Published in final edited form as:

J Mol Biol. 2017 October 27; 429(21): 3168–3180. doi:10.1016/j.jmb.2016.08.031.

R Loops and Links to Human Disease

Patricia Richard and James L. Manley

Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Abstract

Aberrant R-loop structures are increasingly being realized as an important contributor to human disease. R loops, which are mainly co-transcriptional, abundant RNA/DNA hybrids, form naturally and can indeed be beneficial for transcription regulation at certain loci. However, their unwanted persistence elsewhere or in particular situations can lead to DNA double-strand breaks, chromosome rearrangements, and hypermutation, which are all sources of genomic instability. Mutations in genes involved in R-loop resolution or mutations leading to R-loop formation at specific genes affect the normal physiology of the cell. We discuss here the examples of diseases for which a link with R loops has been described, as well as how disease-causing mutations might participate in the development and/or progression of diseases that include repeat-associated conditions, other neurological disorders, and cancers.

Keywords

R loops; genomic instability; chromatin reorganization; transcription; human disease

Introduction

R loops are nucleic acid structures in which a nascent G-rich transcript hybridizes with the DNA template strand, leaving the non-template DNA single stranded [1]. R loops were first described in prokaryotes more than 20 years ago [2]. For a long time, they were considered largely as a by-product of transcription that did not have significant consequences and did not generate much interest. However, during the past decade or so, an increasing number of studies have revealed very important functions of R loops in transcription that when disturbed can be linked to a variety of diseases. Genome-wide mapping studies of R loops in human, mouse, and yeast cells by DNA/RNA immunoprecipitation followed by high-throughput sequencing (DRIP-seq) have tremendously helped to understand how and where R loops form [3–7]. Recent high-resolution DRIP-seq experiments suggest that R loops are abundant structures and are estimated to occupy up to 5% of mammalian genomes and 8% of the budding yeast genome [6,7].

R loops occur naturally during transcription and can have important functions. For example, they are important for class switch recombination of immunoglobulin (Ig) genes in activated B cells [8] and for mitochondrial replication [9,10]. R loops are also found frequently at GC-

rich regions such as typify many promoters and 3' end regions, where they appear to play significant roles in transcription [3,4,7,11,12]. However, in some cases, persistence of R loops can have deleterious effects. For example, they can result in the accumulation of DNA double-strand breaks (DSBs) [13], leading to DNA rearrangements and genome instability [1,14]. It is thus not surprising that R loops have been linked to disease, including some cancers and several neurodegenerative disorders.

Several reviews describing how R loops form and how they can be resolved or lead to genome instability have appeared in the last few years [1,15–18]. The role of R loops in disease has also been recently reviewed [16,17,19]. However, examples of R loops in pathological conditions have continued to bloom, and here, we highlight recent advances in understanding how R loops are connected to disease.

R Loops and Genome Stability

Before delving into the interconnections between R loops and disease, we briefly highlight the many connections between these structures and genome instability, as this is central to their role in many pathological conditions. The first strong links to genome instability were discovered over 10 years ago, in yeast and chicken cells. In yeast, deletion of *hpr1*, which encodes a component of the THO complex that couples transcription with mRNA export, was found to induce the inhibition of transcription elongation, DNA damage, and hyper-recombination through the formation of R loops [20]. The same connection was shown later with the mammalian THO complex [21]. In chicken DT40 cells, depletion of the splicing factor SRSF1 was shown to cause DSBs and DNA hypermutation, again due to the formation of R loops during transcription [22,23]. In fact, it is now known that many factors involved in RNA metabolism and processing play roles in preventing R-loop formation by packaging nascent RNAs into ribonucleoproteins (RNPs) that prevent its hybridization to the DNA template [20,23–26]. In the presence of R loops, the displaced ssDNA becomes sensitive to nucleases such as activation-induced cytidine deaminase, which leads to DSBs necessary to create hypermutations and diversity at the *Ig* locus in B cells [27]. XPF and XPG, two endonucleases of the nucleotide excision repair machinery, have also been found to process R loops into DSBs [13]. Such DSB-inducing mechanisms could participate in R-loop-dependent genome instability. Loss of topoisomerase I, which plays roles in replication and transcription, favors R-loop formation through increased levels of negative supercoiling. This can drive the stalling of both RNA polymerase (RNAP) and replication forks, leading to chromosome breaks [2,28,29]. Stalled replication forks are unstable structures prone to recombination and genome instability [30]. R loops have been found to cause replication impairments in various conditions and organisms, from bacteria to humans [31–34]. Collision between the transcription and replication machineries at common fragile sites in higher eukaryotes, the hot spot of chromosomal rearrangements, has been found responsible for R-loop formation and genomic instability [35]. However, it is unclear what comes first. Does the collision between RNAP and the replisome cause RNAP stalling and lead to R-loop formation? Or does R-loop formation cause RNAP stalling and then collision that arrests the replication fork? While evidence point to both possibilities, genomic instability appears to be the common outcome in the development of several diseases.

R-loop formation at trinucleotide repeat-associated diseases

More than 40 genetic disorders are caused by gene-specific repeat expansions. These include Huntington's disease [CAG repeats in *huntingtin* (*HTT*)], myotonic dystrophy type 1 [CTG repeats in *dystrophia myotonica protein kinase* (*DMPK*)], spinocerebellar ataxia type 1 [CAG repeats in *ataxin 1* (*ATXN1*)], fragile X mental retardation or fragile X syndrome (FXS) [CGG repeats in *fragile X mental retardation 1* (*FMR1*)], and Friedreich ataxia [GAA repeats in *frataxin* (*FXN*)] [36,37]. Those extended repeats can lead to various outcomes, such as transcription inhibition and expression of toxic RNAs or toxic polyglutamine proteins. A significant number of these have been linked to R loops (Table 1). For example, such repeats often have high GC content [36] and can potentially form R loops [38]. Indeed, *in vitro* transcription of CAG, CTG, GAA (not so high GC content), CGG, and CCG repeats can form R loops in *Escherichia coli* [39]. R loops formed *in vitro* and in bacteria at GAA repeats have been linked to RNA polymerase arrest [40]. Using a similar system, CTG repeats lead to the formation of R loops, promoting repeat instability in bacteria and human cells carrying transgenic repeats [41]. Repeat instability leading to expansion or deletion of the repeats has been shown in other disease models in bacteria [42], flies [43], and human cells [44]. Very often, the repeats are transcribed in both directions, which is believed to stimulate repeat instability [39,44–47]. This increased instability has been proposed to result from the formation of R loops on both strands, referred to as double R loops [39,48]. Interestingly, R-loop composition affects instability by leading to repeat expansion or deletion [48].

R loops have been well documented in the *FXN* and *FMR1* genes. Using transformed lymphoblastoid cell lines derived from FRDA and FXS patient cells, Groh *et al.* confirmed that highly stable R loops form at the expanded repeats of both genes [49]. FRDA is the most common inherited ataxia, caused by the GAA expansion in the first intron of *FXN*, leading to transcriptional silencing of *FXN* [50]. Importantly, R-loop enrichment has been found to be associated with the repressive H3K9me2 chromatin mark at *FXN* in Friedreich ataxia cells [49]. Interestingly, H3K9me2 is also found to be associated with R loops forming at gene 3' ends, perhaps contributing to efficient transcription termination by promoting RNAPII pausing and release from the DNA template [11,12]. While a decrease in H3K9me2 does not affect R-loop or mRNA levels, increased R-loop formation, by the addition of the topoisomerase I inhibitor camptothecin, leads to increased H3K9me2 and transcription inhibition, suggesting that histone modification is a consequence of R-loop formation [49]. Expanded repeats at *FXN* appear to be another example of R-loop-dependent transcription elongation inhibition [20,51].

The scenario at *FMR1* appears quite different. The *FMR1* promoter and repeat region (in the 5' UTR) are hypermethylated in FXS patient cells containing more than 200 repeats (full mutation), resulting in the repression of *FMR1* transcription [52,53]. The absence of the FMR1 protein is the leading cause of the intellectual disability of the disease. In contrast, premutation alleles, which contain between 55 and 200 repeats, trigger higher transcriptional activity, as seen in fragile X-associated tremor/ataxia syndrome (FXTAS) for which a toxic gain-of-function of the protein is suspected to cause disease [53,54]. However, R loops form at *FMR1* in both FXS and FXTAS [49,55,56]. Indeed, after transcription reactivation by

methylation inhibition, R loops accumulate at the *FMR1* locus in FXS cells [49]. A study monitoring the differentiation of FXS embryonic stem cells into neurons [56,57] showed that the formation of an R loop leads to epigenetic silencing of the *FMR1* promoter [56]. Altogether, these data indicate that R-loop formation may play a role in switching off *FMR1* expression early during development in FXS patients through epigenetic silencing. Furthermore, R loops forming at premutation loci trigger the formation of a hairpin-like structure in the displaced ssDNA that might promote the transcriptional regulation seen in such alleles [55]. R loops are suspected to protect CpG island promoters from methylation and epigenetic silencing [3], and perhaps, a similar mechanism is used at premutation loci to stimulate transcription. However, it is difficult to understand how R-loop formation at the same gene can have two different outcomes. Is it possible that R-loop formation at shorter repeats triggers different epigenetic changes or different DNA structures, such as hairpins [55,58] or G-quadruplexes (G-Qs), which have the potential to stabilize R loops [59]? G-Qs are highly stable stacks of planar tetramers of four guanines connected by Hoogsteen hydrogen bonds [60]. Similar to R loops, they are enriched in 5'UTRs, first introns, and 3'UTRs [61,62]. Additionally, G-Q and hairpin structures have different protein binding specificity that might trigger distinct processes [63].

Trinucleotide repeat-associated diseases make a clear connection between R loops and epigenetic changes. R loops have been shown to associate with particular chromatin states, which can be open [7,64,65] or condensed [12,66]. It will be important to understand further how R-loop formation can trigger those differences not only in normal and disease cells but also in different cell types or physiological states.

R loops and cancer

Genome instability is a hallmark of almost all cancer cells [67–69]. Replication stress and genome instability at common fragile sites have been linked to cancer development [70]. As we discuss in the following paragraphs, inappropriate accumulation of R loops and the resultant genomic instability appear to play a role in a number of human cancers.

Mutations in the tumor suppressor genes *BRCA1* and *BRCA2* increase the risk of breast and ovarian cancers and some other cancers [71–73]. *BRCA1* and *BRCA2*, which function in DSB repair by homologous recombination, have both been shown to prevent R-loop formation. Indeed, *BRCA1* and *BRCA2* knockdown (KD) leads to the formation of γ H2AX foci, indicative of DNA damage, and this correlates with an increase in R-loop foci [74–76]. Later, Hatchi *et al.* showed that *BRCA1* accumulates at certain RNAPII termination regions that appear to depend on R-loop formation, where it prevents the accumulation of R loops and resultant DNA damage [77]. Interestingly, mutational profiling of 21 breast cancers showed that *BRCA1*-deficient tumors have significantly more insertions-deletions near R-loop-dependent terminators than *BRCA2*-dependent or *BRCA1/2*-independent breast cancers [77]. This data indicate that while *BRCA1* and *BRCA2* both play a role in R-loop prevention, they might use different mechanisms and perhaps have different locus specificities.

Mutations in the mRNA polyadenylation machinery can also result in R-loop formation and DNA damage. For example, mutations affecting the polyadenylation factor FIP1 result in

increasing DNA damage foci and R loops in yeast [25]. Interestingly, 10–20% of eosinophilic leukemias express a truncated version of FIP1, resulting from an oncogenic fusion between FIP1 and PDGFR α [78]. While truncated FIP1 induces R-loop-mediated DNA damage, FIP1 KD increases 53BP1 foci, an indicator of DNA damage [25]. Altogether, those data indicate that partial loss of function of FIP1 and a subsequent increase in R-loop formation could be responsible for genomic instability in eosinophilic leukemia.

The E3 ligase Bre1 (RNF20/40) monoubiquitinates histone H2B and plays a role in transcription elongation control and DNA damage repair. Bre1-depleted cells show increased γ H2AX levels that are RNase H1 sensitive, suggesting that the DSBs and genomic rearrangements associated with Bre1 KD are R-loop-dependent [79]. Importantly, Bre1, which is normally highly expressed in testis, is downregulated in a type of testicular germ cell cancer (called seminoma) known to accumulate high levels of genome instability. It is thus possible that R-loop formation plays a role in this cancer. Notably, the CpG island-rich *RNF20* promoter is often hypermethylated in breast and prostate cancers [80]. RNF20 expression is often found down regulated in metastatic prostate cancer, and DNA methylation inhibition leads to an increase in RNF20 mRNA in prostate cancer cells [80].

CpG islands are characteristic of promoter regions of many human genes, especially the so-called “housekeeping” genes. These regions are subject to DNA methylation, but when hypomethylated, they have a tendency to form R loops [3]. This can be relevant to cancer, as the following two examples illustrate.

The vimentin (*VIM*) gene encodes an intermediate filament protein, and its CpG island promoter region is hypermethylated, and expression is silenced in several cancers, including colon cancer [81]. Interestingly, in normal cells, an antisense RNA hybridizes with the *VIM* promoter region to form R loops, which help maintain an open chromatin state and increase *VIM* transcription [65]. It will be important in the future to explore the relationship between the observed DNA hypermethylation and R-loop formation in *VIM* transcriptional regulation and its relevance to cancer.

Tudor domain-containing protein 3 (TDRD3) is a transcriptional co-activator that facilitates gene expression by recognizing specific methyl-histone marks at promoters. TDRD3 forms a complex with the DNA topoisomerase TOP3B, which is recruited to the *c-MYC* CpG island promoter, where it prevents R-loop formation, likely by relaxing negative super-coiling, and facilitates *c-MYC* transcription [82]. *TDRD3*-null mice show increased DSBs and R-loop formation at *c-MYC* and at the Ig heavy chain (*Igh*) locus, which correlates with increased *c-MYC-Igh* translocations [82]. This suggests that R loops may have a role in the oncogenic *c-MYC-Igh* translocation commonly seen in Burkitt's lymphoma and multiple myeloma [83,84]. Consistent with this, *Igh* and *c-MYC* transgenes have been found to be DSB and recombination hot spots in an *hpr1* yeast strain that accumulates R loops [20,85].

Proteins from the Fanconi anemia (FA) pathway are important to ensure proper DNA replication. Loss of any FA genes leads to FA, a condition characterized by bone marrow failure and an increased risk of cancers [86]. Interestingly, R loops accumulate in KD and knockout (KO) cells for proteins from the FA pathway and in FA patients cells and result in

an increase in replication fork stalling and DNA breaks [87,88]. Significantly, BRCA1 and BRCA2 are also part of the FA pathway. Mutations affecting another FA pathway protein that interacts with BRCA1, the FANCD1 helicase, are linked to breast cancer and FA. Interestingly, stabilization of G-Qs that can normally be unwound by FANCD1 leads to DNA damage in FANCD1-depleted cells [89].

Kaposi's sarcoma-associated herpesvirus is an oncogenic virus associated with several AIDS-associated malignancies. The virus has been shown to cause DSBs and chromosome instability in infected cells [90,91]. We noted above that the THO/TREX complex prevents R-loop accumulation in human cells [21]. Intriguingly, Kaposi's sarcoma-associated herpesvirus-infected cells express a protein, ORF57, implicated in viral mRNA processing, which binds and inactivates the hTREX complex, making it no longer available for cellular mRNA processing and leading to R-loop formation and genome instability [92]. Expression of the THO component THOC1 is decreased in testis and skin tumors [93], although the significance of this remains to be determined.

All the above examples establish a connection between R-loop accumulation and genomic instability in cancer. The actual steps that lead to instability are still largely unknown.

R-loop connections to neurological disorders

ALS/FTD

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease, is a very complex disease for which many genes have been implicated in pathogenesis [94]. ALS is the most common motor neuron disease, characterized by a loss of motor neurons in the brain and spinal cord that leads to a rapid loss of control muscle movement and death typically 3 to 5 years following diagnosis [95]. About 90% of ALS cases are sporadic, and less than 10% are familial [96]. While the exact molecular mechanism(s) leading to disease are unknown, and indeed multiple mechanisms undoubtedly exist, evidence supporting a possible contribution from R loops in certain cases emerged [97]. We discuss below several genes with ALS mutations in which R loops have been implicated.

C9ORF72—The most common known cause of both familial and sporadic ALS is a hexanucleotide expansion in the *C9ORF72* (*C9*) gene. In fact, the expansion, GGGGCC (G4C2; located in the first intron of *C9*), is frequent not only in ALS but also in the related syndrome frontotemporal dementia (FTD) [98–100], which is characterized by the degeneration of the frontal and temporal lobes of the brain and is a common cause of early onset dementia [101]. While normal individuals have a number of G4C2 repeats below 30 [102], ALS patients can have up to several thousand repeats. As would be expected from the extreme G-richness of this sequence, R loops form at *in vitro*-transcribed repeat expansions, and this occurs concomitantly with RNAP II stalling and the accumulation of abortive RNAs containing G-Qs, coupled with a decrease in full-length mRNAs [48,63,103–105]. G-Qs also form on the displaced DNA strand and are believed to stabilize R loops [63].

How might R-loop formation affect *C9* transcription? *C9* patients exhibit repressive chromatin marks such as H3K9me3 or H3K26me3 on the *C9* gene [106], although as noted

above, the relationship between R loops and chromatin modification is poorly understood. Again, consistent with the sequence of the C9 hexanucleotide, DNA methylation has also been found to be associated with the repeat expansion [107–111]. Since bidirectional transcription also occurs at the expansion [112–116] and that R loops are enriched at genes producing antisense RNA [5,12], it is possible that R loops reduce C9 expression through the regulation of bidirectional transcription and chromatin modification (reviewed in Ref. [117]). While reduced C9 protein levels may contribute to disease [104,118,119], other mechanisms, such as the sequestration of one or more RNA binding protein (e.g., Ref. [63]) and/or the accumulation of toxic dipeptide repeat proteins generated by Repeat Associated Non-AUG (RAN) translation [112,120,121], may also be important. Thus, considerable work is required to understand how the C9 repeats cause ALS/FTD and how R-loop formation at the locus may be involved.

G4C2 repeats have not been found at other genes implicated in ALS or FTD [122]. While this could indicate that the formation of R loops and/or G-Qs is unique to the C9 expansion, many other G/C combination repeats can lead to these structures [39], and it is conceivable that they may be found associated with some of the ~50% of ALS cases for which no mutation is known. Notably, the C9 expansion has also been linked to other neurological disorders, such as Alzheimer's [123–126] and Huntington's [127] diseases.

Ataxin-2—PolyQ expansions in Ataxin-2 (ATXN2) have been linked to ALS [128,129] and spinocerebellar ataxia type 2 (SCA2) [130–132]. SCA2 results from the expansion of CAG repeats in *ATXN2*, which normally contains 22 repeats. It has recently been shown that intermediate-length repeats (27–33 repeats) might play a role in the development of ALS in C9 patients and others [128,129,133]. The ortholog of ATXN2 in yeast, Pbp1, maintains ribosomal DNA stability through the inhibition of the deleterious formation of R loops at G-Q DNA sites in the intergenic spacer regions of the ribosomal DNA repeats [134]. Pbp1 also suppresses R-loop formation at telomeres [134]. It is possible that ATXN2 inhibits the formation of R loops at some loci, although ATXN2 is predominantly a cytoplasmic protein [128]. It will be interesting to determine whether SCA2 cells accumulate R loops. In fact, ATXN2 interacts with TAR DNA-binding protein 43 (TDP-43), and the longer the ATXN2 polyQs, the stronger the interaction [129]. TDP-43 is associated with ALS and FTD [135], and while normally nuclear, mutated TDP-43 aggregates in the cytoplasm and associates with stress granules in ALS neurons, a hallmark of ALS [135,136]. Interestingly, ATXN2 and TDP-43 both relocate to cytoplasmic stress granules after stress induction [136,137]. It is then possible that intermediate-length polyQ expansions in ATXN2 help in sequestering TDP-43 into cytoplasmic aggregates and thus contribute to the disease. This scenario implies that TDP-43 could have a role in R-loop formation/resolution in the nucleus. Similarly, ATXN2 also co-localizes with Fused in Sarcoma (FUS), which is also mutated in ALS, in cytoplasmic inclusions [133]. Determining whether TDP-43 and FUS have a direct role in R-loop regulation might bring some light on the link between ATXN2 polyQ expansions and ALS.

SETX—Senataxin (SETX) is an RNA/DNA helicase that is mutated in two distinct neurological disorders, ataxia with oculomotor apraxia type 2 (AOA2) [138] and a juvenile

form of ALS known as ALS4 [139]. Interestingly, AOA2 mutations are recessive, while ALS4 mutations all appear to be dominant. Mutations in *SETX* have also been linked to atypical forms of AOA2 or ALS4 that have variant phenotypes [140–145]. Sen1, the *SETX* homolog in yeast, plays a role in transcription termination of non-coding RNAs and some mRNAs [146,147] and, importantly, was shown to suppress R-loop formation and transcription-associated recombination [148]. *SETX* was also shown to resolve R loops that form normally at the transcription termination pause site of the human β -actin gene [11]. Recently, *SETX* association with R-loop-dependent 3' pause sites was extended to several other genes and was shown to be BRCA1-dependent [77]. *SETX* and BRCA1 co-localize on meiotic chromosomes and have been found to interact genetically and physically [75,77,149]. *SETX*/BRCA1 recruitment at 3' pause site prevents the accumulation of deleterious R loops that would otherwise lead to DNA damage [77]. Interestingly, *SETX* maps to deleted regions in ovarian cancer, which might suggest a tumor suppressor role for *SETX* [75]. While this seems possible, AOA2 patients do not show a higher susceptibility to cancer than the rest of the population.

How do mutations in *SETX* lead to disease? Both sen1 and *SETX* prevent genome instability by suppressing DSBs arising from R loops, reflecting the collisions between the transcription and replication machineries [150,151]. In response to such replication stress, *SETX* relocates to R-loop-containing nuclear foci [151,152], where it likely resolves the RNA:DNA moiety of the R loop. Unexpectedly, the nuclear exosome, a 3' to 5' exoribonuclease complex that functions in RNA quality control, was found to relocate with *SETX*, mediated by direct interaction with an exosome subunit, Rrp45 [152]. Importantly, the *SETX*-Rrp45 interaction is dependent on *SETX* sumoylation, which is disrupted by AOA2, but not ALS4, mutations [152,153]. Thus, it seems that after *SETX* unwinds the R loops, the exosome is poised to degrade the released RNA, preventing rehybridization or other possible toxic effects. How and if this is critical to AOA2 specifically remain to be determined.

Perhaps not unexpectedly, *SETX* KO mice accumulate R loops. However, *SETX* KO does not lead to any neurological characteristics of AOA2, but it does cause male sterility [149,154]. It is still unclear at what level R-loop accumulation is related to disease. Interestingly, AOA2 cells have reduced telomere length, and since R loops and G-Qs form at telomeres, it is possible that *SETX* also plays a role in telomere stability [155–157]. Finally, most data show an accumulation of R loops after *SETX* KO or KD, which relates more to AOA2. It will then be of great interest in the future to explore the R-loop pattern in cells containing ALS4 mutations.

The 5'-3' exonuclease XRN2 also participates in transcription termination by degrading nascent RNA downstream from the 3' cleavage site [158,159]. Interestingly, XRN2 KD was recently found to lead to the accumulation of R loops and DSBs, which are in turn associated with increased replication stress and genomic instability [160]. Similar to *SETX*, XRN2 prevents DSB accumulation at 3' pause sites of a subset of genes. Thus, while R loops are necessary at certain transcriptional terminators, their persistence appears deleterious to the cell.

Other possible R-loop-related neurological disorders

Aicardi–Goutières syndrome (AGS) is a severe childhood inflammatory disorder causing neurological damage [161]. Genome-wide analysis of AGS patient fibroblasts (characterized by mutations in *TREX1*, which encodes a 3′-5′ ssDNA exonuclease or *RNASEH2*, which encodes a second RNase H enzyme) revealed a global increase in R loops [162].

Interestingly, however, the overall pattern at the expected GC-skewed regions was similar between control and AGS patient cells [4]. Thus, the excess of R loops specific to AGS cells was not at the 5′ or 3′ end of genes but instead at the intergenic or gene body regions [162]. This suggests that R loops formed in AGS patient cells are probably of a different type than the more naturally forming R loops. AGS patient cells also show a global DNA hypomethylation pattern, which corresponds to the AGS-specific R-loop loci, particularly in cells carrying an *RNASEH2* mutation [162]. These data strongly suggest that R-loop formation in AGS patient cells is linked to epigenetic changes that might trigger the immune response stimulation that is typical of the disease. Interestingly, mutations in *TREX1* and *RNASEH2* are also linked to systemic lupus erythematosus [163,164], which might indicate an R-loop contribution to another autoimmune disease.

Prader–Willi syndrome (PWS) leads to obesity in children, is associated with motor and mental retardation, [165] and involves genes subject to imprinting in the brain [166,167]. The locus includes the maternally expressed gene encoding the E3 ubiquitin-protein ligase *Ube3a* and a long paternally expressed transcript, which includes the antisense transcript to *Ube3a*, *Ube3a-ATS*, and the GC-skewed *SNORD116* locus that encodes a cluster of non-coding RNAs, including small nucleolar RNAs (snoRNAs). Paternally inherited deletions of *SNORD116* and the loss of snoRNAs are linked to PWS [168]. Normally, R loops form at the *SNORD116* intronic region in neurons and have been linked to chromatin decondensation that allows the expression of *Ube3a-ATS*, which in turn silences *Ube3a* on the paternal allele [64]. However, treatment with Topotecan, a topoisomerase inhibitor, leads to excess levels of R loops over the *SNORD116* locus that inhibit transcription through *Ube3a-ATS*, allowing the expression of *Ube3a* [64]. This example thus reveals a versatile and subtle form of transcriptional regulation mediated by R loops and is consistent with the possibility that the loss of R-loop formation at *SNORD116* may contribute to PWS.

Conclusions

R loops have gone from being a “ghost structure” that did not receive much attention to a master player of transcription regulation, chromatin modification, and genomic stability. While certain cancers and a considerable number of neurological disorders have been linked in one way or another to R loops, it is still very difficult to determine whether R loops are the cause of the disease or a secondary effect due to other issues. Nevertheless, it seems that resolving R loops or, in some cases, inducing them might have a therapeutic value. However, considerably more research to obtain a clear understanding of how and under what circumstances R loops form and to distinguish a “good” R loop from a “bad” R loop is needed. How can some R loops be deleterious to the cell while some are necessary and protective? It also seems that many loci prone to R-loop formation might also be prone to DNA and RNA G-Qs. Many proteins have been suggested to bind G-Qs, and it will be

interesting to determine whether any of these proteins also bind R loops. And then, what is the deleterious structure for the cell? The R loops or the G-Qs or both? Is it the actual change in chromatin accessibility, resulting from R loops and/or G-Qs, that is responsible for genomic instability? Could chromatin modification and compaction generated through R-loop formation be the obstacle leading to instability and to replication fork progression [34,169]?

Many intriguing properties of R loops remain to be investigated. While essentially all described R loops are co-transcriptional, and thus form in *cis*, R loops forming in *trans*, far from their site of transcription, have been reported in yeast [170]. The possibility that R loops can form in *trans* in the human genome would potentially extend both their regulatory functions and deleterious effects to a considerable yet unknown locus and might be linked to more diseases than the ones described here. The identification of R-loop-binding proteins will provide important knowledge on how they form, are maintained, and impact genome stability. Finally, many diseases are caused by gene-specific repeat expansions for which a possible role for R loops has not yet been investigated [36]. It is thus likely that more examples of R loops playing important and unexpected roles in disease remain to be uncovered.

Acknowledgments

P.R. is supported in part by a Young Investigator Research Grant awarded by the National Ataxia Foundation, and relevant work in our lab is supported by NIH grant R35 GM118136 to J.L.M.

References

1. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell*. 2012; 46:115–124. [PubMed: 22541554]
2. Drolet M, Bi X, Liu LF. Hypernegative supercoiling of the DNA template during transcription elongation *in vitro*. *J Biol Chem*. 1994; 269:2068–2074. [PubMed: 8294458]
3. Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell*. 2012; 45:814–825. [PubMed: 22387027]
4. Ginno PA, Lim YW, Lott PL, Korf I, Chedin F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res*. 2013; 23:1590–1600. [PubMed: 23868195]
5. Chan YA, Aristizabal MJ, Lu PY, Luo Z, Hamza A, Kobor MS, et al. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet*. 2014; 10:e1004288. [PubMed: 24743342]
6. Wahba L, Costantino L, Tan FJ, 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. [PubMed: 27298336]
7. Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, Ginno PA, et al. Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol Cell*. 2016; 63:167–178. [PubMed: 27373332]
8. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol*. 2003; 4:442–451. [PubMed: 12679812]
9. Xu B, Clayton DA. RNA–DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA–DNA hybrids serving as primers. *EMBO J*. 1996; 15:3135–3143. [PubMed: 8670814]

10. Pohjoismaki JL, Holmes JB, Wood SR, Yang MY, Yasukawa T, Reyes A, et al. Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of RNA/DNA hybrid. *J Mol Biol.* 2010; 397:1144–1155. [PubMed: 20184890]
11. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell.* 2011; 42:794–805. [PubMed: 21700224]
12. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature.* 2014; 516:436–439. [PubMed: 25296254]
13. Sollier J, Stork CT, Garcia-Rubio ML, Paulsen RD, Aguilera A, Cimprich KA. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell.* 2014; 56:777–785. [PubMed: 25435140]
14. Li X, Manley JL. Cotranscriptional processes and their influence on genome stability. *Genes Dev.* 2006; 20:1838–1847. [PubMed: 16847344]
15. Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* 2014; 28:1384–1396. [PubMed: 24990962]
16. Sollier J, Cimprich KA. Breaking bad: R-loops and genome integrity. *Trends Cell Biol.* 2015; 25:514–522. [PubMed: 26045257]
17. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat Rev Genet.* 2015; 16:583–597. [PubMed: 26370899]
18. Hamperl S, Cimprich KA. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair.* 2014; 19:84–94. [PubMed: 24746923]
19. Groh M, Gromak N. Out of balance: R-loops in human disease. *PLoS Genet.* 2014; 10:e1004630. [PubMed: 25233079]
20. Huertas P, Aguilera A. Cotranscriptionally formed DNA: RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell.* 2003; 12:711–721. [PubMed: 14527416]
21. Dominguez-Sanchez MS, Barroso S, Gomez-Gonzalez B, Luna R, Aguilera A. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* 2011; 7:e1002386. [PubMed: 22144908]
22. Li X, Manley JL. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell.* 2005; 122:365–378. [PubMed: 16096057]
23. Wahba L, Amon JD, 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. [PubMed: 22195970]
24. Paulsen RD, Soni DV, Wollman R, Hahn AT, Yee MC, Guan A, et al. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell.* 2009; 35:228–239. [PubMed: 19647519]
25. Stirling PC, Chan YA, Minaker SW, Aristizabal MJ, Barrett I, Sipahimalani P, et al. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev.* 2012; 26:163–175. [PubMed: 22279048]
26. Li X, Niu T, Manley JL. The RNA binding protein RNPS1 alleviates ASF/SF2 depletion-induced genomic instability. *RNA.* 2007; 13:2108–2115. [PubMed: 17959926]
27. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell.* 2011; 144:353–363. [PubMed: 21255825]
28. Tuduri S, Crabbe L, Conti C, Tourriere H, Holtgreve-Grez H, Jauch A, et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol.* 2009; 11:1315–1324. [PubMed: 19838172]
29. El Hage A, French SL, Beyer AL, Tollervey D. Loss of topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* 2010; 24:1546–1558. [PubMed: 20634320]
30. Tourriere H, Pasero P. Maintenance of fork integrity at damaged DNA and natural pause sites. *DNA Repair.* 2007; 6:900–913. [PubMed: 17379579]

31. Castellano-Pozo M, Garcia-Muse T, Aguilera A. R-loops cause replication impairment and genome instability during meiosis. *EMBO Rep.* 2012; 13:923–929. [PubMed: 22878416]
32. Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL, et al. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* 2011; 25:2041–2056. [PubMed: 21979917]
33. Wellinger RE, Prado F, Aguilera A. Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. *Mol Cell Biol.* 2006; 26:3327–3334. [PubMed: 16581804]
34. Herrera-Moyano E, Mergui X, Garcia-Rubio ML, 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. [PubMed: 24636987]
35. Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell.* 2011; 44:966–977. [PubMed: 22195969]
36. Pearson CE, Nichol Edamura K, Cleary JD. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet.* 2005; 6:729–742. [PubMed: 16205713]
37. Lopez Castel A, Cleary JD, Pearson CE. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol.* 2010; 11:165–170. [PubMed: 20177394]
38. McIvor EI, Polak U, Napierala M. New insights into repeat instability: role of RNA*DNA hybrids. *RNA Biol.* 2010; 7:551–558. [PubMed: 20729633]
39. Reddy K, Tam M, Bowater RP, Barber M, Tomlinson M, Nichol Edamura K, et al. Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res.* 2011; 39:1749–1762. [PubMed: 21051337]
40. Grabczyk E, Mancuso M, Sammarco MC. A persistent RNA.DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP *in vitro*. *Nucleic Acids Res.* 2007; 35:5351–5359. [PubMed: 17693431]
41. Lin Y, Dent SY, Wilson JH, Wells RD, Napierala M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc Natl Acad Sci U S A.* 2010; 107:692–697. [PubMed: 20080737]
42. Bowater RP, Jaworski A, Larson JE, Parniewski P, Wells RD. Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in *Escherichia coli*. *Nucleic Acids Res.* 1997; 25:2861–2868. [PubMed: 9207036]
43. Jung J, Bonini N. CREB-binding protein modulates repeat instability in a *Drosophila* model for polyQ disease. *Science (New York, NY).* 2007; 315:1857–1859.
44. Nakamori M, Pearson CE, Thornton CA. Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. *Hum Mol Genet.* 2011; 20:580–588. [PubMed: 21088112]
45. Batra R, Charizanis K, Swanson MS. Partners in crime: bidirectional transcription in unstable microsatellite disease. *Hum Mol Genet.* 2010; 19:R77–R82. [PubMed: 20368264]
46. Lin Y, Leng M, Wan M, Wilson JH. Convergent transcription through a long CAG tract destabilizes repeats and induces apoptosis. *Mol Cell Biol.* 2010; 30:4435–4451. [PubMed: 20647539]
47. Lin Y, Wilson JH. Nucleotide excision repair, mismatch repair, and R-loops modulate convergent transcription-induced cell death and repeat instability. *PLoS One.* 2012; 7:e46807. [PubMed: 23056461]
48. Reddy K, Schmidt MH, Geist JM, Thakkar NP, Panigrahi GB, Wang YH, et al. Processing of double-R-loops in (CAG).(CTG) and C9orf72 (GGGGCC).(GGCCCC) repeats causes instability. *Nucleic Acids Res.* 2014; 42:10,473–10,487.
49. Groh M, Lufino MM, 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. [PubMed: 24787137]
50. Wells RD. DNA triplexes and Friedreich ataxia. *FASEB J.* 2008; 22:1625–1634. [PubMed: 18211957]
51. Tous C, Aguilera A. Impairment of transcription elongation by R-loops *in vitro*. *Biochem Biophys Res Commun.* 2007; 360:428–432. [PubMed: 17603014]

52. Santoro MR, Bray SM, Warren ST. Molecular mechanisms of fragile X syndrome: a twenty-year perspective. *Annu Rev Pathol.* 2012; 7:219–245. [PubMed: 22017584]
53. Usdin K, Kumari D. Repeat-mediated epigenetic dysregulation of the FMR1 gene in the fragile X-related disorders. *Front Genet.* 2015; 6:192. [PubMed: 26089834]
54. Hagerman P. Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. *Acta Neuropathol.* 2013; 126:1–19. [PubMed: 23793382]
55. Loomis EW, Sanz LA, Chedin F, Hagerman PJ. Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.* 2014; 10:e1004294. [PubMed: 24743386]
56. Colak D, Zaninovic N, Cohen MS, Rosenwaks Z, Yang WY, Gerhardt J, et al. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science (New York, NY).* 2014; 343:1002–1005.
57. Eiges R, Urbach A, Malcov M, Frumkin T, Schwartz T, Amit A, et al. Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplantation genetically diagnosed embryos. *Cell Stem Cell.* 2007; 1:568–577. [PubMed: 18371394]
58. Zumwalt M, Ludwig A, Hagerman PJ, Dieckmann T. Secondary structure and dynamics of the r(CGG) repeat in the mRNA of the fragile X mental retardation 1 (FMR1) gene. *RNA Biol.* 2007; 4:93–100. [PubMed: 17962727]
59. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* 2004; 18:1618–1629. [PubMed: 15231739]
60. Gellert M, Lipsett MN, Davies DR. Helix formation by guanylic acid. *Proc Natl Acad Sci U S A.* 1962; 48:2013–2018. [PubMed: 13947099]
61. Huppert JL, Bugaut A, Kumari S, Balasubramanian S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* 2008; 36:6260–6268. [PubMed: 18832370]
62. Eddy J, Maizels N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res.* 2008; 36:1321–1333. [PubMed: 18187510]
63. Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature.* 2014; 507:195–200. [PubMed: 24598541]
64. Powell WT, Coulson RL, Gonzales ML, Crary FK, Wong SS, Adams S, et al. R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proc Natl Acad Sci U S A.* 2013; 110:13,938–13,943.
65. Boque-Sastre R, Soler M, Oliveira-Mateos C, Portela A, Moutinho C, Sayols S, et al. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc Natl Acad Sci U S A.* 2015; 112:5785–5790. [PubMed: 25902512]
66. Castellano-Pozo M, Santos-Pereira JM, Rondon AG, Barroso S, Andujar E, Perez-Alegre M, et al. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol Cell.* 2013; 52:583–590. [PubMed: 24211264]
67. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability—an evolving hallmark of cancer. *Nat Rev Mol Cell Biol.* 2010; 11:220–228. [PubMed: 20177397]
68. Loeb LA. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nat Rev Cancer.* 2011; 11:450–457. [PubMed: 21593786]
69. Schvartzman JM, Sotillo R, Benezra R. Mitotic chromosomal instability and cancer: mouse modelling of the human disease. *Nat Rev Cancer.* 2010; 10:102–115. [PubMed: 20094045]
70. Gorgoulis VG, Vassiliou LV, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, et al. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature.* 2005; 434:907–913. [PubMed: 15829965]
71. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (New York, NY).* 1994; 266:66–71.
72. Schlacher K, Wu H, Jasin M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell.* 2012; 22:106–116. [PubMed: 22789542]

73. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The breast cancer linkage consortium. *Am J Hum Genet.* 1998; 62:676–689. [PubMed: 9497246]
74. Savage KI, Gorski JJ, Barros EM, Irwin GW, Manti L, Powell AJ, et al. Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. *Mol Cell.* 2014; 54:445–459. [PubMed: 24746700]
75. Hill SJ, Rolland T, Adelmant G, Xia X, Owen MS, Dricot A, et al. Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev.* 2014; 28:1957–1975. [PubMed: 25184681]
76. Bhatia V, Barroso SI, Garcia-Rubio ML, 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. [PubMed: 24896180]
77. Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L, Yen A, Kamieniarz-Gdula K, et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol Cell.* 2015; 57:636–647. [PubMed: 25699710]
78. Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, et al. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med.* 2003; 348:1201–1214. [PubMed: 12660384]
79. Chernikova SB, Razorenova OV, Higgins JP, Sishc BJ, Nicolau M, Dorth JA, et al. Deficiency in mammalian histone H2B ubiquitin ligase Bre1 (Rnf20/Rnf40) leads to replication stress and chromosomal instability. *Cancer Res.* 2012; 72:2111–2119. [PubMed: 22354749]
80. Shema E, Tirosch I, Aylon Y, Huang J, Ye C, Moskovits N, et al. The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes Dev.* 2008; 22:2664–2676. [PubMed: 18832071]
81. Chen WD, Han ZJ, Skoletsy J, Olson J, Sah J, Myeroff L, et al. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J Natl Cancer Inst.* 2005; 97:1124–1132. [PubMed: 16077070]
82. Yang Y, McBride KM, Hensley S, Lu Y, Chedin F, Bedford MT. Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol Cell.* 2014; 53:484–497. [PubMed: 24507716]
83. Kupperts R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene.* 2001; 20:5580–5594. [PubMed: 11607811]
84. Shou Y, Martelli ML, Gabrea A, Qi Y, Brents LA, Roschke A, et al. Diverse karyotypic abnormalities of the c-myc locus associated with c-myc dysregulation and tumor progression in multiple myeloma. *Proc Natl Acad Sci U S A.* 2000; 97:228–233. [PubMed: 10618400]
85. Ruiz JF, Gomez-Gonzalez B, Aguilera A. AID induces double-strand breaks at immunoglobulin switch regions and c-MYC causing chromosomal translocations in yeast THO mutants. *PLoS Genet.* 2011; 7:e1002009. [PubMed: 21383964]
86. Kee Y, D'Andrea AD. Molecular pathogenesis and clinical management of Fanconi anemia. *J Clin Invest.* 2012; 122:3799–3806. [PubMed: 23114602]
87. Schwab RA, Nieminszczy J, Shah F, Langton J, Lopez Martinez D, Liang CC, et al. The Fanconi anemia pathway maintains genome stability by coordinating replication and transcription. *Mol Cell.* 2015; 60:351–361. [PubMed: 26593718]
88. Garcia-Rubio ML, Perez-Calero C, Barroso SI, Tumini E, Herrera-Moyano E, Rosado IV, et al. The Fanconi anemia pathway protects genome integrity from R-loops. *PLoS Genet.* 2015; 11:e1005674. [PubMed: 26584049]
89. Wu Y, Shin-ya K, Brosh RM Jr. FANCD1 helicase defective in Fanconi anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol.* 2008; 28:4116–4128. [PubMed: 18426915]
90. Xiao Y, Chen J, Liao Q, Wu Y, Peng C, Chen X. Lytic infection of Kaposi's sarcoma-associated herpesvirus induces DNA double-strand breaks and impairs non-homologous end joining. *J Gen Virol.* 2013; 94:1870–1875. [PubMed: 23677788]

91. Pan H, Zhou F, Gao SJ. Kaposi's sarcoma-associated herpesvirus induction of chromosome instability in primary human endothelial cells. *Cancer Res.* 2004; 64:4064–4068. [PubMed: 15205312]
92. Jackson BR, Noerenberg M, Whitehouse A. A novel mechanism inducing genome instability in Kaposi's sarcoma-associated herpesvirus infected cells. *PLoS Pathog.* 2014; 10:e1004098. [PubMed: 24788796]
93. Dominguez-Sanchez MS, Saez C, Japon MA, Aguilera A, Luna R. Differential expression of THOC1 and ALY mRNP biogenesis/export factors in human cancers. *BMC Cancer.* 2011; 11:77. [PubMed: 21329510]
94. Renton AE, Chio A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci.* 2014; 17:17–23. [PubMed: 24369373]
95. Haverkamp LJ, Appel V, Appel SH. Natural history of amyotrophic lateral sclerosis in a database population. Validation of a scoring system and a model for survival prediction. *Brain.* 1995; 118:707–719. [PubMed: 7600088]
96. Kiernan MC, Vucic S, Cheah BC, Turner MR, Eisen A, Hardiman O, et al. Amyotrophic lateral sclerosis. *Lancet.* 2011; 377:942–955. [PubMed: 21296405]
97. Salvi JS, Mekhail K. R-loops highlight the nucleus in ALS. *Nucleus.* 2015; 6:23–29. [PubMed: 25587791]
98. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron.* 2011; 72:245–256. [PubMed: 21944778]
99. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron.* 2011; 72:257–268. [PubMed: 21944779]
100. Majounie E, Renton AE, Mok K, Doppler EG, Waite A, Rollinson S, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 2012; 11:323–330. [PubMed: 22406228]
101. Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology.* 1998; 51:1546–1554. [PubMed: 9855500]
102. Rutherford NJ, Heckman MG, DeJesus-Hernandez M, Baker MC, Soto-Ortolaza AI, Rayaprolu S, et al. Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. *Neurobiol Aging.* Dec; 2012 33(12):2950.e5–2950.e7.
103. Fratta P, Mizielska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, et al. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci Rep.* 2012; 2:1016. [PubMed: 23264878]
104. Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol.* 2013; 74:180–187. [PubMed: 23720273]
105. Waite AJ, Baumer D, East S, Neal J, Morris HR, Ansorge O, et al. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol Aging.* Jul; 2014 35(7):1779.e5–1779.e13.
106. Belzil VV, Bauer PO, Prudencio M, Gendron TF, Stetler CT, Yan IK, et al. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* 2013; 126:895–905. [PubMed: 24166615]
107. Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, et al. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet.* 2013; 92:981–989. [PubMed: 23731538]
108. Liu EY, Russ J, Wu K, Neal D, Suh E, McNally AG, et al. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathol.* 2014; 128:525–541. [PubMed: 24806409]

109. Russ J, Liu EY, Wu K, Neal D, Suh E, Irwin DJ, et al. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathol.* 2015; 129:39–52. [PubMed: 25388784]
110. Belzil VV, Bauer PO, Gendron TF, Murray ME, Dickson D, Petrucelli L. Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res.* 2014; 1584:15–21. [PubMed: 24530272]
111. Xi Z, Rainero I, Rubino E, Pinessi L, Bruni AC, Maletta RG, et al. Hypermethylation of the CpG-island near the C9orf72 G(4)C(2)-repeat expansion in FTL D patients. *Hum Mol Genet.* 2014; 23:5630–5637. [PubMed: 24908669]
112. Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, Lewis J, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A.* 2013; 110:E4968–E4977. [PubMed: 24248382]
113. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, et al. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* 2013; 126:829–844. [PubMed: 24129584]
114. Mori K, Arzberger T, Grasser FA, Gijssels I, May S, Rentzsch K, et al. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* 2013; 126:881–893. [PubMed: 24132570]
115. Mizielińska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, et al. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* 2013; 126:845–857. [PubMed: 24170096]
116. Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A.* 2013; 110:E4530–E4539. [PubMed: 24170860]
117. Wang J, Haeusler AR, Simko EA. Emerging role of RNA*DNA hybrids in C9orf72-linked neurodegeneration. *Cell Cycle.* 2015; 14:526–532. [PubMed: 25590632]
118. Therrien M, Rouleau GA, Dion PA, Parker JA. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*. *PLoS One.* 2013; 8:e83450. [PubMed: 24349511]
119. Sellier C, Campanari ML, Julie Corbier C, Gaucherot A, Kolb-Cheynel I, Oulad-Abdelghani M, et al. Loss of C9ORF72 impairs autophagy and synergizes with polyQ ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* 2016; 35:1276–1297. [PubMed: 27103069]
120. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL D/ALS. *Science (New York, NY).* 2013; 339:1335–1338.
121. Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron.* 2013; 77:639–646. [PubMed: 23415312]
122. He F, Jones JM, Figueroa-Romero C, Zhang D, Feldman EL, Goutman SA, et al. Screening for novel hexanucleotide repeat expansions at ALS- and FTD-associated loci. *Neurol Genet.* 2016; 2:e71. [PubMed: 27274540]
123. Rollinson S, Halliwell N, Young K, Callister JB, Toulson G, Gibbons L, et al. Analysis of the hexanucleotide repeat in C9ORF72 in Alzheimer's disease. *Neurobiol Aging.* 1846; 2012(33):e5–e6.
124. Kohli MA, John-Williams K, Rajbhandary R, Naj A, Whitehead P, Hamilton K, et al. Repeat expansions in the C9ORF72 gene contribute to Alzheimer's disease in Caucasians. *Neurobiol Aging.* May; 2013 34(5):1519.e5–1519.e12.
125. Majounie E, Abramzon Y, Renton AE, Perry R, Bassett SS, Pletnikova O, et al. Repeat expansion in C9ORF72 in Alzheimer's disease. *N Engl J Med.* 2012; 366:283–284. [PubMed: 22216764]
126. Harms M, Benitez BA, Cairns N, Cooper B, Cooper P, Mayo K, et al. C9orf72 hexanucleotide repeat expansions in clinical Alzheimer disease. *JAMA Neurol.* 2013; 70:736–741. [PubMed: 23588422]

127. Hensman Moss DJ, Poulter M, Beck J, Hehir J, Polke JM, Campbell T, et al. C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. *Neurology*. 2014; 82:292–299. [PubMed: 24363131]
128. Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. 2010; 466:1069–1075. [PubMed: 20740007]
129. van Blitterswijk M, Mullen B, Heckman MG, Baker MC, DeJesus-Hernandez M, Brown PH, et al. Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. *Neurobiol Aging*. Oct; 2014 35(10):2421.e13–2421.e17.
130. Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, et al. Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet*. 1996; 14:269–276. [PubMed: 8896555]
131. Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, et al. Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet*. 1996; 14:285–291. [PubMed: 8896557]
132. Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, et al. Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet*. 1996; 14:277–284. [PubMed: 8896556]
133. Farg MA, Soo KY, Warraich ST, Sundaramoorthy V, Blair IP, Atkin JD. Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. *Hum Mol Genet*. 2013; 22:717–728. [PubMed: 23172909]
134. Salvi JS, Chan JN, Szafranski K, Liu TT, Wu JD, Olsen JB, et al. Roles for Pbp1 and caloric restriction in genome and lifespan maintenance via suppression of RNA–DNA hybrids. *Dev Cell*. 2014; 30:177–191. [PubMed: 25073155]
135. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, NY)*. 2006; 314:130–133.
136. Colombrita C, Zennaro E, Fallini C, Weber M, Sommacal A, Buratti E, et al. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem*. 2009; 111:1051–1061. [PubMed: 19765185]
137. Nonhoff U, Ralser M, Welzel F, Piccini I, Balzereit D, Yaspo ML, et al. Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Mol Biol Cell*. 2007; 18:1385–1396. [PubMed: 17392519]
138. Moreira MC, Klur S, Watanabe M, Nemeth AH, Le Ber I, Moniz JC, et al. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet*. 2004; 36:225–227. [PubMed: 14770181]
139. Chen YZ, Bennett CL, Huynh HM, Blair IP, Puls I, Irobi J, et al. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet*. 2004; 74:1128–1135. [PubMed: 15106121]
140. Bassuk AG, Chen YZ, Batish SD, Nagan N, Opal P, Chance PF, et al. In cis autosomal dominant mutation of senataxin associated with tremor/ataxia syndrome. *Neurogenetics*. 2007; 8:45–49. [PubMed: 17096168]
141. Zhao ZH, Chen WZ, Wu ZY, Wang N, Zhao GX, Chen WJ, et al. A novel mutation in the senataxin gene identified in a Chinese patient with sporadic amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2009; 10:118–122.
142. Hirano M, Quinzii CM, Mitsumoto H, Hays AP, Roberts JK, Richard P, et al. Senataxin mutations and amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2011; 12:223–227. [PubMed: 21190393]
143. Rudnik-Schoneborn S, Arning L, Epplen JT, Zerres K. SETX gene mutation in a family diagnosed autosomal dominant proximal spinal muscular atrophy. *Neuromuscul Disord*. 2012; 22:258–262. [PubMed: 22088787]
144. Saracchi E, Castelli M, Bassi MT, Brighina E, Cereda D, Marzorati L, et al. A novel heterozygous SETX mutation in a patient presenting with chorea and motor neuron disease. *Amyotroph Lateral Scler Frontotemporal Degener*. 2014; 15:138–140. [PubMed: 24533459]

145. Mancini C, Orsi L, Guo Y, Li J, Chen Y, Wang F, et al. An atypical form of AOA2 with myoclonus associated with mutations in SETX and AFG3L2. *BMC Med Genet.* 2015; 16:16. [PubMed: 25927548]
146. Steinmetz EJ, Conrad NK, Brow DA, Corden JL. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature.* 2001; 413:327–331. [PubMed: 11565036]
147. Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell.* 2006; 24:735–746. [PubMed: 17157256]
148. Mischo HE, Gomez-Gonzalez B, Grzechnik P, Rondon AG, Wei W, Steinmetz L, et al. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell.* 2011; 41:21–32. [PubMed: 21211720]
149. Becherel OJ, Yeo AJ, Stellati A, Heng EY, Luff J, Suraweera AM, et al. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet.* 2013; 9:e1003435. [PubMed: 23593030]
150. Alzu A, Bermejo R, Begnis M, Lucca C, Piccini D, Carotenuto W, et al. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell.* 2012; 151:835–846. [PubMed: 23141540]
151. Yuce O, West SC. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Mol Cell Biol.* 2013; 33:406–417. [PubMed: 23149945]
152. Richard P, Feng S, Manley JL. A SUMO-dependent interaction between senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes Dev.* 2013; 27:2227–2232. [PubMed: 24105744]
153. Richard P, Manley JL. SETX sumoylation: a link between DNA damage and RNA surveillance disrupted in AOA2. *Rare Dis.* Jan-Dec;2014 2(1):1–6.
154. Yeo AJ, Becherel OJ, Luff JE, Cullen JK, Wongsurawat T, Jenjaroenpoon P, et al. R-loops in proliferating cells but not in the brain: implications for AOA2 and other autosomal recessive ataxias. *PLoS One.* Mar 17.2014 9(3):e90219. [PubMed: 24637776]
155. Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K, et al. Telomeric RNA–DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol.* 2013; 20:1199–1205. [PubMed: 24013207]
156. 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. [PubMed: 24084588]
157. De Amicis A, Piane M, Ferrari F, Fanciulli M, Delia D, Chessa L. Role of senataxin in DNA damage and telomeric stability. *DNA Repair.* 2011; 10:199–209. [PubMed: 21112256]
158. Kaneko S, Rozenblatt-Rosen O, Meyerson M, Manley JL. The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes Dev.* 2007; 21:1779–1789. [PubMed: 17639083]
159. West S, Gromak N, Proudfoot NJ. Human 5' → 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature.* 2004; 432:522–525. [PubMed: 15565158]
160. Morales JC, Richard P, Patidar PL, Motea EA, Dang TT, Manley JL, et al. XRN2 links transcription termination to DNA damage and replication stress. *PLoS Genet.* 2016; 12:e1006107. [PubMed: 27437695]
161. Rice G, Patrick T, Parmar R, Taylor CF, Aeby A, Aicardi J, et al. Clinical and molecular phenotype of Aicardi–Goutieres syndrome. *Am J Hum Genet.* 2007; 81:713–725. [PubMed: 17846997]
162. Lim YW, Sanz LA, Xu X, Hartono SR, Chedin F. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutieres syndrome. *eLife.* 2015; 4
163. Lee-Kirsch MA, Gong M, Chowdhury D, Senenko L, Engel K, Lee YA, et al. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. *Nat Genet.* 2007; 39:1065–1067. [PubMed: 17660818]

164. Gunther C, Kind B, Reijns MA, Berndt N, Martinez-Bueno M, Wolf C, et al. Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. *J Clin Invest*. 2015; 125:413–424. [PubMed: 25500883]
165. Dykens EM, Lee E, Roof E. Prader–Willi syndrome and autism spectrum disorders: an evolving story. *J Neurodev Disord*. 2011; 3:225–237. [PubMed: 21858456]
166. Chamberlain SJ, Lalande M. Angelman syndrome, a genomic imprinting disorder of the brain. *J Neurosci*. 2010; 30:9958–9963. [PubMed: 20668179]
167. Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader–Willi syndrome. *Genet Med*. 2012; 14:10–26. [PubMed: 22237428]
168. Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, et al. Prader–Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet*. 2008; 40:719–721. [PubMed: 18500341]
169. El Achkar E, Gerbault-Seureau M, Muleris M, Dutrillaux B, Debatisse M. Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites. *Proc Natl Acad Sci U S A*. 2005; 102:18,069–18,074.
170. Wahba L, Gore SK, Koshland D. The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife*. 2013; 2:e00505. [PubMed: 23795288]

Abbreviations

DSB	DNA double-strand break
RNAP	RNA polymerase
FXS	fragile X syndrome
FXTAS	fragile X-associated tremor/ataxia syndrome
ssDNA	single-stranded DNA
G-Q	G-quadruplex
KD	knockdown
TDRD3	tudor domain-containing protein 3
FA	Fanconi anemia
ALS	amyotrophic lateral sclerosis
FTD	frontotemporal dementia
ATXN2	Ataxin-2
SCA2	spinocerebellar ataxia type 2
TDP-43	TAR DNA-binding protein 43
SETX	senataxin
AOA2	ataxia with oculomotor apraxia type 2
AGS	Aicardi–Goutières syndrome
PWS	Prader–Willi syndrome

Table 1
List of diseases and associated genes for which a connection to R loops has been shown or is highly suspected

Disease	R-loop-associated genes (repeats)n	References [In vivo detection of R loops at expanded repeats (#)]
DM1 (myotonic dystrophy type 1)	<i>DMPK</i>	[39,41 ^(#)]
HD (Huntington's disease)	<i>HTT</i> (CAG) <i>n</i>	[39]
FRAXE (fragile X syndrome type E)	<i>FRM2</i> (CCG) <i>n</i>	[39]
FXS or FRAXA (fragile X syndrome type A)	* <i>FMR1</i> (CGG) <i>n</i>	[39,49 ^(#) ,55 ^(#) ,56 ^(#)]
FXTAS (fragile X tremor ataxia syndrome)	* <i>FRM1/FRM4</i> (CGG) <i>n</i>	[39,55 ^(#)]
SCAs (spinocerebellar ataxias)	<i>ATXN1, ATXN2</i> (CAG) <i>n</i>	[39]
FRDA (Friedreich ataxia)	* <i>FXN</i> (GAA) <i>n</i>	[39,40,49 ^(#)]
ALS/FTD (frontotemporal dementia)	<i>C9orf72</i> (GGGGCC) <i>n</i>	[48,63]
Breast and ovarian cancer	<i>BRCA1, BRCA2</i>	[75, 76, 77]
Eosinophilic leukemia	<i>FIP1</i>	[25]
Seminoma	<i>BRE1</i>	[79]
Colon cancer	* <i>VIM</i>	[65]
FA (Fanconi anemia)	FA genes, e.g., <i>FANCM, FANCD2</i>	[87,88]
ALS (amyotrophic lateral sclerosis)	<i>ATXN2</i>	[134] (yeast)
AOA2 (ataxia with oculomotor apraxia)	<i>SETX</i>	[11,149,151,154]
AGS (Aicardi–Goutières syndrome)	<i>TREX1, RNASEH2, SAMHD1</i>	[162]
PWS (Prader–Willi syndrome)	* <i>SNORD116</i>	[64]
Burkitt's lymphoma	* <i>c-MYC</i> and <i>IgH</i>	[82]

Disease	R-loop-associated genes (repeats)n	References [In vivo detection of R loops at expanded repeats (#)]
MM (Multiple myeloma)	<i>*c-MYC and IgH</i>	[82]
AIDS-associated malignancies	<i>hTREX</i>	[92]

In blue, genes where R loops form (*shown at endogenous locus) or may form (lack of in vivo evidence at endogenous locus). In gray, genes encoding proteins involved in R-loop prevention or removal.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript