Author Manuscript Author Manuscript Author Manuscript Author Manuscript

HHS Public Access
Author manuscript
Curr Genet. Author manuscript; available in PMC 2019 August 01.

Published in final edited form as:
Curr Genet. 2018 August ; 64(4): 789–794. doi:10.1007/s00294-018-0806-z.

R-loops: Targets for Nuclease Cleavage and Repeat Instability

Catherine H. Freudenreich[1,2,*]

[1]Department of Biology, Tufts University, Medford MA 02155, USA

[2]Program in Genetics, Tufts University, Medford MA 02155, USA

Abstract

R-loops form when transcribed RNA remains bound to its DNA template to form a stable RNA:DNA hybrid. Stable R-loops form when the RNA is purine rich, and are further stabilized by DNA secondary structures on the non-template strand. Interestingly, many expandable and disease-causing repeat sequences form stable R-loops, and R-loops can contribute to repeat instability. Repeat expansions are responsible for multiple neurodegenerative diseases, including Huntington's disease, myotonic dystrophy, and several types of ataxias. Recently, it was found that R-loops at an expanded CAG/CTG repeat tract cause DNA breaks as well as repeat instability (Su and Freudenreich, 2017). Two factors were identified as causing R-loop-dependent breaks at CAG/CTG tracts: deamination of cytosines and the MutLγ (Mlh1-Mlh3) endonuclease, defining two new mechanisms for how R-loops can generate DNA breaks (Su and Freudenreich, 2017). Following R-loop-dependent nicking, base excision repair resulted in repeat instability. These results have implications for human repeat expansion diseases and provide a paradigm for how RNA:DNA hybrids can cause genome instability at structure-forming DNA sequences. This perspective summarizes mechanisms of R-loop-induced fragility at G-rich repeats and new links between DNA breaks and repeat instability.

Keywords

R-loop; trinucleotide repeat instability; chromosome fragility; cytosine deamination; base excision repair (BER); MutLγ (Mlh1-Mlh3)

R-loops at G-rich Repeats

R-loops are stable RNA:DNA hybrids that can form during transcription by reannealing of the newly synthesized RNA transcript to the DNA template. In addition to these co-transcriptional R-loops, RNA-DNA hybrids can form in trans, with the help of proteins such as Rad51 or Rad52 to assist in the homology search and annealing steps of R-loop formation (Keskin et al., 2014; Wahba et al., 2013). Recent genome-wide mapping approaches have indicated that R-loops are fairly common in eukaryotic genomes. In yeast, they comprise about 8% of the genome (Chan et al., 2014; Wahba et al., 2016). In mammalian genomes, nearly 5% of the genome is engaged in R-loops (Sanz et al., 2016). In particular, promoter

[*]corresponding author: Tel: 617-627-4037, catherine.freudenreich@tufts.edu.

Author Manuscript

and terminator regions are over-represented, and there is evidence that they play a role in transcription termination (reviewed in (Chedin, 2016)).

Certain sequence characteristics favor the formation of R-loops. Thermodynamic analysis indicates that the most stable hybrid forms between a purine rich RNA transcript and a complementary pyrimidine-rich DNA strand, so RNA:DNA hybrids comprised of G-rich RNA are more stable than the corresponding DNA:DNA duplex (Ratmeyer et al., 1994; Roberts and Crothers, 1992). In addition, intra-strand structure formation on the displaced G-rich DNA strand can further stabilize R-loops (Maizels and Gray, 2013; Skourti-Stathaki and Proudfoot, 2014). Mammalian switch regions were one of the first well-characterized G-rich R-loop forming sequences, and have both of these characteristics. For example it was shown that G-quadruplex formation in the non-template strand stabilizes R-loops that form in the Sμ region of the human immunoglobulin heavy chain locus (Ig locus) in mammalian B cells (Duquette et al., 2004) reviewed in (Santos-Pereira and Aguilera, 2015). Other G-rich repetitive DNA has also been shown to form particularly stable R-loops, including G_4C_2/C_4G_2, CGG/GCC and CAG/CTG repeats, all of which can also form secondary structures (G quadruplexes or hairpins) (Belotserkovskii et al., 2010; Belotserkovskii et al., 2013; Groh et al., 2014; Kim and Jinks-Robertson, 2011; Loomis et al., 2014; Reddy et al., 2011; Salvi and Mekhail, 2015). Also, the purine-rich GAA/TTC repeat forms very stable R-loops when transcribed such that the GAA purine strand is the nascent RNA (Grabczyk et al., 2007; Groh et al., 2014; Reddy et al., 2011). The R-loop-forming repeats listed above all have the additional characteristic of expanding once they reach a threshold length, and these expansions can lead to various severe genetic diseases. Expansion of the G_4C_2 repeat is one cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), CGG expansion causes Fragile X syndrome, GAA expansion causes Friedrich's ataxia, and CAG/CTG expansion is the cause of 13 different diseases, including Huntington's disease (HD), Myotonic dystrophy type 1 (DM1), and multiple spinocerebellar ataxias (SCAs) (McMurray, 2010; Usdin et al., 2015). Indeed, R-loops have been shown to cause repeat instability at expandable triplet repeat sequences in several human cell lines as well as *in vitro* systems designed to measure repeat instability. For example, R-loops form *in vivo* and *in vitro* upon transcription through a CAG/CTG (CAG) repeat region, with the level of R-loop formation increasing with CAG tract length (Lin et al., 2010; Reddy et al., 2014; Reddy et al., 2011). R-loop formation destabilizes expanded CAG repeats, as depletion of RNaseH stimulates repeat contractions (Lin et al., 2010; Reddy et al., 2011; Su and Freudenreich, 2017). Similarly, transcription through an expanded GAA/TCC repeat promotes expansions (Ditch et al., 2009; McGinty et al., 2017a; Rindler and Bidichandani, 2011), though a direct link to R-loops has not been made in this case.

Links between R-loops, Cytosine Deamination, BER and Repeat Instability

The formation of stable R-loops within structure-forming G/C-rich repeats has several types of deleterious consequences. The displaced single-strand of DNA (ssDNA) is exposed to base damage or ssDNA nucleases, which can cause nicks in the DNA. During class-switch recombination (CSR), the stable R-loops formed in the Sμ region exposes cytosines on the displaced strand, which are deaminated to uracil by activation-induced deaminase (AID) (Stavnezer et al., 2008). A recent study showed that G-quadruplexes, like those formed at Sμ

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

region R loops, are particularly good AID binding substrates, explaining how the Ig switch region is a preferred AID target (Qiao et al., 2017). Uracil DNA glycosylase (Ung1) dependent base-excision repair (BER) then creates single-stranded breaks (SSBs) by removal of uracil, followed by APE1- and APE2-dependent cleavage of abasic sites (Bregenhorn et al., 2016; Guikema et al., 2007; Schrader et al., 2009). These SSBs can be converted to DSBs when close to each other on opposite strands, and re-ligation of these DSBs achieves CSR. Though in this region the nicks and breaks can lead to a desired outcome, this is not always the case. We recently found that cytosine deamination also leads to breaks in an expanded CAG/CTG tract engaged in a co-transcriptional R-loop (Su and Freudenreich, 2017). In this case, since the experiment was done in a budding yeast model system, the endogenous yeast Fcy1 deaminase was involved. Fcy1 is part of the CDA family of deaminases also found in mammalian cells. Deletion of the *FCY1* gene suppressed approximately 60% of the R-loop induced DSBs and also completely suppressed the R-loop-induced CAG contractions that were observed. We showed that the Fcy1-dependent repeat contractions were occurring due to BER, and required the Ung1 glycosylase and the Apn1 endonuclease (Figure 1).

Ung1 targets uracil in DNA, the product of cytosine deamination, to create an abasic site, which is acted upon by an AP endonuclease (e.g. Apn1/APE) to create a nick. This normally creates a substrate for BER repair, using either the Polβ or Polδ polymerase (Liu and Wilson, 2012). Interestingly, a hypomorphic Polβ mutation reduces expansion frequency in a mouse model of Fragile X-related disorders (Lokanga et al., 2015), and BER has been shown to be a mechanism that can create both CAG repeat contractions and expansions in *in vitro* systems (Lai et al., 2016; Liu and Wilson, 2012). There is evidence that BER-induced triplet repeat expansions can occur downstream of oxidative damage, because mutation of glycosylases OGG1 and NEIL1, that target oxidized bases, suppress somatic CAG repeat expansions in mouse models of HD (Kovtun et al., 2007; Mollersen et al., 2012). Also, treatment of cells with oxidative DNA-damaging agents promotes repeat instability in multiple cell types (Entezam et al., 2010; Kovtun et al., 2007; Polyzos et al., 2016). However, in light of the effects of cytosine deamination and Ung1 in the yeast model, it may be worth investigating whether DNA targeted deaminases could trigger repeat instability in human cells. CAG instability is high in the liver in mouse models of HD, a tissue that expresses human cytidine deaminase (CDA)(www.proteinatlas.org). For AID, cytosines within the single-stranded overhangs adjacent to G-quadruplexes were preferred deamination substrates (Qiao et al., 2017). A similar mechanism could conceivably favor deamination of C's near other types of DNA structures such as CNG hairpins. The known disease-associated repeat expansions in humans are associated with transcribed loci known to form R-loops, thus potentially exposing ssDNA to various types of DNA damage. Strand-specific DNA nicking followed by repair provides a mechanism to explain R-loop-induced repeat instability in non-dividing cells (Figure 1).

The role of MutLγ in R-loop-induced Repeat Fragility and Instability

In mouse models, both intergenerational and somatic CAG repeat expansions can be suppressed by mutation of various proteins normally associated with MMR, indicating that in the context of structure-forming repeats they are playing an alternative and mutagenic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

role. The connection between MMR proteins and repeat instability has been extensively studied and verified in multiple model systems from human cell lines to yeast to *in vitro* repair assays (reviewed in (Polyzos and McMurray, 2017; Schmidt and Pearson, 2016; Usdin et al., 2015)). The MMR components that cause repeat instability include members of the MutSβ recognition complex (Msh2, Msh3) and MutLα nuclease components (Mlh1, Pms2) (see (Usdin et al., 2015) for review). There is a second MutL nuclease complex, MutLγ (Mlh1, Mlh3) that has a less understood cellular role. Interestingly, in a mouse model of HD, mutation of either of the MutLγ nuclease components (Mlh1, Mlh3) completely suppressed somatic repeat expansions (Pinto et al., 2013). This finding was in keeping with the interpretation that an alternative pathway is operative, as MutLγ has only a minor role in canonical mismatch repair (Flores-Rozas and Kolodner, 1998; Romanova and Crouse, 2013). Its most well-characterized function is in processing Holliday junctions in meiosis, and it can bind and supershift a 4-way junction *in vitro* (Ranjha et al., 2014; Rogacheva et al., 2014).

We identified MutLγ nuclease activity as a cause of both R-loop-induced CAG fragility and contractions (Su and Freudenreich, 2017). This provides a possible mechanism for how MutL nucleases are causing instability: by nicking of R-loop-induced structures. The genetic data was consistent with MutLγ action being partially dependent on cytosine deamination and Apn1 activity. Since Apn1 acts on dsDNA, these results suggest that the relevant MutLγ substrate may be created by R-loop displacement, which can result in a slipped-strand structure with hairpins on both strands (Reddy et al., 2014). Based on its inherent endonuclease activity on supercoiled DNA, MutLγ could create nicks at the hairpin junctions, which are open and accessible; nicks on both strands could be converted into a DSB (Figure 1). In addition, results show that the MutLγ endonuclease activity can be directed by a nick to cleave the opposite strand in a concerted mechanism to create a DSB (Manhart et al., 2017). This mechanism could account for the partial dependence on Fcy1: an initial nick created by processing of a deaminated cytosine could direct MutLγ cleavage to the opposite strand to create a DSB (Figure 1). This same concerted mechanism can also occur without a preexisting nick, so another possibility is that recruitment to the DNA structure (perhaps by a MutS complex) initiates MutLγ non-directional cleavage in the vicinity (Manhart et al., 2017).

The state of the Mlh1-Mlh3 nuclease appears to be highly relevant to the propensity of CAG repeats to expand in mammals. A study of two Huntington's disease mouse strain models showed that there was a modifier in the Mlh1 gene that determined the somatic expansion difference between the two strains (Pinto et al., 2013). Further analysis showed that expansions were dependent on both Mlh1 and Mlh3, implicating the MutLγ complex as a key driver of somatic CAG expansions in this mouse model (Pinto et al., 2013). Recently, sequencing of a large human cohort identified the Mlh1 gene as a modifier of Huntington's disease onset (Lee et al., 2017; Lee JM, 2015). MutSβ complex members that could recruit the MutLγ nuclease have also been identified as modifiers of somatic repeat instability in both Huntington's and myotonic dystrophy type 1 patients (Morales et al., 2016). Thus, the mechanism of MutLγ action on R-loops and CAG/CTG repeat structures deserves further study.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Involvement of TCR and other Nucleases in R-loop-induced Breaks

There is evidence that transcription-coupled repair (TCR) and the TCR-associated endonucleases XPG (yeast Rad2) and XPF (yeast Rad1) are also involved in R-loop processing and instability of R-loop-forming repeat sequences, at least in some circumstances. In the human cell line CAG contraction selection system, high levels of induced transcription through CAG repeats were shown to induce contractions through TCR, as contractions were eliminated upon CSB, XPA, XPG or ERCC1 knockdown, but did not require the human APE1 BER nuclease (Lin et al., 2006; Lin and Wilson, 2007, 2012). Also, R-loops formed in human cells in the absence of RNA/DNA helicases Aquarius (AQR) or Senataxin (SETX) were processed into DSBs by TCR nucleases XPG and XPF (Sollier et al., 2014). Thus, it appears that different R-loops, caused by different means or at different sequences, may be processed into DSBs in multiple ways. One possibility is that transient RNA:DNA hybrids that occur in the context of highly induced transcription or in the absence of RNA helicases are more likely to stall RNA polymerase and trigger TCR. Interestingly, at the persistent R-loops occurring at a CAG repeat in the absence of yeast RNaseH1 and RNaseH2, TCR components did not cause fragility. However, a significant decrease in contractions was observed in the absence of either the Rad1 or Rad2 nucleases, though deletion of other TCR and NER proteins had no effect (Su and Freudenreich, 2017). By genetic analysis, the Rad1/Rad2-dependent contractions occurred in the same pathway as Ung1/Apn1, suggesting that the Rad1/XPF and Rad2/XPG nucleases may sometimes work in the context of the BER pathway, for example in the gap processing step (Figure 1). Thus, there may be some cross-talk between the two pathways during R-loop processing.

Unknowns and Future Directions

In the yeast system unlike in human cells, triplet repeat expansions are less common than contractions, thus a reduction of expansions was not evident in the non-selective system used to study the effect of cytosine deamination and MutLγ on repeat instability. Model systems that can select for rare expansion events (Keogh et al., 2017; Kim et al., 2017) could be used to learn more about the role of these factors on repeat expansions. However, the data from mouse models already strongly implicate both BER and the MutLγ complex as causative agents in repeat expansions. Also, evidence cited above shows that co-transcriptional R-loops cause repeat instability, which is exacerbated in RNaseH mutants. The missing connection is what role R-loops play in repeat expansion diseases. How often are R-loops involved in the initiating event that triggers BER or other pathways involving MMR proteins that result in repeat expansions?

A related question is what contribution each of the various mechanisms that can cause an initial nick or break has in initiating repeat instability. At expanded CAG tracts, evidence exists for oxidative damage and cytosine deamination, which could both be exacerbated by R-loops, as initiating mutational events. Evidence cited above also points to a role for endonuclease recognition of either the R-loop itself or resulting DNA structures by MutLγ and the TCR endonucleases. In dividing cells, breaks could also form due to transcription-replication collisions, which are known to cause genome instability within non-repetitive DNA (see for example (Brambati et al., 2015; Oestergaard and Lisby, 2017) for review).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Some of these pathways may work together or in sequence. Though expanded triplet repeats are a useful model as they are especially prone to mutation, these pathways are likely also operating at other R-loop or structure-forming sequences to cause mutations and genome instability.

Lastly, there appear to be multiple mechanisms that can be employed to repair the lesions, whatever the initial cause (R-loops or otherwise), and the interplay among these pathways can determine the mutagenic outcome. For nicks or gaps within CAG tracts, BER is an important pathway of repair. Interestingly, the MutSβ complex interacts with Polβ in the BER pathway to promote trinucleotide repeat expansion by stimulating Polβ synthesis across the repeat tract and influencing flap processing by FEN1 (Lai et al., 2016). Nuclease processing during BER, in gap creation (Figure 1) or at later steps (for example on the template strand to cause a contraction or fix an expansion) likely plays an important role in determining whether expansions or contractions occur (see (Lai et al., 2016; Polyzos and McMurray, 2017; Usdin et al., 2015) for models). It remains to be determined what role, if any, MutL complexes have in these downstream nuclease processing steps. Formation of DSBs by MutLγ or other nucleases opens up another suite of possibilities for repair pathways, including end joining and homologous recombination, that have been shown to generate both repeat instability and chromosomal rearrangements (reviewed in (Polleys et al., 2017), see also (McGinty et al., 2017b; Su and Freudenreich, 2017)). Thus, R-loop-induced breaks at DNA repeats could be a major mutagenic force in cells, able to cause both local instability and larger genome rearrangements.

Acknowledgments

Thanks to Xiaofeng Allen Su for help with the figure. The author's research is supported by the National Science Foundation (MCB1330743) and the National Institute of Health (GM122880 and GM105473).

References

Belotserkovskii BP, Liu R, Tornaletti S, Krasilnikova MM, Mirkin SM, Hanawalt PC. Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. Proc Natl Acad Sci U S A. 2010; 107:12816–12821. [PubMed: 20616059]

Belotserkovskii BP, Mirkin SM, Hanawalt PC. DNA sequences that interfere with transcription: implications for genome function and stability. Chem Rev. 2013; 113:8620–8637. [PubMed: 23972098]

Brambati A, Colosio A, Zardoni L, Galanti L, Liberi G. Replication and transcription on a collision course: eukaryotic regulation mechanisms and implications for DNA stability. Frontiers in genetics. 2015; 6:166. [PubMed: 25972894]

Bregenhorn S, Kallenberger L, Artola-Boran M, Pena-Diaz J, Jiricny J. Non-canonical uracil processing in DNA gives rise to double-strand breaks and deletions: relevance to class switch recombination. Nucleic acids research. 2016; 44:2691–2705. [PubMed: 26743004]

Chan YA, Aristizabal MJ, Lu PY, Luo Z, Hamza A, Kobor MS, Stirling PC, Hieter P. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS genetics. 2014; 10:e1004288. [PubMed: 24743342]

Chedin F. Nascent Connections: R-Loops and Chromatin Patterning. Trends in genetics: TIG. 2016; 32:828–838. [PubMed: 27793359]

Ditch S, Sammarco MC, Banerjee A, Grabczyk E. Progressive GAA.TTC repeat expansion in human cell lines. PLoS genetics. 2009; 5:e1000704. [PubMed: 19876374]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

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 & development. 2004; 18:1618–1629. [PubMed: 15231739]

Entezam A, Lokanga AR, Le W, Hoffman G, Usdin K. Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. Hum Mutat. 2010; 31:611–616. [PubMed: 20213777]

Flores-Rozas H, Kolodner RD. The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. Proc Natl Acad Sci U S A. 1998; 95:12404–12409. [PubMed: 9770499]

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 research. 2007; 35:5351–5359. [PubMed: 17693431]

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 genetics. 2014; 10:e1004318. [PubMed: 24787137]

Guikema JE, Linehan EK, Tsuchimoto D, Nakabeppu Y, Strauss PR, Stavnezer J, Schrader CE. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. J Exp Med. 2007; 204:3017–3026. [PubMed: 18025127]

Keogh N, Chan KY, Li GM, Lahue RS. MutSbeta abundance and Msh3 ATP hydrolysis activity are important drivers of CTG*CAG repeat expansions. Nucleic acids research. 2017; 45:10068–10078. [PubMed: 28973443]

Keskin H, Shen Y, Huang F, Patel M, Yang T, Ashley K, Mazin AV, Storici F. Transcript-RNA-templated DNA recombination and repair. Nature. 2014; 515:436–439. [PubMed: 25186730]

Kim JC, Harris ST, Dinter T, Shah KA, Mirkin SM. The role of break-induced replication in large-scale expansions of (CAG)n/(CTG)n repeats. Nature structural & molecular biology. 2017; 24:55–60.

Kim N, Jinks-Robertson S. Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. DNA repair. 2011; 10:953–960. [PubMed: 21813340]

Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature. 2007; 447:447–452. [PubMed: 17450122]

Lai Y, Budworth H, Beaver JM, Chan NL, Zhang Z, McMurray CT, Liu Y. Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. Nature communications. 2016; 7:12465.

Lee JM, Chao MJ, Harold D, Abu Elneel K, Gillis T, Holmans P, Jones L, Orth M, Myers RH, Kwak S, et al. A modifier of Huntington's disease onset at the MLH1 locus. Hum Mol Genet. 2017; 26:3859–3867. [PubMed: 28934397]

Lee JMWV, Chao MJ, Vonsattel JP, Pinto RM, Lucente D, Abu-Elneel K, Ramos EM, Mysore JS, Gillis T, MacDonald ME, Gusella JF, Harold D, Stone TC, Escott-Price V, Han J, Vedernikov A, Holmans P, Jones L, Kwak S, Mahmoudi M, Orth M, Landwehrmeyer GB, Paulsen JS, Dorsey ER, Shoulson I, Myers RH. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. Cell. 2015; 162:516–526. [PubMed: 26232222]

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]

Lin Y, Dion V, Wilson JH. Transcription promotes contraction of CAG repeat tracts in human cells. Nature structural & molecular biology. 2006; 13:179–180.

Lin Y, Wilson JH. Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. Molecular and cellular biology. 2007; 27:6209–6217. [PubMed: 17591697]

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]

Liu Y, Wilson SH. DNA base excision repair: a mechanism of trinucleotide repeat expansion. Trends in biochemical sciences. 2012; 37:162–172. [PubMed: 22285516]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Lokanga RA, Senejani AG, Sweasy JB, Usdin K. Heterozygosity for a hypomorphic Polbeta mutation reduces the expansion frequency in a mouse model of the Fragile X-related disorders. PLoS genetics. 2015; 11:e1005181. [PubMed: 25886163]

Loomis EW, Sanz LA, Chedin F, Hagerman PJ. Transcription-Associated R-Loop Formation across the Human FMR1 CGG-Repeat Region. PLoS genetics. 2014; 10:e1004294. [PubMed: 24743386]

Maizels N, Gray LT. The G4 genome. PLoS genetics. 2013; 9:e1003468. [PubMed: 23637633]

Manhart CM, Ni X, White MA, Ortega J, Surtees JA, Alani E. The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans. PLoS biology. 2017; 15:e2001164. [PubMed: 28453523]

McGinty RJ, Puleo F, Aksenova AY, Hisey JA, Shishkin AA, Pearson EL, Wang ET, Housman DE, Moore C, Mirkin SM. A Defective mRNA Cleavage and Polyadenylation Complex Facilitates Expansions of Transcribed (GAA)n Repeats Associated with Friedreich's Ataxia. Cell reports. 2017a; 20:2490–2500. [PubMed: 28877480]

McGinty RJ, Rubinstein RG, Neil AJ, Dominska M, Kiktev D, Petes TD, Mirkin SM. Nanopore sequencing of complex genomic rearrangements in yeast reveals mechanisms of repeat-mediated double-strand break repair. Genome research. 2017b

McMurray CT. Mechanisms of trinucleotide repeat instability during human development. Nature reviews Genetics. 2010; 11:786–799.

Mollersen L, Rowe AD, Illuzzi JL, Hildrestrand GA, Gerhold KJ, Tveteras L, Bjolgerud A, Wilson DM 3rd, Bjoras M, Klungland A. Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. Hum Mol Genet. 2012; 21:4939–4947. [PubMed: 22914735]

Morales F, Vasquez M, Santamaria C, Cuenca P, Corrales E, Monckton DG. A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. DNA repair. 2016; 40:57–66. [PubMed: 26994442]

Oestergaard VH, Lisby M. Transcription-replication conflicts at chromosomal fragile sites-consequences in M phase and beyond. Chromosoma. 2017; 126:213–222. [PubMed: 27796495]

Pinto RM, Dragileva E, Kirby A, Lloret A, Lopez E, St Claire J, Panigrahi GB, Hou C, Holloway K, Gillis T, et al. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS genetics. 2013; 9:e1003930. [PubMed: 24204323]

Polleys EJ, House NCM, Freudenreich CH. Role of recombination and replication fork restart in repeat instability. DNA repair. 2017

Polyzos A, Holt A, Brown C, Cosme C, Wipf P, Gomez-Marin A, Castro MR, Ayala-Pena S, McMurray CT. Mitochondrial targeting of XJB-5-131 attenuates or improves pathophysiology in HdhQ150 animals with well-developed disease phenotypes. Hum Mol Genet. 2016; 25:1792–1802. [PubMed: 26908614]

Polyzos AA, McMurray CT. Close encounters: Moving along bumps, breaks, and bubbles on expanded trinucleotide tracts. DNA repair. 2017; 56:144–155. [PubMed: 28690053]

Qiao Q, Wang L, Meng FL, Hwang JK, Alt FW, Wu H. AID Recognizes Structured DNA for Class Switch Recombination. Mol Cell. 2017; 67:361–373. e364. [PubMed: 28757211]

Ranjha L, Anand R, Cejka P. The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. J Biol Chem. 2014; 289:5674–5686. [PubMed: 24443562]

Ratmeyer L, Vinayak R, Zhong YY, Zon G, Wilson WD. Sequence specific thermodynamic and structural properties for DNA.RNA duplexes. Biochemistry. 1994; 33:5298–5304. [PubMed: 7513557]

Reddy K, Schmidt MH, Geist JM, Thakkar NP, Panigrahi GB, Wang YH, Pearson CE. Processing of double-R-loops in (CAG).(CTG) and C9orf72 (GGGGCC).(GGCCCC) repeats causes instability. Nucleic acids research. 2014; 42:10473–10487. [PubMed: 25147206]

Reddy K, Tam M, Bowater RP, Barber M, Tomlinson M, Nichol Edamura K, Wang YH, Pearson CE. Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. Nucleic acids research. 2011; 39:1749–1762. [PubMed: 21051337]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Rindler PM, Bidichandani SI. Role of transcript and interplay between transcription and replication in triplet-repeat instability in mammalian cells. Nucleic acids research. 2011; 39:526–535. [PubMed: 20843782]

Roberts RW, Crothers DM. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. Science. 1992; 258:1463–1466. [PubMed: 1279808]

Rogacheva MV, Manhart CM, Chen C, Guarne A, Surtees J, Alani E. Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease. J Biol Chem. 2014; 289:5664–5673. [PubMed: 24403070]

Romanova NV, Crouse GF. Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast. PLoS genetics. 2013; 9:e1003920. [PubMed: 24204320]

Salvi JS, Mekhail K. R-loops highlight the nucleus in ALS. Nucleus. 2015; 6:23–29. [PubMed: 25587791]

Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nature reviews Genetics. 2015; 16:583–597.

Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, Ginno PA, Xu X, Chedin F. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol Cell. 2016; 63:167–178. [PubMed: 27373332]

Schmidt MH, Pearson CE. Disease-associated repeat instability and mismatch repair. DNA repair. 2016; 38:117–126. [PubMed: 26774442]

Schrader CE, Guikema JE, Wu X, Stavnezer J. The roles of APE1, APE2, DNA polymerase beta and mismatch repair in creating S region DNA breaks during antibody class switch. Philos Trans R Soc Lond B Biol Sci. 2009; 364:645–652. [PubMed: 19010771]

Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes & development. 2014; 28:1384–1396. [PubMed: 24990962]

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]

Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. Annu Rev Immunol. 2008; 26:261–292. [PubMed: 18370922]

Su XA, Freudenreich CH. Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2017; 114:E8392–E8401. [PubMed: 28923949]

Usdin K, House NC, Freudenreich CH. Repeat instability during DNA repair: Insights from model systems. Critical reviews in biochemistry and molecular biology. :1–26.

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 & development. 2016; 30:1327–1338. [PubMed: 27298336]

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]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

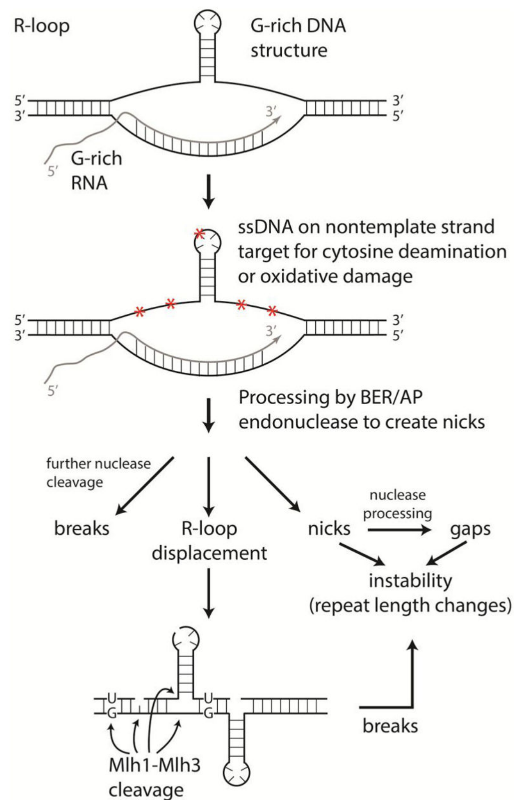

Figure 1. Events at R-loops within G-rich repeats

Stable R-loops can form when a purine-rich RNA stays hybridized to the DNA template, facilitating DNA structure formation on the exposed single-stranded non-template DNA strand (ssDNA). The ssDNA is exposed to cytosine deamination and oxidative damage, which can be processed by base excision repair (BER) and AP endonuclease to create strand-specific nicks. Multiple downstream events can ensue. Right pathway: the nicks can be repaired directly by BER or further processed to gaps by nucleases (e.g. Exo1, Rad1/XPF, Rad2/XPG) and then repaired, with the opportunity for repeat length changes to occur. A pre-formed hairpin near the nick may favor repeat expansions. Left pathway: alternatively, further processing by nucleases (e.g. XPF, XPG or others) could lead to double-strand breaks (DSBs). Middle: R-loop displacement is predicted to create a slipped-strand structure with hairpins on both strands, which may also contain U-G mismatches from cytosine deamination and/or nicks from subsequent processing by Ung1 and other BER enzymes. This is a predicted substrate for cleavage by the Mlh1-Mlh3 (MutLγ) nuclease; potential sites of recognition and cleavage are indicated (e.g. U-G mismatches, nicks, or junctions within structured DNA; cleavage of the R-loop or its nicked product are also possible (not shown)). If the resulting breaks are repaired with mis-alignments, repeat length changes will occur.