNA 3' end sequencing for mapping alternative polyadenylation. *Methods Enzymol.* **2021**, *655*, 37–72. [[CrossRef](#)] [[PubMed](#)]
90. Malik, I.; Kelley, C.P.; Wang, E.T.; Todd, P.K. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 589–607. [[CrossRef](#)] [[PubMed](#)]
91. Paulson, H. Repeat expansion diseases. *Handb. Clin. Neurol.* **2018**, *147*, 105–123. [[CrossRef](#)] [[PubMed](#)]
92. Ellerby, L.M. Repeat Expansion Disorders: Mechanisms and Therapeutics. *Neurotherapeutics* **2019**, *16*, 924–927. [[CrossRef](#)] [[PubMed](#)]
93. Haeusler, A.R.; Donnelly, C.J.; Periz, G.; Simko, E.A.; Shaw, P.G.; Kim, M.S.; Maragakis, N.J.; Troncoso, J.C.; Pandey, A.; Sattler, R.; et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507*, 195–200. [[CrossRef](#)] [[PubMed](#)]
94. Reddy, K.; Schmidt, M.H.; Geist, J.M.; Thakkar, N.P.; Panigrahi, G.B.; Wang, Y.H.; Pearson, C.E. Processing of double-R-loops in (CAG)-(CTG) and C9orf72 (GGGGCC)-(GGCCCC) repeats causes instability. *Nucleic Acids Res.* **2014**, *42*, 10473–10487. [[CrossRef](#)] [[PubMed](#)]
95. McIntosh, C.S.; Li, D.; Wilton, S.D.; Aung-Htut, M.T. Polyglutamine Ataxias: Our Current Molecular Understanding and What the Future Holds for Antisense Therapies. *Biomedicines* **2021**, *9*, 1499. [[CrossRef](#)]
96. Lin, Y.; Dent, S.Y.; Wilson, J.H.; Wells, R.D.; Napierala, M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 692–697. [[CrossRef](#)]
97. Khristich, A.N.; Mirkin, S.M. On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **2020**, *295*, 4134–4170. [[CrossRef](#)] [[PubMed](#)]
98. McMurray, C.T. Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **2010**, *11*, 786–799. [[CrossRef](#)] [[PubMed](#)]
99. Paulson, H.L.; Fischbeck, K.H. Trinucleotide repeats in neurogenetic disorders. *Annu. Rev. Neurosci.* **1996**, *19*, 79–107. [[CrossRef](#)] [[PubMed](#)]
100. Usdin, K.; House, N.C.; Freudenreich, C.H. Repeat instability during DNA repair: Insights from model systems. *Crit. Rev. Biochem. Mol. Biol.* **2015**, *50*, 142–167. [[CrossRef](#)]
101. Groh, M.; Lufino, M.M.; Wade-Martins, R.; Gromak, N. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.* **2014**, *10*, e1004318. [[CrossRef](#)] [[PubMed](#)]
102. Loomis, E.W.; Sanz, L.A.; Chédin, F.; Hagerman, P.J. Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.* **2014**, *10*, e1004294. [[CrossRef](#)] [[PubMed](#)]
103. James, P.A.; Talbot, K. The molecular genetics of non-ALS motor neuron diseases. *Biochim. Biophys. Acta* **2006**, *1762*, 986–1000. [[CrossRef](#)]
104. Palau, F.; Espinós, C. Autosomal recessive cerebellar ataxias. *Orphanet. J. Rare Dis.* **2006**, *1*, 47. [[CrossRef](#)]
105. Sze, C.I.; Su, M.; Pugazhenth, S.; Jambal, P.; Hsu, L.J.; Heath, J.; Schultz, L.; Chang, N.S. Down-regulation of WW domain-containing oxidoreductase induces Tau phosphorylation in vitro. A potential role in Alzheimer's disease. *J. Biol. Chem.* **2004**, *279*, 30498–30506. [[CrossRef](#)] [[PubMed](#)]
106. Lim, Y.W.; Sanz, L.A.; Xu, X.; Hartono, S.R.; Chédin, F. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutières syndrome. *Elife* **2015**, *4*, e08007. [[CrossRef](#)] [[PubMed](#)]
107. Jackson, B.R.; Noerenberg, M.; Whitehouse, A. A novel mechanism inducing genome instability in Kaposi's sarcoma-associated herpesvirus infected cells. *PLoS Pathog.* **2014**, *10*, e1004098. [[CrossRef](#)]
108. Arora, R.; Lee, Y.; Wischnewski, H.; Brun, C.M.; Schwarz, T.; Azzalin, C.M. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* **2014**, *5*, 5220. [[CrossRef](#)] [[PubMed](#)]
109. Miller, M.S.; Rialdi, A.; Ho, J.S.; Tilove, M.; Martinez-Gil, L.; Moshkina, N.P.; Peralta, Z.; Noel, J.; Melegari, C.; Maestre, A.M.; et al. Senataxin suppresses the antiviral transcriptional response and controls viral biogenesis. *Nat. Immunol.* **2015**, *16*, 485–494. [[CrossRef](#)] [[PubMed](#)]
110. Stork, C.T.; Bocek, M.; Crossley, M.P.; Sollier, J.; Sanz, L.A.; Chédin, F.; Swigut, T.; Cimprich, K.A. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *Elife* **2016**, *5*, e17548. [[CrossRef](#)]
111. Hatchi, E.; Skourti-Stathaki, K.; Ventz, S.; Pinello, L.; Yen, A.; Kamieniarz-Gdula, K.; Dimitrov, S.; Pathania, S.; McKinney, K.M.; Eaton, M.L.; et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **2015**, *57*, 636–647. [[CrossRef](#)] [[PubMed](#)]

112. Hill, S.J.; Rolland, T.; Adelmant, G.; Xia, X.; Owen, M.S.; Dricot, A.; Zack, T.I.; Sahni, N.; Jacob, Y.; Hao, T.; et al. Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev.* **2014**, *28*, 1957–1975. [[CrossRef](#)]
113. Wang, Y.; Ma, B.; Liu, X.; Gao, G.; Che, Z.; Fan, M.; Meng, S.; Zhao, X.; Sugimura, R.; Cao, H.; et al. ZFP281-BRCA2 prevents R-loop accumulation during DNA replication. *Nat Commun.* **2022**, *13*, 3493. [[CrossRef](#)] [[PubMed](#)]
114. Bhatia, V.; Barroso, S.I.; García-Rubio, M.L.; Tumini, E.; Herrera-Moyano, E.; Aguilera, A. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **2014**, *511*, 362–365. [[CrossRef](#)]
115. Chen, W.D.; Han, Z.J.; Skoletsky, J.; Olson, J.; Sah, J.; Myeroff, L.; Platzer, P.; Lu, S.; Dawson, D.; Willis, J.; et al. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J. Natl. Cancer Inst.* **2005**, *97*, 1124–1132. [[CrossRef](#)]
116. Boque-Sastre, R.; Soler, M.; Oliveira-Mateos, C.; Portela, A.; Moutinho, C.; Sayols, S.; Villanueva, A.; Esteller, M.; Guil, S. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5785–5790. [[CrossRef](#)]
117. Gorthi, A.; Romero, J.C.; Loranc, E.; Cao, L.; Lawrence, L.A.; Goodale, E.; Iniguez, A.B.; Bernard, X.; Masamsetti, V.P.; Roston, S.; et al. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature* **2018**, *555*, 387–391. [[CrossRef](#)]
118. Becherel, O.J.; Yeo, A.J.; Stellati, A.; Heng, E.Y.; Luff, J.; Suraweera, A.M.; Woods, R.; Fleming, J.; Carrie, D.; McKinney, K.; et al. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet.* **2013**, *9*, e1003435. [[CrossRef](#)]
119. Bruno, T.; Corleone, G.; Catena, V.; Cortile, C.; de Nicola, F.; Fabretti, F.; Gumenyuk, S.; Pisani, F.; Mengarelli, A.; Passananti, C.; et al. AATF/Che-1 localizes to paraspeckles and suppresses R-loops accumulation and interferon activation in Multiple Myeloma. *EMBO J.* **2022**, *41*, e109711. [[CrossRef](#)]
120. Chen, L.; Chen, J.Y.; Huang, Y.J.; Gu, Y.; Qiu, J.; Qian, H.; Shao, C.; Zhang, X.; Hu, J.; Li, H.; et al. The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol. Cell* **2018**, *69*, 412–425.e6. [[CrossRef](#)] [[PubMed](#)]
121. Sagie, S.; Toubiana, S.; Hartono, S.R.; Katzir, H.; Tzur-Gilat, A.; Havazelet, S.; Francastel, C.; Velasco, G.; Chédin, F.; Selig, S. Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat. Commun.* **2017**, *8*, 14015. [[CrossRef](#)] [[PubMed](#)]
122. Thrasher, A.J. New insights into the biology of Wiskott-Aldrich syndrome (WAS). *Hematol. Am. Soc. Hematol. Educ. Program.* **2009**, *2009*, 132–138. [[CrossRef](#)] [[PubMed](#)]
123. Chandra, S.; Bronicki, L.; Nagaraj, C.B.; Zhang, K. WAS-Related Disorders. In *Gene Reviews*; Adam, M.P., Ardinger, H.H., Pagon, R.A., Eds.; University of Washington: Seattle, WA, USA, 1993.
124. Muralimanoharan, S.; Shamy, R.; Stansbury, N.; Schenken, R.; de la Pena Avalos, B.; Javanmardi, S.; Dray, E.; Sung, P.; Boyer, T.G. Aberrant R-loop-induced replication stress in MED12-mutant uterine fibroids. *Sci. Rep.* **2022**, *12*, 6169. [[CrossRef](#)] [[PubMed](#)]
125. Struve, N.; Hoffer, K.; Weik, A.S.; Riepen, B.; Krug, L.; Cetin, M.H.; Burmester, J.; Ott, L.; Liebing, J.; Gatzemeier, F.; et al. Increased replication stress and R-loop accumulation in EGFRvIII-expressing glioblastoma present new therapeutic opportunities. *Neurooncol. Adv.* **2021**, *4*, vdab180. [[CrossRef](#)] [[PubMed](#)]
126. Huang, W.; Qian, T.; Huang, Z.; Liu, Y.; Cui, L.; Zhu, P.; Zhong, Q.; Zeng, T.; Fu, L.; Si, C.; et al. Increased expression of IFI16 predicts adverse prognosis in multiple myeloma. *Pharm. J.* **2021**, *21*, 520–532. [[CrossRef](#)]
127. Giannini, M.; Bayona-Feliu, A.; Sproviero, D.; Barroso, S.I.; Cereda, C.; Aguilera, A. TDP-43 mutations link Amyotrophic Lateral Sclerosis with R-loop homeostasis and R loop-mediated DNA damage. *PLoS Genet.* **2020**, *16*, e1009260. [[CrossRef](#)] [[PubMed](#)]
128. Boguslawski, S.J.; Smith, D.E.; Michalak, M.A.; Mickelson, K.E.; Yehle, C.O.; Patterson, W.L.; Carrico, R.J. Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* **1986**, *89*, 123–130. [[CrossRef](#)]
129. Phillips, D.D.; Garboczi, D.N.; Singh, K.; Hu, Z.; Leppla, S.H.; Leysath, C.E. The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. *J. Mol. Recognit.* **2013**, *26*, 376–381. [[CrossRef](#)] [[PubMed](#)]
130. Hu, Z.; Zhang, A.; Storz, G.; Gottesman, S.; Leppla, S.H. An antibody-based microarray assay for small RNA detection. *Nucleic Acids Res.* **2006**, *34*, e52. [[CrossRef](#)] [[PubMed](#)]
131. Yeo, A.J.; Becherel, O.J.; Luff, J.E.; Cullen, J.K.; Wongsurawat, T.; Jenjaroenpun, P.; Kuznetsov, V.A.; McKinnon, P.J.; Lavin, M.F. R-loops in proliferating cells but not in the brain: Implications for AOA2 and other autosomal recessive ataxias. *PLoS ONE* **2014**, *9*, e90219. [[CrossRef](#)] [[PubMed](#)]
132. Wahba, L.; Costantino, L.; Tan, F.J.; Zimmer, A.; Koshland, D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* **2016**, *30*, 1327–1338. [[CrossRef](#)] [[PubMed](#)]
133. Sanz, L.A.; Chédin, F. High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* **2019**, *14*, 1734–1755. [[CrossRef](#)] [[PubMed](#)]
134. Nadel, J.; Athanasiadou, R.; Lemetre, C.; Wijetunga, N.A.; Broin, P.O.; Sato, H.; Zhang, Z.; Jeddeloh, J.; Montagna, C.; Golden, A.; et al. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics Chromatin* **2015**, *8*, 46. [[CrossRef](#)] [[PubMed](#)]
135. Dumelie, J.G.; Jaffrey, S.R. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. *Elife* **2017**, *6*, e28306. [[CrossRef](#)]

136. Crossley, M.P.; Bocek, M.J.; Hamperl, S.; Swigut, T.; Cimprich, K.A. qDRIP: A method to quantitatively assess RNA-DNA hybrid formation genome-wide. *Nucleic Acids Res.* **2020**, *48*, e84. [[CrossRef](#)]
137. Malig, M.; Hartono, S.R.; Giafaglione, J.M.; Sanz, L.A.; Chedin, F. Ultra-deep Coverage Single-molecule R-loop Footprinting Reveals Principles of R-loop Formation. *J. Mol. Biol.* **2020**, *432*, 2271–2288. [[CrossRef](#)]
138. El Hage, A.; Webb, S.; Kerr, A.; Tollervey, D. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* **2014**, *10*, e1004716. [[CrossRef](#)]
139. Yan, Q.; Shields, E.J.; Bonasio, R.; Sarma, K. Mapping Native R-Loops Genome-wide Using a Targeted Nuclease Approach. *Cell Rep.* **2019**, *29*, 1369–1380.e5. [[CrossRef](#)]
140. Wang, K.; Wang, H.; Li, C.; Yin, Z.; Xiao, R.; Li, Q.; Xiang, Y.; Wang, W.; Huang, J.; Chen, L.; et al. Genomic profiling of native R loops with a DNA-RNA hybrid recognition sensor. *Sci. Adv.* **2021**, *7*, eabe3516. [[CrossRef](#)]
141. Yan, Q.; Wulfridge, P.; Doherty, J.; Fernandez-Luna, J.L.; Real, P.J.; Tang, H.Y.; Sarma, K. Proximity labeling identifies a repertoire of site-specific R-loop modulators. *Nat. Commun.* **2022**, *13*, 53. [[CrossRef](#)]
142. Jenjaroenpun, P.; Wongsurawat, T.; Yenamandra, S.P.; Kuznetsov, V.A. QmRLFS-finder: A model, web server and stand-alone tool for prediction and analysis of R-loop forming sequences. *Nucleic Acids Res.* **2015**, *43*, W527–W534. [[CrossRef](#)] [[PubMed](#)]
143. Brázda, V.; Havlík, J.; Kolomazník, J.; Trenz, O.; Št'astný, J. R-Loop Tracker: Web Access-Based Tool for R-Loop Detection and Analysis in Genomic DNA Sequences. *Int. J. Mol. Sci.* **2021**, *22*, 2857. [[CrossRef](#)] [[PubMed](#)]
144. Xu, W.; Li, K.; Li, S.; Hou, Q.; Zhang, Y.; Liu, K.; Sun, Q. The R-Loop Atlas of Arabidopsis Development and Responses to Environmental Stimuli. *Plant Cell* **2020**, *32*, 888–903. [[CrossRef](#)] [[PubMed](#)]
145. Jenjaroenpun, P.; Wongsurawat, T.; Sutheeworapong, S.; Kuznetsov, V.A. R-loopDB: A database for R-loop forming sequences (RLFS) and R-loops. *Nucleic Acids Res.* **2017**, *45*, D119–D127. [[CrossRef](#)]
146. Chen, L.; Chen, J.Y.; Zhang, X.; Gu, Y.; Xiao, R.; Shao, C.; Tang, P.; Qian, H.; Luo, D.; Li, H.; et al. R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol. Cell* **2017**, *68*, 745–757.e5. [[CrossRef](#)]
147. Shaw, N.N.; Xi, H.; Arya, D.P. Molecular recognition of a DNA:RNA hybrid: Sub-nanomolar binding by a neomycin-methidium conjugate. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4142–4145. [[CrossRef](#)]
148. Lin, R.; Zhong, X.; Zhou, Y.; Geng, H.; Hu, Q.; Huang, Z.; Hu, J.; Fu, X.D.; Chen, L.; Chen, J.Y. R-loopBase: A knowledgebase for genome-wide R-loop formation and regulation. *Nucleic Acids Res.* **2022**, *50*, D303–D315. [[CrossRef](#)]