Published in final edited form as:

DNA Repair (Amst). 2017 August ; 56: 144–155. doi:10.1016/j.dnarep.2017.06.017.

Close encounters: Moving along bumps, breaks, and bubbles on expanded trinucleotide tracts

Aris Polyzos and Cynthia McMurray*

MBIB Division, Lawrence Berkeley Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, United States

Abstract

Expansion of simple triplet repeats (TNR) underlies more than 30 severe degenerative diseases. There is a good understanding of the major pathways generating an expansion, and the associated polymerases that operate during gap filling synthesis at these “difficult to copy” sequences. However, the mechanism by which a TNR is repaired depends on the type of lesion, the structural features imposed by the lesion, the assembled replication/ repair complex, and the polymerase that encounters it. The relationships among these parameters are exceptionally complex and how they direct pathway choice is poorly understood. In this review, we consider the properties of polymerases, and how encounters with GC-rich or abnormal structures might influence polymerase choice and the success of replication and repair. Insights over the last three years have highlighted new mechanisms that provide interesting choices to consider in protecting genome stability.

Keywords

DNA repair; Trinucleotide repeats; Trinucleotide expansions; DNA structure; DNA; RNA; R-loops; Polymerase; Translesion; Polymerases; Replicative polymerases; Base excision repair; Triple helix; Quadruplex; Hairpins

1. Introduction

Alterations in simple repetitive sequences lie at the center of DNA evolution and sequence diversity, which drives adaptation [1–6]. However, rapid changes in repetitive sequences can also result in deleterious effects on gene expression and function, leading to disease. Simple trinucleotide repeats (TNR) have taken on special significance in this regard since their expansion underlies more than 30 human neurodegenerative and neuromuscular diseases [7–10].

The structure-forming potential of TNRs lies at the heart of the instability (Fig. 1A–C) [7–10], and the structural details of their features were established early on [11,12].

Trinucleotide repeats are quasi-palindromes, and form hydrogen bonded, hairpins harboring

*Corresponding author. ctmcmurray@lbl.gov (C. McMurray).

Conflict of interest

The authors have no conflicts of interest.

A/A, T/T, or C/C mismatches (Fig. 1A, N = A,T,G) within their stem [11–13], or, depending on the sequence, triplex [14–17] (Fig. 1B) and quadruplex structures [18,19] (Fig. 1C), among others [20–22] (Fig. 1B,C). Indeed, some TNRs can form quadruplexes comprising both DNA and RNA [23,24]. Regardless of the type of structure, it must be removed to avoid mutation [25], and if not, the gene grows by roughly the size of the extrahelical DNA loop that forms during a biological transaction (e.g., Fig. 1C). Indeed, in yeast, triplet repeats capable of forming secondary structure escape repair [26] and increase the frequency of expansion by 5–1000-fold [27–29]. Equivalent small unstructured loops are rapidly removed and do not expand [26].

The location of the expanded TNRs can occur anywhere in or around a gene (Fig. 1D), but whether it occurs in coding or non-coding regions determines the size of the expansion [7–10]. One can easily predict that genetic selection would rid organisms of a cells expressing a toxic, expanded protein, while non-coding expansions are better tolerated [9]. However, there is an unexplained linkage between the TNR gene position of disease-length allele and the parent who transmits the expansion. Short expansions in coding alleles tend to be transmitted through the male germline, while long expansions in non-coding alleles tend to be transmitted through the female germline [9]. In HD, for example, it was noted early in patients develop disease above a premutation threshold of roughly 35 CAGs (Fig. 1E) [30–33]. Transmitted CAG expansions rarely grow beyond 100 CAG repeats in a living patient [9,34], and intergenerational “jumps” are most often between 3 and 15 TNRs [30,31,34,35]. If transmitted maternally, the size of the CAG expansion changes little or even shortens [36]. The same patterns are observed in HD primate models [37] and in mouse models [38–40] of HD, although the threshold for expansion occurs at longer tract lengths.

In non-coding TNR alleles in Fragile X syndrome (FXS) [41–43], myotonic dystrophy 1 gene (DM1) [44,45], and Friedreich's ataxia (FRDA) [46,47] patients (Fig. 1C), the expansions are larger. The premutations range for CGG, CTG and GAA repeats, respectively, is around 55–200, and TNR expansion to the full mutation occurs in the oocytes with gains of hundreds of repeats and thousands in congenital cases [48–50]. In the non-coding TNRs, contractions typically occur during male transmission [51,52], indicating that the sperm is far more restrictive in the TNR lengths it will tolerate [53]. For example, the full mutation allele in FXS is deleted back to its premutation length in the sons of premutation mothers around fetal day 13, in their dividing spermatogonia [54,55]. Sons of premutation mothers cannot transmit a disease-length allele, yet it is maintained in their somatic cells and causes disease [54,55]. In oocytes from FXS mothers, the large expanded allele continues to expand with age even though cell division has ceased.

TNRs not only expand during transmission in the germ cell [9,30,32], TNRs continue to expand in somatic cells throughout life [56,57]. In HD, for example, the somatic expansion exceeds the size of the inherited lengths, and rare alleles of up to 1000 have been detected in animals [57] (Fig. 1F). The somatic changes in the long TNRs are observed in most non-coding genes, but the size of the somatic lengths are small by comparison to the large jumps that occur in early germ cell development. Importantly, both the germ cells and the somatic expansions contribute to toxicity [58–61]. In this regard, the somatic expansion has raised the possibility that shortening the repeats during life may be therapeutic [62,63]. Indeed, use

of genetic and pharmacological methods to reduce the TNR length in somatic cells has improved pathophysiology in HD [34,35,64], FXS [65], GAA [66,67], and DM1 [68].

1.1. The polymerase problem

Not surprisingly, these tantalizing features of expansion continue to spark interest into this fascinating mutation, and raise issues as to whether there are multiple underlying mechanisms. Secondary structure within the TNRs is at the root of the problem; it can be constitutive or arise actively during polymerase passage, but in both cases create a need to process and remove the impediment. Self-folding of a structure-capable DNA is stable enough to either block polymerase passage, or to peel away from the duplex just long enough that the advancing polymerase prevents reannealing of the extrahelical strand [69]. In either case, the structure is trapped, and the cell faces a decision to unwind, to remove, to move past, or to incorporate the looped DNA. Genome stability or instability ensues depending on the choice. By this simple definition, polymerase fidelity and processivity over these tracts are likely to be as important as the stability of the structure in determining the mechanisms and the extent of mutation.

Surprisingly little is known about the role of polymerases in completing error-prone synthesis, or what features of a lesion might determine or exclude the use of a particular polymerase at TNR tracts. Based on structure, mammalian polymerases fall into three of five classes (Table 1), which play pivotal roles in DNA replication (pols γ , ϕ , ν , α , δ , ϵ) (family A,B), base excision repair (pol β) (family X), replication and repair in mitochondrial DNA (pol γ) (family A), non-homologous end-joining and immunological diversity (pols λ , μ , and terminaldeoxynucleotidyl transferase) (family X), and DNA damage tolerance including translesion synthesis (η , κ , ζ , Rev1) (family Y) (Table 1). Some of their known roles are summarized (Table 2), and each polymerase is described in detail in excellent reviews [70–72].

Over the past 3 years, however, there have been surprising new insights into the mechanisms that build on our understanding of expansion and highlight the importance of polymerases. Here, we discuss polymerase encounters on abnormal structures or “difficult to copy” TNRs. In the following sections, we will focus on polymerase encounters at three representative DNA barriers, which we refer to as bumps (small chemical lesions), bubbles (RNA loops, and secondary structures), and breaks. Each of these present distinct perturbation contexts of the DNA helix and the outcome of expansion relies on distinct genomic and polymerase contexts. Three questions will be considered: (1) is expansion a process of replication or repair, i.e., would expansion arise from use of a repair or replicative polymerase, (2) do disease-length repeats alter the ways in which they are viewed by a polymerase when damage is encountered, (3) how might secondary structures alter the polymerase that is chosen.

1.2. To replicate or to repair, which is the question

So, does expansion occur during replication or repair, and how do we distinguish among the polymerases that might be used? The hereditary nature of the expansion diseases prompted early focus on the germ cell, and early replication-dependent expansion in spermatogonia

was a popular hypothesis. However, over the last 20 years, the consensus viewpoint has shifted to DNA repair as the major expansion mechanism in both dividing and non-dividing cells [7–10]. Expansion occurs in neurons and quiescent oocytes [38,73,74]. Neither cell type divides, but widespread age-dependent expansion occurs in both [49–52]. No new oocytes are created after birth, indicating that age-dependent expansion in adult oocytes cannot arise from a new round of proliferation [75]. Proliferating patient fibroblasts or lymphoblasts harboring disease-length alleles do not typically expand in culture. However, in one cell line isolated from a DM1 transgenic mouse, the rate of CTG tracts is unrelated to the rate of replication [76]. On long CGG tracts, replicative polymerases stall [77], and frequently break in cells [78,79]. As modeled in yeast and in mammalian cells, expansions during proliferation arise from rescue of stalled forks by recombination [80], break-induced replication (BIR) [81], and strand switching [82,83]. Thus, unlike early models, the consensus viewpoint is that expansion is a process of repair even in proliferating cells.

If expansion occurs during repair, what polymerases are used? All DNA replication and repair polymerases carry out gap-filling synthesis in three basic reactions, and pair the template with incoming nucleotide using a common structural framework, depicted as a finger, palm, and thumb structure (Fig. 2) [84]. Moreover, DNA secondary structures apparently influence polymerase function [85–87]. High-fidelity polymerases have precise active site geometry and closely monitor the DNA minor groove during translocation of the nascent base pair, which can trigger the 3′–5′ exonuclease proofreading activity if a mismatch or abnormal base pair is sensed (Fig. 2A,B) [70–72,86,88,89]. However, their rigid structural constraints require correctly matched base pairs, and unusual DNA structures are not accepted in the active site [70–72,86,88–93]. A DNA translesion polymerase (TLP) will typically lack exonuclease activity and has evolved larger active sites to accommodate the damage template, as well as the incoming nucleotide in the optimal geometry [70–72,86,88,89,91–93]. However, the impact of encountering a TNR secondary structure for these error-prone polymerases is unknown in most cases [86].

2. Encountering bumps within triplet repeat tracts

The cell's main strategy to deal with an oxidized base is to remove it and fill the gap with the correct base using base excision repair (BER), and to a lesser extent NER/TCR [7–10,93]. In BER, the oxidized bases are recognized and removed by a DNA glycosylase, followed by Apolendonuclease 1 (APE1) processing to generate a 3′OH suitable for polymerase extension [7–10]. Gap-filling synthesis is performed by polymerase β (pol β) [90,93] (or back-up by other X family members [84,85,90]) for short or intermediate patch repair. For longer gaps, proliferating cell nuclear antigen (PCNA) and the flap endonuclease 1 (FEN-1) are required, and may be catalyzed by pol β or by pol δ or pol ϵ [84,90,93]. Indeed, TNRs expansion is reduced in transgenic mice lacking 8-oxo-guanine glycosylase (OGG1) [94], NEIL1 [95] pol β with normal activity [96] and XPA [97], strongly supporting the notion that excision repair machinery mediates the instability during removal of oxidative damage *in vivo* (Fig. 3A,B).

However, new mechanistic insight into polymerase encounters with “bumps” in TNRs came from an unlikely source. Recent, emerging evidence reveals that BER and MMR cooperate

in a hybrid pathway to correct oxidative lesion in TNR tracts and expand them [98,99] (Fig. 3B). Specifically, MSH2-MSH3 physically interacts with DNA and stimulates pol β -catalyzed synthesis on CAG and GAA repeats to complete long patch repair and strand displacement gap filling synthesis [98,99] (Fig. 3A). Although it has well-described roles in postreplicative repair, unexpectedly, MSH2-MSH3 adopts an entirely new function as a stimulatory factor when it operates with a low processivity enzyme, pol β , in the context of BER [98,99]. Indeed, both pol β and MSH2-MSH3 are critical to the expansion mechanisms. In mice, a heterozygous Y265C mutation in pol β is sufficient to significantly reduce both the number of expansions in paternal gametes, and the extent of somatic expansion in some tissues of the FXS mouse, which harbor CGG repeats [96]. In mice and/or cells, loss of either MSH2 or MSH3 (and some polymorphisms) [100] also abolishes expansion within exon 1 of the human HD allele [100–104], in the DM1 allele [105–109], in the CGG Fragile X syndrome (FXS) [110,111]. Expansion and fragility occur in mammalian cells lines and in yeast expressing the GAA in the Friedreich's ataxia (FRDA) gene [112–115], but not in mice for reasons that are poorly understood [113]. Thus, in general, BER and MMR genetically interact and appear to provide means to the same end.

Perhaps, as a low processivity enzyme (X family), it is not too surprising that pol β has a difficult time copying TNRs. However, pol β nearly stops in its tracks as it encounters the TNR tract, and on its own, is barely able to traverse them [98,99]. What pol β cannot copy, however, it skips over, leaving behind large uncopied TNR loops, which are deleted (Fig. 3B). Association with MSH2-MSH3 facilitates loading of pol β to the oxidized base [98,99], reduces deletions that normally would occur by BER alone [99,116], and stimulates pol β to complete strand displacement synthesis on CAG repeat tracts *in vitro*. Cooperation of the two enzymes has a direct impact on moving pol β through the repeat tract.

This could be a good outcome since repair is complete, but by a series of unfortunate events, successful repair of the hybrid pathway is also its downfall. MSH2-MSH3-pol β complex drives the completion of long patch BER at the cost of strand displacement and hairpin formation [98,99]. FEN1 cleavage requires a specific geometry between the 3' end of the synthesized strand [117]. However, TNR hairpins hide the 5' end of the template, which are refractory to cleavage by FEN-1 [29,118]. Since these strands can be ligated successfully *in vitro*, presumably, the MSH2-MSH3-pol β protects the flap and must rearrange it to produce a correctly processed end for ligation (Fig. 3B) [118]. Indeed, the ability to bind to the TNR loops provide MSH2-MSH3 the opportunity to re-equilibrate the displaced flap to allow partial reannealing and production of a shorter unblocked 5' end that is suitable for FEN1 processing (Fig. 3B). These steps are consistent with known features of MMR and BER, and undoubtedly occur through an intermediate in which MSH2-MSH3 binds to and alters its hairpin junction [119,120]. Indeed, ligase requires FEN1 activity and coordination of APE1, pol β and FEN1 to modulate CAG repeat expansion [118,121,122].

How these loops are finally incorporated into DNA to complete expansion is poorly understood. However, an endonuclease clip “in trans” on the strand opposite the TNR hairpin would extrude the loop, and gap-filling synthesis would complete TNR incorporation into duplex DNA. Indeed, such a mechanism has been reported. At CTG hairpins, MSH2-MSH3 recruitment of endonucleolytic MLH1-PMS2 results in incision both “in cis” and “in

trans” in small loops, and “in trans” nicking of TNR loops leads to insertion products (Fig. 3C) [123]. In mice, loss of Pms2 [124,125], and MLH3 [126], and MLH-related [127] also suppress expansion in mice, bolstering the notion that Mut-like proteins can contribute to expansion. MLH1-PMS2 “in trans” incision becomes inefficient when TNR loop sizes are larger than 4 repeats [123], however, indicating that other processes or other endonucleases would need to be recruited if larger loops form from a displaced strand. MSH2-MSH3 as a processivity factor is an appealing mechanism since it promotes completion of synthesis on difficult sequence after removal of oxidized bases, but also conveniently positions MSH2-MSH3 to do its job in loop recognition as well (Fig. 3A).

2.1. Structural models for MSH2-MSH3-pol β complex

The structure of MSH2-MSH3 [119] and pol β [128–132] independently have been solved, but the interaction of the pol β -MSH2-MSH3 complex remains unknown. Thus, how MSH2-MSH3 stimulates pol β synthesis can only be speculated (Fig. 2B). Pol β fills gaps of 2–6 nucleotides processively by grabbing simultaneously both the 5′ and 3′ ends of the gap [128–131]. The polymerase domain is positioned at duplex template-primer DNA while the single strand and 5′-terminal phosphate is bound near an 8 kDa domain containing the AP lyase active site [130,131]. The DNA template is kinked 90° precisely at the 5′-phosphodiester linkage of the template base in the polymerase active site. This sharp bend exposes the base pair on the downstream side of the active site, allowing the template-strand base to stack with a histidine residue in the 8-kDa domain rather than with its neighboring base in the active site [128–131,133]. Thus, keeping the downstream base out of the way is clearly important. Indeed, pol λ a closely related X-family back-up for pol β , performs a similar task by “scrunching” the additional uncopied template base in an extrahelical position within a 3 aa. binding pocket (Fig. 2C) [134]. Mutations in the pocket impair pol β processivity indicating that preventing the neighboring base interference is intimately connected to how long a polymerase can hang onto the duplex DNA.

Since pol β lacks both the tight fitting active site and 3′–5′ exonuclease activity associated with replicative polymerases, there seems no obvious need to adopt a new geometry to accommodate the gapped TNR template and incoming nucleotide. Moreover, the lyase activity of pol β is not needed for strand displacement synthesis during long patch repair. Thus, MSH2-MSH3 may interact with the lyase domain to increase its affinity for incoming template and 5′-phosphate at larger gaps. Interestingly, pol β *in vivo* has a high rate of deletion errors (24×10^{-4}) on mononucleotide templates [134], indicating a tendency for strand slippage during the dissociation-reassociation phase of a polymerization reaction [134]. Since 2–6 nucleotides are filled processively by pol β , it will bind tightly to 1–2 triplet repeats before falling off, perhaps accounting for the fact that deletion of TNR is notably in units of three nucleotides. Given the large gap size, it is conceivable that the MSH2-MSH3 may prevent internal reannealing of the single strand template to ensure that robust pairing of the uncopied template with the incoming nucleotide occurs “in single file” as it enters the active site. Indeed, MSH2-MSH3 protects the DNA from S1 cleavage around the abasic site [99], consistent with an end binding activity that might retain pol β tightly associated with the 5′ phosphate, facilitating gap repair.

In all of the above scenarios, expansion size is much smaller than the large leaps observed in FXS or DM1, which expands in oocytes. Thus, other mechanisms may promote them, or larger expansion may require accumulation of smaller expansions in non-dividing cells. While oxidation of any particular base in a TNR is infrequent, the human lifetime is decades, creating opportunity for a “toxic oxidation cycle” in which continuous damage-removal-repair cycles lead to increasing expansions with age (Fig. 3D) [94]. In this type of model, each new round of damage-removal-repair generates a bigger expansion and a larger target (Fig. 2C). MMR during repair by Family B polymerases improves the fidelity of replicative DNA synthesis by 100–1000-fold, and lowers the overall mutation rate to one error per 10^{10} nucleotides synthesized [135]. Loss of MMR activity is a hallmark of genome-wide genetic instability referred to as a mutator phenotype. The inability for MSH2-MSH3 to stimulate low processivity polymerases at difficult sequences would also contribute to the mutation rate by increasing the error-prone nature of TLPs.

2.2. Pathway choice?

Is that just the tip of the iceberg? MSH2-MSH3 plays roles as a co-factor in BER [94,98,99], transcription-coupled repair (TCR) [136–138] and double strand break repair (DSBR) [137,139,140]. Each of these pathways has been implicated in CAG instability in mice [58,77–79, 81–83, 94], and classically operates with distinct Family A, B, X or Y polymerases, depending on the pathway (Table 2). Does MSH2-MSH3 stimulate them all?

It is unlikely to be the case for A or B family replicative polymerases. Despite its role in post-replicative repair, there is no direct evidence that MSH2-MSH3 associates with highly processive pol δ or pol ϵ , the two most common replication enzymes, both of which operate with their own PCNA processivity factors. There is a paucity of data as to whether MMR functionally alters the synthetic activity of other X or Y family polymerases at TNRs (Table 2). However, a smattering of reports over the last 15 years finds that physical interactions between MMR recognition complexes and TLPs not only occur but also have functional consequences in other systems. For example, MSH2-MSH6 physically interacts with DNA polymerase (pol η), and stimulates the catalytic activity of pol η at bases located downstream of the initial dU lesion in somatic hypermutability [141].

Another example is pol κ . Each of the Y-family polymerases exhibits a preference for the replicative bypass of specific types of base damage in DNA (Table 2). Pol κ supports accurate bypass of benzo[*a*]pyrene diol epoxide guanine adducts (B[a]P-dG) and ultraviolet (UV) radiation- induced cis-syn thymine-thymine cyclobutane-pyrimidine dimers (CPDs), respectively [142]. Both pol κ and REV1 associate physically with the mismatch repair (MMR) protein MSH2 [143]. Depletion of MSH2 impairs PCNA monoubiquitination, reduces pol κ and REV1 focus formation, and reduces the bypass of CPD lesions after exposure of cells to UV radiation. Thus, MSH2 appears to cooperate with at least some of the family Y polymerases in post-UV cellular responses. From the perspective of TNR tracts, it is interesting that functional interactions of X and Y family members occur, minimally, with MSH2-MSH6, which has little impact on expansion [7–10]. However, MSH2-MSH6 may act as a masked “competitor” or impose additional selectivity for polymerase stimulation of TLPs or broaden this class of regulation under some conditions.

Considering the number of possible interactions, however, it is remarkable that the MSH2-MSH3-pol β complex is constitutive in the cells both on and off TNR DNA. The complex is recruited together at the damage site on TNR tracts in cells upon oxidative damage [98,99]. Under these conditions, BER appears to be “primed” by forming the MSH2-MSH3-pol β complex. This would not be an unreasonable design considering the fact that oxidation is the most common lesion in DNA. However, the scenario poses provocative challenges. MSH2, by many rough “guestimates”, is present between 1000 to at 100,000 molecules in cells [144], which is distributed between the MSH3 and MSH6 heterodimers. However, if the MSH2-MSH3-pol β complex is constitutive, the 50,000 oxidized lesions per day [145] would appear to compete heavily for the attention of MSH2-MSH3, given its putative roles in BER, TCR, or DSBRR and a host of TLPs.

Could MSH2-MSH3 guide the choice of repair among redundant pathways by competition and selective polymerase stimulation? This type of competition is used consistently in regulating the PCNA clamp loaders for the replicative polymerases. For example, PCNA interacts with DNA polymerases δ and ϵ , as well as other prominent PCNA-interacting proteins including flap endonuclease 1 (FEN1), DNA ligase I, Topo I, WRN helicase, Uracil-DNA glycosylase, cyclin, P21, as well as pol η and pol κ [146]. However, the p66 subunit of pol δ and FEN1 bind to the same site on PCNA as the p21 peptide [146–148]. Given the large number of proteins interacting with PCNA, the fine-tuned affinity provides a way of coordinating the sequential binding events on a common platform.

Being a clamp itself, MSH2-MSH3 might operate similarly. For example, both BER and TCR can remove oxidative damage in TNR DNA, but with opposite effects. CSB *in vivo* protects CAG repeats and CGG repeats from expansion [149,150], and an oxidized form of XRCC1 is also protective against extreme oxidative stress. On the hand, OGG1 tends to expand them [94]. A mechanism for CSB/OGG1 antagonism is unclear. Although CSB is part of the TCR machinery, it also interacts with BER machinery such as APE1 [151] and poly (ADP-ribose) polymerase-1 [152]. Gap filling synthesis by NER/TCR is thought to occur using replicative polymerase machinery. However, it is possible that CSB may co-opt MSH2-MSH3, and prevent its ability to stimulate BER, thereby favoring reduction in expansion via TCR (or vice versa). Such crosstalk may explain the globally protective effects of CSB on expansion, and the observations that treatment with oxidizing agents can induce expansions and contractions [153]. There has been some suggestion that MSH2-MSH3 has differential roles in the excision repair pathways. Knock-down of TC-NER components increases convergent transcription- induced cell death, while depletion of MSH2 decreases cell death [154,155].

Crosstalk solutions may involve other pathways. For example, XRCC1 is involved in the efficient repair of DNA single-strand breaks formed by exposure to ionizing radiation and alkylating agents [156]. The XRCC1 protein does not have enzymatic activity, but acts as a scaffolding protein that interacts with multiple repair enzymes including pol β and poly (ADP-ribose) polymerase. XRCC1's scaffolding function is thought to aid SSBR, BER, and TCR, particularly in extreme cases of oxidative damage [156]. How MMR might cooperate in such a system is only speculative, but implementation of crosstalk pathways provides an

interesting strategy to share machinery and orchestrate pathways choice by optimal use of whatever machinery is available or that best fits the situation.

3. Encountering bubbles and breaks within TNRs

But what happens if a polymerase encounters a stable TNR secondary structure or an R-loop? We refer to both of these structures as DNA “bubbles” since the two anti-parallel DNA strands remain open and cannot reanneal to reform duplex DNA. They differ, however, in that a secondary structure blocks polymerase by intra-strand pairing while the R-loop uses interstrand pairing between DNA.

3.1. Possible roles of R-loops at TNRs

Let's consider R-loops first. R-loops, many of which are transcriptionally- induced, occur naturally during DNA transcription [157,158] and DNA repair [159], and are present at most TNRs, many of which are bidirectionally transcribed [160]. R-loops are involved in mitochondrial DNA replication, class-switch recombination, gene expression, and DNA repair [157,158], and must be preserved for normal function. At least three RNA-DNA helicases, including senataxin, are capable of unwinding R-loops [161], and RNase H removes the RNA in the RNA-DNA hybrid, which is re-established during the next round of transcription. Indeed, senataxin is of particular interest to TNRs since deficits in its unwinding activity leads to microsatellite instability in Oculomotor Apraxia (AOA) [162,163]. As TNRs expand, these regions become more difficult to maneuver, and leave behind stable R-loops that cannot be removed.

Indeed, most analysis of disease-length TNRs support a detrimental effect of R-loop formation, which has been particularly well studied in long non-coding CGG and GAA TNRs in the FXS [164,165] and Friedreich's ataxia (FRDA) alleles [166]. R-loops formed at disease-length CGG and GAA repeats appear to cause transcriptional silencing, a mechanism for toxicity, as well as TNR instability [66,164–168]. For example, in a mouse model, expression and instability in the CAG tract of Spinocerebellar Ataxia 7 gene (*SCA7*) gene depends on expression of an adjacent alternative promoter and a convergently transcribed antisense noncoding RNA, SCAANT1 (SCA antisense 1) [169]. SCA7 is a subunit of a transcriptional coactivator complex (called STAGA or TFTC) that has histone acetyltransferase activity [169]. Loss of SCAANT1 depressed ataxin-7 sense transcription in a *cis*-dependent fashion and is accompanied by chromatin remodeling and a rise in TNR instability, supporting the idea that instability is an anti-sense, RNA-mediated process [169]. Thus, a bi-directional RNA synthesis and convergent bubbles appear to mediate the instability in this gene locus.

Various *in vitro* tests for convergent transcription support a cause-effect relationship between anti-sense R-loops and instability. Indeed, in the selectable cell system for contractions [170,171], DIT7 cells carry a single integrated copy of an *HPRT* minigene that carries a CAG₉₅ tract in its intron and is flanked by the promoters pTRE-CMV^{mini} and pNERB-X1, which drive inducible sense and antisense transcription, respectively. Co-induction of each gene increases the rate of CGG, GAA, CAG, CCTG contraction and increases toxicity relative to one active promoter alone [170]. Using an extract system, R-loops and double-R-

loops also form at expanded (CAG)₇₉·(CTG)₇₉ repeat tract and in (GGGGCC)_n·(GGCCCC)_n repeat tracts from the *C9orf72* gene in the absence of DNA replication, and results in instability [172]. Interestingly, the mRNA and not the antisense is detected at silenced CGG repeats in the FMR-1 gene without obvious instability [164], and in episomal plasmids [167]. Although the relationship between expansion and deletion is still not clear, collectively, these results raise the possibility that gene silencing is caused by R-loops from mRNA, while instability depends on bi-directional transcription, which has also been supported in other systems [173,174].

3.2. Polymerase strategies to by-pass R-loops and large impassable structures

So, how might an R-loop cause unstable DNA? Normally, at an impassable block, the TCR/NER machinery is used to remove the block [175,176]. The CSA and CSB modeling factors recruits the NER machinery including XPA and the XPF/ERCC1 and XPG endonucleases, which would remove the offending strand [175,176]. Loop formation might occur during gap filling synthesis or by fold back of the displaced strand after TCR, ending in an expansion [9]. However, at transcriptionally repressed FMR1 genes, at least in the CGG repeats, RNase H cannot remove the R-loop [164], indicating that it is inaccessible. If TCR is the underlying cause of expansion, however, the RNA-DNA hybrid should be an accessibility structure for TCR bubble removal by XPF and XPG endonucleases [175,176]. Furthermore, loss of XPA (in DM1 mice) suppresses expansion [97], while loss of CSB (FRDA and HD) increases expansion [149,150]. This feature is exceptionally puzzling in that XPA and CSB are members of the same TCR machinery. Loss of XPC has no effects on expansion arguing against the idea that XPA operates within GG-NER [108].

There answer to the conundrum is unknown. However, a single strand break intermediate left behind after successful TCR can be converted into DSB [158,177] by either double incision of the bubble (Fig. 4), or if the TCR product converges with a replication fork before the polymerase can complete gap-filling synthesis (Fig. 4). Indeed, in a convergent transcription model for CAG contraction, siRNA knockdown in mammalian cells reveals a contribution from RFC, ATR, and junction endonucleases components of the Fanconi Anemia pathway [170,171], and FANCD1 facilitates replication through common fragile sites [178], all of which are signals for DSB, or uncoupling of leading and lagging strand replication. If correct, what began as TCR, in these cases, ended up as a DSB, which can be repaired by a different pathway.

There is as yet no general consensus as to how a DSB is repaired at TNRs among humans, mice and yeast, and there are likely to be multiple mechanisms. However, secondary structures often block common biological transactions like transcription [179]. Freudenreich will cover this topic in detail in this special issue. However, two mechanisms deserve some mention. Strand switching is one way that polymerases get past a block. Indeed, in patient fibroblasts, single molecule combing analysis in the FXS allele reveals that the polymerase pauses at CGG repeats and switches to using the lagging-strand as the template (Fig. 5A) [82]. Upon every fork rearrangement, at least part of the CGG repeat impediment has been bypassed on the leading strand until, the polymerase can bootstrap its way through a longer repeat, albeit with instability- causing loop formation along the way. Such a mechanism is

also reported at long GAA repeats in yeast [180,181]. Indeed, yeast systems have been particularly powerful in identifying factors that promote expansion [180–182]. In genome-wide yeast screens, the rates of GAA repeat expansions were 19-fold increased in the Rad27 (FEN-1). Both Tof1 and Mrc1 proteins alleviate the replication fork stalling at DNA hairpins, and were protective against expansion [180,182]. In yeast, a complete set of 4786 deletion mutations and 800 essential genes evaluated by knockdown or induction led to the identification of 33 mutants, which strongly increase GAA/TTC fragility and expansions [181]. These included replisome components and replication-pausing checkpoint surveillance, indicating that problems imposed by the triplex structure during DNA replication are a major triggering event in chromosomal fragility and size in actively dividing cells. Mutants lacking Sgs1 (3'-to-5' DNA helicase homologous to the human Bloom's syndrome DNA helicase), Rad5 (SWI/SNF family), and Rad6 (ubiquitin-conjugating enzyme that regulates strand switching in yeast) promoted expansion. Thus, fork stabilizing precluded expansion while template switching and fork restart promoted repeat expansions in yeast [180–182].

More recently, however, break-induced repair (BIR) has emerged as an attractive mechanism, at least at long CAG repeats (Fig. 5B,C) [183–185]. BIR is distinct from HR and NHEJ (Fig. 5A,C) [81,183,185] in that the latter two mechanisms deal with two broken ends, while DSBs that arise by replication fork collapse or by erosion of uncapped telomeres have only one broken end. In this case, a single end can invade the homologous chromosome and copy to the chromosome end [81,183,185]. In the engineered yeast system harboring large CAG expansions [185], loss of Rad51, Rad52, Mre11, Pif1, and Mus81 and/or Yen1 proteins suppressed TNR expansion, all of which are proteins of BIR and recombination pathways [183,184]. Pif1 prevents replication- fork stalling at G-quadruplexes [186]. The factors that promote expansion are members of the recombination machinery. Rad51 and Rad52 are required for filament formation during HR and BIR. Mus81 and Yen-1 are structure specific endonucleases that, like XPF, play overlapping roles in resolving structures such as D-loops during recombination [187]. YEN-1 is a Holliday junction resolvase, which promotes template switching during BIR [187]. Notably, defects in the replicative DNA polymerases δ and ϵ strongly increased rates for both repeat expansions, but mutants that impair translesion DNA polymerases have no effects, indicating the specificity of the reaction for B family polymerases for the expansions [184]. These results are consistent with previous reports that DNA polymerases η and ζ lack significant effects on triplet repeat instability [188], that Mre11-Rad50-Xrs2 complex promotes trinucleotide repeat expansions independently of homologous recombination [189], and the fact that the catalytic subunit of DNA PK has no impact on expansion in mice [108].

Much more investigation will be required to fully explore BIR and its relationship to expansions. However, the properties of expansion at long CAG repeats in yeast are remarkably reminiscent of those of common fragile sites (CFS) in mammalian cells [190,191]. Like TNRs, CFSs are difficult-to-replicate loci, which often occur at repetitive DNAs, and manifest as gaps or breaks on metaphase chromosomes. Entry into mitotic prophase apparently triggers the recruitment of MUS81 to CFSs, and the structure-specific endonuclease is also used to promote the appearance of chromosome gaps or breaks. Furthermore, the nuclease activity of MUS81 promotes replicative pol δ -dependent DNA

synthesis at the break, leaving behind sites with intrinsically high levels of chromosomal instability, as is observed for BIR.

4. Concluding remarks and future perspectives

Although our understanding of instability is ever increasing, as many questions emerge as are solved. For example, if stable R-loops lead to transcriptional silencing and hypermethylation, why do oocytes harboring full mutations expand in smaller steps with age? DNA repair nuclease factors must be able to access the site to cause instability, yet these sites are generally “nuclease resistant” in heterochromatin. If BIR occurs during replication stalling and causes expansion, how does a break form during a decades-long quiescent state of oocytes. Treatment of FRDA animals with potassium bromide, a powerful oxidant, induces age-dependent expansion in the oocytes of these animals, and could be repaired by BER or TCR. These latter processes are typically associated with smaller expansions and not large jumps. If either of these processes leads to BIR, is the same mechanism responsible for both large-scale and small-scale expansions? It is also important to consider organism biases. *Saccharomyces cerevisiae* lacks homologs of several mammalian short-patch BER proteins, including pol β , DNA ligase III, XRCC1, and the kinase domain of polynucleotide kinase [192]. If a BER pathway were responsible for expansion in non-dividing mammalian cells, it would be unstable in yeast if there were no back-up systems. Instability in oocytes increases with maternal age. In mice, loss of Rad 52 and 54 have little impact for expansion, which is used for BIR, HR and NHEJ. Is there another process that does not require these enzymes? Resolving these issues continue to fuel investigation that will certainly uncover hidden secrets of the genome, and are worth the wait.

Acknowledgments

This work was supported by National Institutes of Health grants NS060115 (to CTM), and CA092584 (to CTM). The authors would like to thank S. Mirkin, L. Symington, K. Cimprich, S. Double, S. Wallace, S. Wilson, and T. Kunkel who generously gave us permission or use images for illustration in this review.

References

1. Demers JE, Jiménez-Gasco MdM. Evolution of nine microsatellite loci in the fungus *Fusarium oxysporum*. *J. Mol. Evol.* 2015; 82:27–37. [PubMed: 26661928]
2. Gáspári Z, Ortutay C, Tóth G. Divergent microsatellite evolution in the human and chimpanzee lineages. *FEBS Lett.* 2007; 581:2523–2526. [PubMed: 17498704]
3. Jurka J, Pethiyagoda C. Simple repetitive DNA sequences from primates: compilation and analysis. *J. Mol. Evol.* 1995; 40:120–126. [PubMed: 7699718]
4. Kovtun IV, McMurray CT. Features of trinucleotide repeat instability in vivo. *Cell Res.* 2008; 18:198–213. [PubMed: 18166978]
5. Plohl M, Meštrović N, Mravinac B. Satellite DNA evolution. *Genome Dynamics*, S. Karger AG. : 2012:126–152.
6. Toth G. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* 2000; 10:967–981. [PubMed: 10899146]
7. Evans-Galea MV, Hannan AJ, Carroddus N, Delatycki MB, Saffery R. Epigenetic modifications in trinucleotide repeat diseases. *Trends Mol. Med.* 2013; 19:655–663. [PubMed: 23953480]

8. Kim JC, Mirkin SM. The balancing act of DNA repeat expansions. *Curr. Opin. Genet. Dev.* 2013; 23:280–288. [PubMed: 23725800]
9. McMurray CT. Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* 2010; 11:786–799. [PubMed: 20953213]
10. Usdin K, House NCM, Freudenreich CH. Repeat instability during DNA repair: insights from model systems. *Crit. Rev. Biochem. Mol. Biol.* 2015; 50:142–167. [PubMed: 25608779]
11. Kiliszek A, Rypniewski W. Structural studies of CNG repeats. *Nucleic Acids Res.* 2014; 42:8189–8199. [PubMed: 24939898]
12. Saini N, Zhang Y, Usdin K, Lobachev KS. When secondary comes first – The importance of non-canonical DNA structures. *Biochimie.* 2013; 95:117–123. [PubMed: 23084930]
13. Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell.* 1995; 81:533–540. [PubMed: 7758107]
14. Frank-Kamenetskii MD, Mirkin SM. Triplex DNA structures. *Annu. Rev. Biochem.* 1995; 64:65–95. [PubMed: 7574496]
15. Gacy AM, Goellner GM, Spiro C, Chen X, Gupta G, Bradbury EM, Dyer RB, Mikesell MJ, Yao JZ, Johnson AJ, Richter A, Melançon SB, McMurray CT. GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol. Cell.* 1998; 1:583–593. [PubMed: 9660942]
16. Jain A, Rajeswari MR, Ahmed F. Formation and thermodynamic stability of intermodular (R•R•Y) DNA triplex In GAA/TTC repeats associated with Friedreich's ataxia. *J. Biomol. Struct. Dyn.* 2002; 19:691–699. [PubMed: 11843630]
17. Mariappan SVS, Catasti P, Silks LA, Bradbury EM, Gupta G. The high-resolution structure of the triplex formed by the GAA/TTC triplet repeat associated with Friedreich's ataxia. *J. Mol. Biol.* 1999; 285:2035–2052. 11 Edited by I. Tinoco. [PubMed: 9925783]
18. Madireddy A, Purushothaman P, Loosbroock CP, Robertson ES, Schildkraut CL, Verma SC. G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Res.* 2016; 44:3675–3694. [PubMed: 26837574]
19. Zamiri B, Mirceta M, Bomsztyk K, Macgregor RB, Pearson CE. Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)8•(GGCCCC)8 repeat: effect of CpG methylation. *Nucleic Acids Res.* 2015; 43:10055–10064. [PubMed: 26432832]
20. Day HA, Pavlou P, Waller ZAE. i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.* 2014; 22:4407–4418. [PubMed: 24957878]
21. Wang YX, Luo C, Zhao D, Beck J, Nassal M. Extensive mutagenesis of the conserved box E motif in duck hepatitis B virus P protein reveals multiple functions in replication and a common structure with the primer grip in HIV-1 reverse transcriptase. *J. Virol.* 2012; 86:6394–6407. [PubMed: 22514339]
22. Zhang Y, Roland C, Sagui C. Structure and dynamics of DNA and RNA double helices obtained from the GGGGCC and CCCC GG hexanucleotide repeats that are the hallmark of C9FTD/ALS diseases. *ACS Chem. Neurosci.* 2016 [Epub ahead of print].
23. Gudanis D, Popena L, Szpotkowski K, Kierzek R, Gdaniec Z. Structural characterization of a dimer of RNA duplexes composed of 8-bromoguanosine modified CGG trinucleotide repeats: a novel architecture of RNA quadruplexes. *Nucleic Acids Res.* 2016; 44:2409–2416. [PubMed: 26743003]
24. Malgowska M, Gudanis D, Kierzek R, Wyszko E, Gabelica V, Gdaniec Z. Distinctive structural motifs of RNA G-quadruplexes composed of AGG, CGG and UGG trinucleotide repeats. *Nucleic Acids Res.* 2014; 42:10196–10207. [PubMed: 25081212]
25. León-Ortiz AM, Svendsen J, Boulton SJ. Metabolism of DNA secondary structures at the eukaryotic replication fork. *DNA Repair.* 2014; 19:152–162. [PubMed: 24815912]
26. Moore H, Greenwell PW, Liu CP, Arnheim N, Petes TD. Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc. Natl. Acad. Sci.* 1999; 96:1504–1509. [PubMed: 9990053]
27. Miret JJ, Pessoa-Brandao L, Lahue RS. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* 1998; 95:12438–12443. [PubMed: 9770504]

28. Pelletier R, Krasilnikova MM, Samadashwily GM, Lahue R, Mirkin SM. Replication and expansion of trinucleotide repeats in yeast. *Mol. Cell. Biol.* 2003; 23:1349–1357. [PubMed: 12556494]
29. Spiro C, Pelletier R, Rolfsmeier ML, Dixon MJ, Lahue RS, Gupta G, Park MS, Chen X, Mariappan SVS, McMurray CT. Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell.* 1999; 4:1079–1085. [PubMed: 10635332]
30. Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, Nance M, Ross CA, Scahill RI, Wetzel R, Wild EJ, Tabrizi SJ. Huntington disease. *Nat. Rev. Dis. Primers.* 2015:15005. [PubMed: 27188817]
31. Giovannone B, Sabbadini G, Di Maio L, Calabrese O, Castaldo I, Frontali M, Novelletto A, Squitieri F. Analysis of (CAG)_n size heterogeneity in somatic and sperm cell DNA from intermediate and expanded Huntington disease gene carriers. *Hum. Mutat.* 1997; 10:458–464. [PubMed: 9401009]
32. Kremer B, Goldberg P, Andrew SE, Theilmann J, Telenius H, Zeisler J, Squitieri F, Lin B, Bassett A, Almqvist E, Bird TD, Hayden MR. A worldwide study of the Huntington's disease mutation: the sensitivity and specificity of measuring CAG repeats. *N. Engl. J. Med.* 1994; 330:1401–1406. [PubMed: 8159192]
33. Trottier Y, Biancalana V, Mandel JL. Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. *J. Med. Genet.* 1994; 31:377–382. [PubMed: 8064815]
34. Andrew SE, Paul Goldberg Y, Kremer B, Telenius H, Theilmann J, Adam S, Starr E, Squitieri F, Lin B, Kalchman MA, Graham RK, Hayden MR. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* 1993; 4:398–403. [PubMed: 8401589]
35. Nørremølle A, Sørensen SA, Fenger K, Hasholt L. Correlation between magnitude of CAG repeat length alterations and length of the paternal repeat in paternally inherited Huntington's disease. *Clin. Genet.* 2008; 47:113–117.
36. Kremer B, Almqvist E, Theilmann J, Spence N, Telenius H, Goldberg YP, Hayden MR. Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. *Am. J. Hum. Genet.* 1995; 57:343–350. [PubMed: 7668260]
37. Putkhao K, Kocerha J, Cho I-K, Yang J, Parnpai R, Chan AWS. Pathogenic cellular phenotypes are germline transmissible in a transgenic primate model of Huntington's disease. *Stem Cells Dev.* 2013; 22:1198–1205. [PubMed: 23190281]
38. Kovtun IV. Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene. *Hum. Mol. Genet.* 2000; 9:2767–2775. [PubMed: 11063736]
39. Kovtun IV. Somatic deletion events occur during early embryonic development and modify the extent of CAG expansion in subsequent generations. *Hum. Mol. Genet.* 2004; 13:3057–3068. [PubMed: 15496421]
40. Mangiarini L, Sathasivam K, Mahal A, Mott R, Seller M, Bates GP. Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat. Genet.* 1997; 15:197–200. [PubMed: 9020849]
41. Fu Y-H, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkert AJMH, Holden JJA, Fenwick RG, Warren ST, Oostra BA, Nelson DL, Caskey CT. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell.* 1991; 67:1047–1058. [PubMed: 1760838]
42. Quartier A, Poquet H, Gilbert-Dussardier B, Rossi M, Casteleyn A-S, Portes Vd, Feger C, Nourisson E, Kuentz P, Redin C, Thevenon J, Mosca-Boidron A-L, Callier P, Muller J, Lesca G, Huet F, Geoffroy V, El Chehadeh S, Jung M, Trojak B, Le Gras S, Lehalle D, Jost B, Maury S, Masurel A, Edery P, Thauvin-Robinet C, Gérard B, Mandel J-L, Faivre L, Piton A. Intragenic FMR1 disease-causing variants: a significant mutational mechanism leading to Fragile-X syndrome. *Eur. J. Hum. Genet.* 2017 [epub ahead of print].
43. Zhao X-N, Usdin K. Ups and downs mechanisms of repeat instability in the fragile X-related disorders. *Genes.* 2016; 7:70.

44. Santoro M, Masciullo M, Silvestri G, Novelli G, Botta A. Myotonic dystrophy type 1: role of CCG, CTC and CGG interruptions within DMPK alleles in the pathogenesis and molecular diagnosis. *Clin. Genet.* 2017 [epub ahead of print].
45. Smith CA, Gutmann L. Myotonic dystrophy type 1 management and therapeutics. *Curr. Treat. Options Neurol.* 2016; 18:52. [PubMed: 27826760]
46. Chiang S, Kovacevic Z, Sahni S, Lane DJR, Merlot AM, Kalinowski DS, Huang MLH, Richardson DR. Frataxin and the molecular mechanism of mitochondrial iron-loading in Friedreich's ataxia. *Clin. Sci.* 2016; 130:853–870. [PubMed: 27129098]
47. Kearney, M., Orrell, RW., Fahey, M., Brassington, R., Pandolfo, M. *Cochrane Database of Systematic Reviews.* Wiley-Blackwell; 2016. Pharmacological treatments for Friedreich ataxia.
48. Dean NL, Tan SL, Ao A. Instability in the transmission of the myotonic dystrophy CTG repeat in human oocytes and preimplantation embryos. *Fertil. Steril.* 2006; 86:98–105. [PubMed: 16716318]
49. Rife M. Analysis of CGG variation through 642 meioses in Fragile X families. *Mol. Hum. Reprod.* 2004; 10:773–776. [PubMed: 15322225]
50. Temmerman ND, Sermon K, Seneca S, De Rycke M, Hilven P, Lissens W, Van Steirteghem A, Liebaers I. Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. *Am. J. Hum. Genet.* 2004; 75:325–329. [PubMed: 15185171]
51. Ashizawa T, Anvret M, Baiget M, Barcelo JM, Brunner H, Cobo AM, Dallapiccola B, Fenwick RG Jr, Grandell U, Harley H, et al. Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. *Am. J. Hum. Genet.* 1994; 54:414–423. [PubMed: 8116611]
52. Lavedan C, Hofmann-Radvanyi H, Shelbourne P, Rabes JP, Duros C, Savoy D, Dehaupas I, Luce S, Johnson K, Junien C. Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am. J. Hum. Genet.* 1993; 52:875–883. [PubMed: 8098180]
53. Jansen G, Willems P, Coerwinkel M, Nillesen W, Smeets H, Vits L, Howeler C, Brunner H, Wieringa B. Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)_n repeat variation and selection against extreme expansion in sperm. *Am. J. Hum. Genet.* 1994; 54:575–585. [PubMed: 8128954]
54. Malter HE, Iber JC, Willemsen R, Graaff Ed, Tarleton JC, Leisti J, Warren ST, Oostra BA. Characterization of the full fragile X syndrome mutation in fetal gametes. *Nat. Genet.* 1997; 15:165–169. [PubMed: 9020841]
55. Reyniers E, Vits L, De Boule K, Roy BV, Velzen DV, de Graaff E, Verkerk AJMH, Jorens HZJ, Darby JK, Oostra B, Willems PJ. The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm. *Nat. Genet.* 1993; 4:143–146. [PubMed: 8348152]
56. Kennedy L. Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Hum. Mol. Genet.* 2000; 9:2539–2544. [PubMed: 11030759]
57. Kennedy L. Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Hum. Mol. Genet.* 2003; 12:3359–3367. [PubMed: 14570710]
58. Alexandrov V, Brunner D, Menalled LB, Kudwa A, Watson-Johnson J, Mazzella M, Russell I, Ruiz MC, Torello J, Sabath E, Sanchez A, Gomez M, Filipov I, Cox K, Kwan M, Ghavami A, Ramboz S, Lager B, Wheeler VC, Aaronson J, Rosinski J, Gusella JF, MacDonald ME, Howland D, Kwak S. Large-scale phenome analysis defines a behavioral signature for Huntington's disease genotype in mice. *Nat. Biotechnol.* 2016; 34:838–844. [PubMed: 27376585]
59. Budworth H, Harris FR, Williams P, Lee DY, Holt A, Pahnke J, Szczesny B, Acevedo-Torres K, Ayala-Peña S, McMurray CT. Suppression of somatic expansion delays the onset of pathophysiology in a mouse model of Huntington's disease. *PLoS Genet.* 2015; 11:e1005267. [PubMed: 26247199]
60. Budworth H, McMurray CT. Problems and solutions for the analysis of somatic CAG repeat expansion and their relationship to Huntington's disease toxicity. *Rare Dis.* 2016; 4:e1131885. [PubMed: 27141411]
61. Swami M, Hendricks AE, Gillis T, Massood T, Mysore J, Myers RH, Wheeler VC. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum. Mol. Genet.* 2009; 18:3039–3047. [PubMed: 19465745]

62. McMahon MA, Cleveland DW. Gene therapy: gene-editing therapy for neurological disease. *Nat. Rev. Neurol.* 2017; 13:7–9. [PubMed: 27982043]
63. Richard GF, Viterbo D, Khanna V, Mosbach V, Castelain L, Dujon B. Highly specific contractions of a single CAG/CTG trinucleotide repeat by TALEN in yeast. *PLoS One.* 2014; 9:e95611. [PubMed: 24748175]
64. Malankhanova TB, Malakhova AA, Medvedev SP, Zakian SM. Modern genome editing technologies in Huntington's disease research. *J. Huntingt. Dis.* 2017
65. Park CY, Halevy T, Lee DR, Sung JJ, Lee JS, Yanuka O, Benvenisty N, Kim DW. Reversion of FMR1 methylation and silencing by editing the triplet repeats in fragile X iPSC-derived neurons. *Cell Rep.* 2015; 13:234–241. [PubMed: 26440889]
66. Li Y, Lu Y, Polak U, Lin K, Shen J, Farmer J, Seyer L, Bhalla AD, Rozwadowska N, Lynch DR, Butler JS, Napierala M. Expanded GAA repeats impede transcription elongation through the FXN gene and induce transcriptional silencing that is restricted to the FXN locus. *Hum. Mol. Genet.* 2015; 6932–6943. [PubMed: 26401053]
67. Ouellet DL, Cherif K, Rousseau J, Tremblay JP. Deletion of the GAA repeats from the human frataxin gene using the CRISPR-Cas9 system in YG8R-derived cells and mouse models of Friedreich ataxia. *Gene Ther.* 2017
68. van Agtmael EL, Andre LM, Willemse M, Cumming SA, van Kessel ID, van den Broek WJ, Gourdon G, Furling D, Mouly V, Monckton DG, Wansink DG, Wieringa B. CRISPR/Cas9-Induced (CTGCAG)_n repeat instability in the myotonic dystrophy type 1 locus: implications for therapeutic genome editing. *Mol. Ther.* 2017; 25:24–43. [PubMed: 28129118]
69. Lee D-Y, McMurray CT. Trinucleotide expansion in disease: why is there a length threshold? *Curr. Opin. Genet. Dev.* 2014; 26:131–140. [PubMed: 25282113]
70. Barnes R, Eckert K. Maintenance of genome integrity: how mammalian cells orchestrate genome duplication by coordinating replicative and specialized DNA polymerases. *Genes.* 2017; 8:19.
71. Lange SS, Takata K-i, Wood RD. DNA polymerases and cancer. *Nat. Rev. Cancer.* 2011; 11:96–110. [PubMed: 21258395]
72. McVey M, Khodaverdian VY, Meyer D, Cerqueira PG, Heyer W-D. Eukaryotic DNA polymerases in homologous recombination. *Annu. Rev. Genet.* 2016; 50:393–421. [PubMed: 27893960]
73. Gonitel R, Moffitt H, Sathasivam K, Woodman B, Detloff PJ, Faull RLM, Bates GP. DNA instability in postmitotic neurons. *Proc. Natl. Acad. Sci.* 2008; 105:3467–3472. [PubMed: 18299573]
74. Kovtun IV, McMurray CT. Trinucleotide expansion in haploid germ cells by gap repair. *Nat. Genet.* 2001; 27:407–411. [PubMed: 11279522]
75. Von Stetina JR, Orr-Weaver TL. Developmental control of oocyte maturation and egg activation in metazoan models. *Cold Spring Harbor Perspect. Biol.* 2011; 3:a005553.
76. Gomes-Pereira M. Mouse tissue culture models of unstable triplet repeats: in vitro selection for larger alleles, mutational expansion bias and tissue specificity, but no association with cell division rates. *Hum. Mol. Genet.* 2001; 10:845–854. [PubMed: 11285250]
77. Krasilnikova MM, Mirkin SM. Replication stalling at Friedreich's ataxia (GAA)_n repeats in vivo. *Mol. Cell. Biol.* 2004; 24:2286–2295. [PubMed: 14993268]
78. Freudenreich CH. Expansion and length-dependent fragility of CTG repeats in yeast. *Science.* 1998; 279:853–856. [PubMed: 9452383]
79. Kumari D, Hayward B, Nakamura AJ, Bonner WM, Usdin K. Evidence for chromosome fragility at the frataxin locus in Friedreich ataxia. *Mutat. Res./Fundam. Mol. Mech. Mutagen.* 2015; 781:14–21.
80. Su XA, Dion V, Gasser SM, Freudenreich CH. Regulation of recombination at yeast nuclear pores controls repair and triplet repeat stability. *Genes Dev.* 2015; 29:1006–1017. [PubMed: 25940904]
81. Saini N, Ramakrishnan S, Elango R, Ayyar S, Zhang Y, Deem A, Ira G, Haber JE, Lobachev KS, Malkova A. Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature.* 2013; 502:389–392. [PubMed: 24025772]
82. Gerhardt J, Tomishima Mark J, Zaninovic N, Colak D, Yan Z, Zhan Q, Rosenwaks Z, Jaffrey Samie R, Schildkraut Carl L. The DNA replication program is altered at the FMR1 locus in fragile X embryonic stem cells. *Mol. Cell.* 2014; 53:19–31. [PubMed: 24289922]

83. Mirkin, Ekaterina V., Mirkin, Sergei M. To switch or not to switch: at the origin of repeat expansion disease. *Mol. Cell.* 2014; 53:1–3. [PubMed: 24411078]
84. Doublié S, Zahn KE. Structural insights into eukaryotic DNA replication. *Front. Microbiol.* 2014; 5:444. [PubMed: 25202305]
85. Branzei D, Szakal B. DNA damage tolerance by recombination: molecular pathways and DNA structures. *DNA Repair.* 2016; 44:68–75. [PubMed: 27236213]
86. Wickramasinghe CM, Arzouk H, Frey A, Maiter A, Sale JE. Contributions of the specialised DNA polymerases to replication of structured DNA. *DNA Repair.* 2015; 29:83–90. [PubMed: 25704659]
87. Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η . *Biochemistry.* 2014; 53:2793–2803. [PubMed: 24716551]
88. Pellegrini L, Costa A. New insights into the mechanism of DNA duplication by the eukaryotic replisome. *Trends Biochem. Sci.* 2016; 41:859–871. [PubMed: 27555051]
89. Zhao L, Washington MT. Translesion synthesis insights into the selection and switching of DNA polymerases. *Genes (Basel).* 2017; 8:24.
90. Belousova EA, Lavrik OI. DNA polymerases β and λ and their roles in cell. *DNA Repair.* 2015; 29:112–126. [PubMed: 25715666]
91. Goodman MF, Woodgate R. Translesion DNA polymerases. *Cold Spring Harbor Perspect. Biol.* 2013; 5:a010363.
92. Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η . *Biochemistry.* 2014; 53:2793–2803. [PubMed: 24716551]
93. Zahn KE, Wallace SS, Doublié S. DNA polymerases provide a canon of strategies for translesion synthesis past oxidatively generated lesions. *Curr. Opin. Struct. Biol.* 2011; 21:358–369. [PubMed: 21482102]
94. 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]
95. Mollersen L, Rowe AD, Illuzzi JL, Hildrestrand GA, Gerhold KJ, Tveteras L, Bjolgerud A, Wilson DM, Bjoras M, Klungland A. Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1mice. *Hum. Mol. Genet.* 2012; 21:4939–4947. [PubMed: 22914735]
96. Lokanga RA, Senejani AG, Sweasy JB, Usdin K. Heterozygosity for a hypomorphic pol β mutation reduces the expansion frequency in a mouse model of the fragile X-related disorders. *PLoS Genet.* 2015; 11:e1005181. [PubMed: 25886163]
97. Hubert L, Lin Y, Dion V, Wilson JH. Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Hum. Mol. Genet.* 2011; 20:4822–4830. [PubMed: 21926083]
98. Guo J, Gu L, Leffak M, Li G-M. MutS β promotes trinucleotide repeat expansion by recruiting DNA polymerase β to nascent (CAG) $_n$ or (CTG) $_n$ hairpins for error-prone DNA synthesis. *Cell Res.* 2016; 26:775–786. [PubMed: 27255792]
99. Lai Y, Budworth H, Beaver JM, Chan NLS, Zhang Z, McMurray CT, Liu Y. Crosstalk between MSH2–MSH3 and pol β promotes trinucleotide repeat expansion during base excision repair. *Nat. Commun.* 2016; 7:12465. [PubMed: 27546332]
100. Tomé S, Manley K, Simard JP, Clark GW, Slean MM, Swami M, Shelbourne PF, Tillier ERM, Monckton DG, Messer A, Pearson CE. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. *PLoS Genet.* 2013; 9:e1003280. [PubMed: 23468640]
101. Dragileva E, Hendricks A, Teed A, Gillis T, Lopez ET, Friedberg EC, Kucherlapati R, Edelmann W, Lunetta KL, MacDonald ME, Wheeler VC. Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol. Dis.* 2009; 33:37–47. [PubMed: 18930147]
102. Kovalenko M, Dragileva E, St Claire J, Gillis T, Guide JR, New J, Dong H, Kucherlapati R, Kucherlapati MH, Ehrlich ME, Lee J-M, Wheeler VC. Msh2 acts in medium-spiny striatal neurons as an enhancer of CAG instability and mutant Huntingtin phenotypes in Huntington's disease knock-in mice. *PLoS One.* 2012; 7:e44273. [PubMed: 22970194]

103. Manley K, Shirley TL, Flaherty L, Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenicmice. *Nat. Genet.* 1999; 23:471–473. [PubMed: 10581038]
104. Owen BAL, Yang Z, Lai M, Gajek M, Badger JD, Hayes JJ, Edelmann W, Kucherlapati R, Wilson TM, McMurray CT. CAG_n-hairpin DNA binds to Msh2–Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol.* 2005; 12:663–670. [PubMed: 16025128]
105. Foiry L, Dong L, Savouret C, Hubert L, Riele Ht, Junien C, Gourdon G. Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenicmice. *Hum. Genet.* 2006; 119:520–526. [PubMed: 16552576]
106. Gomes-Pereira M, Foiry L, Nicole A, Huguet A, Junien C, Munnich A, Gourdon G. CTG trinucleotide repeat big jumps: large expansions, small mice. *PLoS Genet.* 2007; 3:e52. [PubMed: 17411343]
107. Nakatani R, Nakamori M, Fujimura H, Mochizuki H, Takahashi MP. Large expansion of CTG•CAG repeats is exacerbated by MutS β in human cells. *Sci. Rep.* 2015; 5:11020. [PubMed: 26047474]
108. Savouret C. CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J.* 2003; 22:2264–2273. [PubMed: 12727892]
109. van den Broek WJAA. Somatic expansion behaviour of the (CTG)_n repeat in myotonic dystrophy knock-in mice is differentiallyaffected by Msh3 and Msh6 mismatch-repair proteins. *Hum. Mol. Genet.* 2002; 11:191–198. [PubMed: 11809728]
110. Lokanga RA, Zhao X-N, Usdin K. The mismatch repair protein MSH2 is rate limiting for repeat expansion in a fragile X premutation mousemodel. *Hum. Mutat.* 2013; 35:129–136.
111. Zhao X-N, Lokanga R, Allette K, Gazy I, Wu D, Usdin K. A mutS β -dependent contribution of MutSa to repeat expansions in fragile X premutationmice? *PLoS Genet.* 2016; 12:e1006190. [PubMed: 27427765]
112. Du J, Campau E, Soragni E, Ku S, Puckett JW, Dervan PB, Gottesfeld JM. Role of mismatch repair enzymes in GAA{middle dot}TTC triplet-repeat expansion in Friedreich ataxiainduced pluripotent stem cells. *J. Biol. Chem.* 2012; 287:29861–29872. [PubMed: 22798143]
113. Ezzatizadeh V, Pinto RM, Sandi C, Sandi M, Al-Mahdawi S, te Riele H, Pook MA. The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreichataxia mouse model. *Neurobiol. Dis.* 2012; 46:165–171. [PubMed: 22289650]
114. Halabi A, Ditch S, Wang J, Grabczyk E. DNA mismatch repair complex MutS promotes GAA{middle dot}TTC repeat expansion in human cells. *J. Biol. Chem.* 2012; 287:29958–29967. [PubMed: 22787155]
115. Kim H-M, Narayanan V, Mieczkowski PA, Petes TD, Krasilnikova MM, Mirkin SM, Lobachev KS. Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatchrepair. *EMBO J.* 2008; 27:2896–2906. [PubMed: 18833189]
116. Slean MM, Panigrahi GB, Castel AL, Pearson AB, Tomkinson AE, Pearson CE. Absence of MutS β leads to the formation of slipped-DNA for CTG/CAG contractions at primatereplication forks. *DNA Repair.* 2016; 42:107–118. [PubMed: 27155933]
117. Tsutakawa, Susan E., Classen, S., Chapados, Brian R., Arvai, AS., Finger, LD., Guenther, G., Tomlinson, Christopher G., Thompson, P., Sarker, Altaf H., Shen, B., Cooper, Priscilla K., Grasby, Jane A., Tainer, John A. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1superfamily. *Cell.* 2011; 145:198–211. [PubMed: 21496641]
118. Liu Y, Bambara RA. Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeatexpansion. *J. Biol. Chem.* 2003; 278:13728–13739. [PubMed: 12554738]
119. Gupta S, Gellert M, Yang W. Mechanism of mismatch recognition revealed by human MutS β bound to unpaired DNA loops. *Nat. Struct. Mol. Biol.* 2011; 19:72–78. [PubMed: 22179786]
120. Lang WH, Coats JE, Majka J, Hura GL, Lin Y, Rasnik I, McMurray CT. Conformational trapping of mismatch recognition complex MSH2/MSH3 on repair-resistant DNA loops. *Proc. Natl. Acad. Sci.* 2011; 108:E837–E844. [PubMed: 21960445]

121. Liu Y, Prasad R, Beard WA, Kedar PS, Hou EW, Shock DD, Wilson SH. Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta. *J. Biol. Chem.* 2007; 282:13532–13541. [PubMed: 17355977]
122. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, McMurray CT, Wilson SH. Coordination between polymerase β and FEN1 can modulate CAG repeat expansion. *J. Biol. Chem.* 2009; 284:28352–28366. [PubMed: 19674974]
123. Pluciennik A, Burdett V, Baitinger C, Iyer RR, Shi K, Modrich P. Extrahelical (CAG)/(CTG) triplet repeat elements support proliferating cell nuclear antigen loading and MutL endonuclease activation. *Proc. Natl. Acad. Sci.* 2013; 110:12277–12282. [PubMed: 23840062]
124. Bourn RL, De Biase I, Pinto RM, Sandi C, Al-Mahdawi S, Pook MA, Bidichandani SI. Pms2 suppresses large expansions of the (GAA·TTC)_n sequence in neuronal tissues. *PLoS One.* 2012; 7:e47085. [PubMed: 23071719]
125. Gomes-Pereira M. Pms2 is a genetic enhancer of trinucleotide CAG·CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Hum. Mol. Genet.* 2004; 13:1815–1825. [PubMed: 15198993]
126. Pinto RM, Dragileva E, Kirby A, Lloret A, Lopez E, St Claire J, Panigrahi GB, Hou C, Holloway K, Gillis T, Guide JR, Cohen PE, Li G-M, Pearson CE, Daly MJ, Wheeler VC. Mismatch repair genes *mlh1* and *mlh3* modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. *PLoS Genet.* 2013; 9:e1003930. [PubMed: 24204323]
127. Ezzatizadeh V, Sandi C, Sandi M, Anjomani-Virmouni S, Al-Mahdawi S, Pook MA. MutL α heterodimers modify the molecular phenotype of Friedreich ataxia. *PLoS One.* 2014; 9:e100523. [PubMed: 24971578]
128. Batra VK, Shock DD, Beard WA, McKenna CE, Wilson SH. Binary complex crystal structure of DNA polymerase reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc. Natl. Acad. Sci.* 2011; 109:113–118. [PubMed: 22178760]
129. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase β . *Chem. Rev.* 2006; 106:361–382. [PubMed: 16464010]
130. Howard MJ, Rodriguez Y, Wilson SH. DNA polymerase β uses its lyase domain in a processive search for DNA damage. *Nucleic Acids Res.* 2017; 1–11. [PubMed: 27899559]
131. Prasad R, Lavrik OI, Kim SJ, Kedar P, Yang XP, Vande Berg BJ, Wilson SH. DNA polymerase-mediated long patch base excision repair: poly(ADP-ribose) polymerase-1 stimulates strand displacement DNA synthesis. *J. Biol. Chem.* 2001; 276:32411–32414. [PubMed: 11440997]
132. Singhal RK, Wilson SH. Short gap-filling synthesis by DNA polymerase beta is processive. *J. Biol. Chem.* 1993; 268:15906–15911. [PubMed: 8340415]
133. Prasad R, Beard WA, Chyan JY, Maciejewski MW, Mullen GP, Wilson SH. Functional analysis of the amino-terminal 8-kDa domain of DNA polymerase as revealed by site-directed mutagenesis: DNA binding and 5'-deoxyribose phosphate lyase activities. *J. Biol. Chem.* 1998; 273:11121–11126. [PubMed: 9556598]
134. Osheroff WP, Jung HK, Beard WA, Wilson SH, Kunkel TA. The fidelity of DNA polymerase during distributive and processive DNA synthesis. *J. Biol. Chem.* 1999; 274:3642–3650. [PubMed: 9920913]
135. Prindle MJ, Loeb LA. DNA polymerase delta in dna replication and genome maintenance. *Environ. Mol. Mutagen.* 2012; 53:666–682. [PubMed: 23065663]
136. Bertrand P, Tishkoff DX, Filosi N, Dasgupta R, Kolodner RD. Physical interaction between components of DNA mismatch repair and nucleotide excision repair. *Proc. Natl. Acad. Sci.* 1998; 95:14278–14283. [PubMed: 9826691]
137. Lv L, Wang F, Ma X, Yang Y, Wang Z, Liu H, Li X, Liu Z, Zhang T, Huang M, Friedberg EC, Tang TS, Guo C. Mismatch repair protein MSH2 regulates translesion DNA synthesis following exposure of cells to UV radiation. *Nucleic Acids Res.* 2013; 41:10312–10322. [PubMed: 24038355]
138. Mellon I, Rajpal DK, Koi M, Boland CR, Champe GN. Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science.* 1996; 272:557–560. [PubMed: 8614807]

139. Evans E, Sugawara N, Haber JE, Alani E. The *saccharomyces cerevisiae* msh2 mismatch repair protein localizes to recombination intermediates *in vivo*. *Mol. Cell.* 2000; 5:789–799. [PubMed: 10882115]
140. Sugawara N, Paques F, Colaiacovo M, Haber JE. Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl. Acad. Sci.* 1997; 94:9214–9219. [PubMed: 9256462]
141. Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, Hanaoka F, Yasui A, Woodgate R, Gearhart PJ. MSH2–MSH6 stimulates DNA polymerase η , suggesting a role for A:T mutations in antibody genes. *J. Exp. Med.* 2005; 201:637–645. [PubMed: 15710654]
142. Liu Y, Ma X, Guo C. Effects of the N terminus of mouse DNA polymerase κ on the bypass of a guanine-benzo[a]pyrenyl adduct. *J. Biochem.* 2015; 159:471–479. [PubMed: 26634445]
143. Guo C. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.* 2003; 22:6621–6630. [PubMed: 14657033]
144. Crouse GF. Non-canonical actions of mismatch repair. *DNA Repair.* 2016; 38:102–109. [PubMed: 26698648]
145. Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, Starr TB. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol. Sci.* 2010; 120:S130–S145. [PubMed: 21163908]
146. De Biasio, A., Blanco, FJ. Protein-Nucleic Acids Interactions. Elsevier BV; Proliferating cell nuclear antigen structure and interactions; p. 20131–36.
147. Boehm, EM., Gildenberg, MS., Washington, MT. DNA Replication Across Taxa. Elsevier BV; The many roles of PCNA in eukaryotic DNA replication; p. 2016231–254.
148. Kirchmaier AL. Ub-family modifications at the replication fork: regulating PCNA-interacting components. *FEBS Lett.* 2011; 585:2920–2928. [PubMed: 21846465]
149. Kovtun IV, Johnson KO, McMurray CT. Cockayne Syndrome B protein antagonizes OGG1 in modulating CAG repeat length *in vivo*. *Aging.* 2011; 3:509–514. [PubMed: 21566259]
150. Zhao X-N, Usdin K. The transcription-coupled repair protein ERCC6/CSB also protects against repeat expansion in a mouse model of the fragile X premutation. *Hum. Mutat.* 2015; 36:482–487. [PubMed: 25726753]
151. Wong HK, Muftuoglu M, Beck G, Imam SZ, Bohr VA, Wilson DM. Cockayne syndrome B protein stimulates apurinic endonuclease 1 activity and protects against agents that introduce base excision repair intermediates. *Nucleic Acids Res.* 2007; 35:4103–4113. [PubMed: 17567611]
152. Thorslund T, von Kobbe C, Harrigan JA, Indig FE, Christiansen M, Stevnsner T, Bohr VA. Cooperation of the cockayne syndrome group B protein and poly(ADP-ribose) polymerase 1 in the response to oxidative stress. *Mol. Cell. Biol.* 2005; 25:7625–7636. [PubMed: 16107709]
153. 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 n/a–n/a.
154. Chatterjee N, Lin Y, Wilson JH. Fanconi anemia pathway regulates convergent transcription-induced cell death at trinucleotide repeats in human cells. *Postdoc. J.* 2016; 4
155. Chatterjee N, Lin Y, Wilson JH. Mismatch repair enhances convergent transcription-induced cell death at trinucleotide repeats by activating ATR. *DNA Repair.* 2016; 42:26–32. [PubMed: 27131875]
156. Horton JK, Seddon HJ, Zhao M-L, Gassman NR, Janoshazi AK, Stefanick DF, Wilson SH. Role of the oxidized form of XRCC1 in protection against extreme oxidative stress. *Free Radic. Biol. Med.* 2017 [epub ahead of print].
157. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* 2015; 16:583–597. [PubMed: 26370899]
158. Sollier J, Cimprich KA. Breaking bad: R-loops and genome integrity. *Trends Cell Biol.* 2015; 25:514–522. [PubMed: 26045257]
159. Lees-Miller SP, Beattie TL, Tainer JA. Noncoding RNA joins Ku and DNA-PKcs for DNA-break resistance in breast cancer. *Nat. Struct. Mol. Biol.* 2016; 23:509–510. [PubMed: 27273637]
160. Budworth H, McMurray CT. Bidirectional transcription of trinucleotide repeats: roles for excision repair. *DNA Repair.* 2013; 12:672–684. [PubMed: 23669397]

161. Groh M, Albulescu LO, Cristini A, Gromak N. Senataxin: genome guardian at the interface of transcription and neurodegeneration. *J. Mol. Biol.* 2016;1–15.
162. Hoch NC, Hanzlikova H, Rulten SL, Tetreault M, Komulainen E, Ju L, Hornyak P, Zeng Z, Gittens W, Rey SA, Staras K, Mancini GM, McKinnon PJ, Wang ZQ, Wagner JD, Care4Rare Canada C, Yoon G, Caldecott KW. XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nature.* 2017; 541:87–91. [PubMed: 28002403]
163. 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.* 2012; 33:406–417. [PubMed: 23149945]
164. Colak D, Zaninovic N, Cohen MS, Rosenwaks Z, Yang WY, Gerhardt J, Disney MD, Jaffrey SR. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science.* 2014; 343:1002–1005. [PubMed: 24578575]
165. Groh M, Gromak N. Out of balance: R-loops in human disease. *PLoS Genet.* 2014; 10:e1004630. [PubMed: 25233079]
166. Groh M, Lufino MMP, Wade-Martins R, Gromak N. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragileX syndrome. *PLoS Genet.* 2014; 10:e1004318. [PubMed: 24787137]
167. Loomis EW, Sanz LA, Chédin F, Hagerman PJ. Transcription-associated R-Loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.* 2014; 10:e1004294. [PubMed: 24743386]
168. Zhou Y, Kumari D, Sciascia N, Usdin K. CGG-repeat dynamics and FMR1 gene silencing in fragile X syndrome stem cells and stem cell-derived neurons. *Mol. Autism.* 2016; 7:42. [PubMed: 27713816]
169. Sopher, Bryce L., Ladd, Paula D., Pineda, Victor V., Libby, Randell T., Sunkin, Susan M., Hurley, James B., Thienes, Cortlandt P., Gaasterland, T., Filippova, Galina N., La Spada, Albert R. CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncodingRNA. *Neuron.* 2011; 70:1071–1084. [PubMed: 21689595]
170. Lin Y, Dent SYR, Wilson JH, Wells RD, Napierala M. R loops stimulate genetic instability of CTG{middle dot}CAG repeats. *Proc. Natl. Acad. Sci.* 2009; 107:692–697. [PubMed: 20080737]
171. 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]
172. Reddy K, Schmidt MHM, Geist JM, Thakkar NP, Panigrahi GB, Wang YH, Pearson CE. Processing of double-R-loops in (CAG){middle dot}(CTG) and C9orf72 (GGGGCC){middle dot}(GGCCCC) repeats causes instability. *Nucleic Acids Res.* 2014; 42:10473–10487. [PubMed: 25147206]
173. 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]
174. Stork CT, Bocek M, Crossley MP, Sollier J, Sanz LA, Chédin F, Swigut T, Cimprich KA. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *eLife.* 2016; 5:e17548. [PubMed: 27552054]
175. Hanawalt PC, Spivak G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* 2008; 9:958–970. [PubMed: 19023283]
176. Spivak G. Transcription-coupled repair: an update. *Arch. Toxicol.* 2016; 90:2583–2594. [PubMed: 27549370]
177. Sollier J, Stork Caroline T, García-Rubio María L, Paulsen Renee D, Aguilera A, Cimprich Karlene A. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell.* 2014; 56:777–785. [PubMed: 25435140]
178. Madireddy A, Kosiyatrakul Settapong T, Boisvert Rebecca A, Herrera-Moyano E, García-Rubio María L, Gerhardt J, Vuono Elizabeth A, Owen N, Yan Z, Olson S, Aguilera A, Howlett Niall G, Schildkraut Carl L. FANCD2 facilitates replication through common fragile sites. *Mol. Cell.* 2016; 64:388–404. [PubMed: 27768874]

179. Pandey S, Ogloblina AM, Belotserkovskii BP, Dolinnaya NG, Yakubovskaya MG, Mirkin SM, Hanawalt PC. Transcription blockage by stable H-DNA analogs in vitro. *Nucleic Acids Res.* 2015; 43:6994–7004. [PubMed: 26101261]
180. Shishkin AA, Voineagu I, Matera R, Cherng N, Chernet BT, Krasilnikova MM, Narayanan V, Lobachev KS, Mirkin SM. Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol. Cell.* 2009; 35:82–92. [PubMed: 19595718]
181. Zhang Y, Shishkin Alexander A, Nishida Y, Marcinkowski-Desmond D, Saini N, Volkov Kirill V, Mirkin Sergei M, Lobachev Kirill S. Genome-wide screen identifies pathways that govern GAA/TTC repeat fragility and expansions in dividing and nondividing yeast cells. *Mol. Cell.* 2012; 48:254–265. [PubMed: 22959270]
182. Voineagu I, Narayanan V, Lobachev KS, Mirkin SM. Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc. Natl. Acad. Sci.* 2008; 105:9936–9941. [PubMed: 18632578]
183. Costantino L, Sotiriou SK, Rantala JK, Magin S, Mladenov E, Helleday T, Haber JE, Iliakis G, Kallioniemi OP, Halazonetis TD. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science.* 2013; 343:88–91. [PubMed: 24310611]
184. 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. *Nat. Struct. Mol. Biol.* 2016; 24:55–60. [PubMed: 27918542]
185. Llorente B, Smith CE, Symington LS. Break-induced replication: what is it and what is it for? *Cell Cycle.* 2008; 7:859–864. [PubMed: 18414031]
186. Geronimo CL, Zakian VA. Getting it done at the ends: Pif1 family DNA helicases and telomeres. *DNA Repair.* 2016; 44:151–158. [PubMed: 27233114]
187. Wild P, Matos J. Cell cycle control of DNA joint molecule resolution. *Curr. Opin. Cell Biol.* 2016; 40:74–80. [PubMed: 26970388]
188. Dixon M. Examining the potential role of DNA polymerases η and ζ in triplet repeat instability in yeast. *DNA Repair.* 2002; 1:763–770. [PubMed: 12509280]
189. Ye Y, Kirkham-McCarthy L, Lahue RS. The *Saccharomyces cerevisiae* Mre11-Rad50-Xrs2 complex promotes trinucleotide repeat expansions independently of homologous recombination. *DNA Repair.* 2016; 43:1–8. [PubMed: 27173583]
190. Minocherhomji S, Ying S, Bjerregaard VA, Bursomanno S, Aleliunaite A, Wu W, Mankouri HW, Shen H, Liu Y, Hickson ID. Replication stress activates DNA repair synthesis in mitosis. *Nature.* 2015; 528:286–290. [PubMed: 26633632]
191. Sarni D, Kerem B. The complex nature of fragile site plasticity and its importance in cancer. *Curr. Opin. Cell Biol.* 2016; 40:131–136. [PubMed: 27062332]
192. Memisoglu A, Samson L. Base excision repair in yeast and mammals. *Mutat. Res.* 2000; 451:39–51. [PubMed: 10915864]
193. Brautigam CA, Steitz TA. Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* 1998; 8:54–63. [PubMed: 9519297]
194. Garcia-Diaz M, Bebenek K, Larrea AA, Havener JM, Perera L, Krahn JM, Pedersen LC, Ramsden DA, Kunkel TA. Template strand scrunching during DNA gap repair synthesis by human polymerase λ . *Nat. Struct. Mol. Biol.* 2009; 16:967–972. [PubMed: 19701199]

Fig. 1.

Features of trinucleotide expansion. (A–C) Schematic diagram of three common secondary structures formed by TNRs: hairpins (A), triple helix (B) or quadruplex (C). (D) Schematic diagram of the TNR location of some representative genes, as indicated. FMR-1 is Fragile X mental retardation; HD is Huntington disease; SCA 1–7, spinocerebellar ataxia types 1–7; SMBA, spino-muscular bulbar atrophy; DRPLA, dentatorubral-pallidoluysian atrophy; FRDA, Friedreich's ataxia; and DM, myotonic dystrophy 1. Pink indicates exons and lines indicates non-coding regions, intron, 5'-untranslated regions (5'utr), and 3'-untranslated regions (3'utr). (E) Schematic representation of the relationship between HD disease severity and expansion as indicated by the triangle. The CAG repeats lengths in the HD gene that correspond to the normal, intermediate (premutation) and full-length mutation are indicated. The larger the repeat tract length the more severe the disease; red triangle indicates the disease range and corresponds to neuronal death in the affected areas of the brain, also shown in red. (F) Gel resolution of individual cells from small pool PCR of striatal tissue of *HdhQ72* mouse at 3 (left) and 24 (right) months. The numbers on the left side of the panel indicate the number of CAG repeats of the normal of seven CAG repeats, and mutated allele of 71 CAG repeats. With time, the disease-length allele grows longer, due to somatic expansions. Taken from Ref. [57].

Fig. 2.

Polymerase features and model of MSH2-MSH3-pol β complex. (A,B) All DNA polymerases carry out three common steps using a common structural framework, depicted as a finger (blue)(F), palm (brown)(P), and thumb (green)(T), which positions incoming nucleotide triphosphates in relation to the template strand, holds in place the elongating DNA duplex, and the enzyme's active site, respectively. (C–F) Four examples are shown which have been similarly oriented with respect to each other by superposition of the first two base pairs at the primer terminus. (C) *Taq* DNA polymerase bound to DNA. The DNA stacks against the fingers and is contacted across the minor groove by the thumb domain. (D) The binary complex of HIV-1 RT, and DNA. This structure does not have a nucleotide-binding α -helix in the fingers domains. Instead, a β -hairpin probably performs this function.

(E) RB69 gp43. A likely DNA-binding α -helix has been highlighted. The thumb domain would have to move toward the primer terminus to bind DNA analogously to the other polymerases. (F) The ternary complex of rat pol β with DNA and dideoxy-NTP. Structures D-F are taken from Ref. [193]. (G) Cartoon of hypothetical pol β -MSH2-MSH3 interaction. In this model, MSH2-MSH3 stabilizes the lyase domain of pol β and binds on ssDNA near the 5' phosphate. (H) Schematic of pol λ scrunching mechanism. The DNA polymerase is represented as a yellow surface, and each of the three DNA binding sites (I, II and III) are represented. (H) Cartoon version of the pol λ mechanism. Pol λ accomplishes primer alignment by use of "scrunching". This scrunching is accompanied by repositioning of the uncopied template nucleotide to an extrahelical position, where it binds within a pocket comprised of three conserved amino acids. This pocket seems to contribute to processive gap filling, which may help to prevent premature release of repair intermediates that are potentially toxic and/or mutagenic. Taken from Ref. [194].

Fig. 3.

Proposed mechanism for MSH2-MSH3-pol β crosstalk. (A) MSH2-MSH3 promotes TNR expansion via suppression of repeat deletion during BER. Oxidative stress induces an oxidized DNA base in a TNR tract such as 8-oxoG. OGG1 removes the 8-oxoG and leaves an abasic site that is subsequently 5'-incised by APE1. (A) In the absence of MSH2-MSH3 ('-MSH2-MSH3'), pol β opens the template to generate single strand DNA, and deletion of the template strand occurs after endonuclease excision. Thus MSH2-MSH3 suppresses deletion and promotes expansion. (B) The pol β •MSH2-MSH3 complex loads onto DNA at the APE1 incision, and promotes gap-filling synthesis and flap formation. MSH2-MSH3 inhibits FEN1 removal of the flap. However, MSH2-MSH3 interaction with the loop reorients MSH2-MSH3 and allows flap re-alignment on the damaged strand to generate a new and shorter flap suitable for FEN1 cleavage and ligation. Incorporation of the loop by an endonuclease results in expansion. Taken from Ref. [99]. (C) "in cis" error free loop removal after gap filling. MSH2-MSH3 (purple and blue balls) recruits an endonuclease (orange), which clips on the side of the hairpin. Gap filling (purple is polymerase) restores the correct sequence. (C) "in trans" Error-prone loop incorporation. Taken from Ref. [9]. The hairpin loop is not removed. Rather, MSH2-MSH3 (blue-purple dimer) recruits or misdirects an endonuclease (orange ball) to clip the strand across the hairpin. The loop opens and it is incorporated into DNA after gap filling synthesis and becomes an expansion. (D) The toxic oxidation cycle for somatic age-dependent expansion. Continuous oxidation-excision-repair cycles allow for increasing somatic expansion with age. 1 is mitochondrial

stress that increases oxidative damage on DNA; 2 is excision by the BER machinery; 3 is strand displacement synthesis; 4 is hairpin formation and stabilization by MSH2-MSH3. Taken from Ref. [94].

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4.

Three models for conversion of a TCR intermediate to a double strand break at a deleterious R-loop. (A) The NER/TCR endonucleases XPF and XPG excise the RNA-DNA hybrid of an R-loop. The single-stranded gap may be converted into a DSB either by “double cutting” on both of the bubble strands, and can involve the same or different set of structure-specific endonucleases. (B) The NER/TCR endonucleases XPF and XPG excise the RNA-DNA hybrid of an R-loop, and encounter with the replication fork leads to a DSB. The endonucleases XPF and XPG can recognize both strands of an R-loop and generate a DSB. (C) The TC-NER machinery may be recruited to R-loops during replication, and the concerted action of the endonucleases XPF and XPG with other structure-specific endonucleases may generate a DSB. HR is homologous recombination, NHEJ is non-homologous end joining, and BIR is break-induced repair. Taken from Ref. [158].

Fig. 5.

Mechanisms of large-scale expansions at TNRs by template switching and BIR. (A) Template switch model. (a) Replication-fork progression proceeds through the repetitive region. Closed circles denote the 5' end of the template strand, and closed arrows denote the 3' end of the nascent strand. Polymerase stalls when it encounters a TNR. (b) The lagging strand forms a hairpin. (c) Isomerization of the resulting four-way junction (chicken-foot structure). (d) The Four-way junction from c isomerizes to form a Holliday junction, which can be cleaved by resolvase proteins (gray arrows). (e) A one-ended DSB is subject to 5'-to-3' resection (dotted line), thus resulting in a single-stranded repetitive sequence. (B) BIR at a TNR. In a BIR model, an endonuclease may act directly on the hairpin structure at the stalled fork, and yield a one-ended DSB. Gray arrow denotes the potential site of endonuclease cleavage, which could allow the intermediate to proceed directly to f. (f) The single-stranded repetitive sequence is coated by replication protein A, then Rad51 (dark gray circles) and forms stable hairpin structure(s). (g) Out-of-register invasion results in formation of a D loop. A leftward-moving convergent replication fork shows the leading strand (green) and lagging strand (yellow) emanating from the adjacent replication origin. (h) Pol32-dependent DNA synthesis continues, thus further extending D-loop progression. (i) The D loop converges with the leftward-moving replication fork, thus resulting in a Holliday junction, which must be resolved (gray arrows). (j) The nascent DNA strand (bottom) would have accumulated extra CAG repeats at a length equivalent to that of the out-of-register invasion step, thus resulting in a large-scale expansion. Taken from Ref. [185]. (B) Cartoon of normal BIR. Taken from Ref. [184].

Table 1Polymerase Families.^a

Family	Types of DNA polymerase	Species	Examples
A	Replicative and Repair Polymerases	Eukaryotic and Prokaryotic	T7 DNA polymerase, Pol I, Polv, Polφ, and DNA Polγ
B	Replicative and Repair Polymerases	Eukaryotic and Prokaryotic	Pol II, Pol B, Polζ, Polα, Polδ and Pole
C	Replicative Polymerases	Prokaryotic	Pol III
D	Replicative Polymerases	Euryarchaeota	Not well-characterized
X	Replicative and Repair Polymerases	Eukaryotic	Polβ, Polλ, Polμ and Terminal deoxynucleotidyl transferase
Y	Replicative and Repair Polymerases	Eukaryotic and Prokaryotic	Polη, Polκ, Pol η, Pol IV, and Pol V
RT	Replicative and Repair Polymerases	Viruses, Retroviruses, and Eukaryotic	Telomerase, Hepatitis B virus

^aDetail of Polymerase Families are described in detail [70–72].

Table 2

Polymerase Families and Functions.^a

DNA pols (Family)	Primary Known Function	Rep	BER	NER	HR NHEJ	MMR	By-pass	H. Fid	Ly	Exo
Pol γ (A)	Mitochondrial DNA replication and repair	X						X	X	X
Pol ϕ (A)	Defense against ionizing radiation-induced DNA damage		X	X	X	X	X	X	X	X
Pol ν (A)	ICL repair or testis-specific function?			X						
Pol α (B)	DNA replication priming	X						X		
Pol δ (B)	DNA replication, NER, and MMR		X	X		X	X	X	X	X
Pol ϵ (B)	DNA replication, NER, and MMR	X	X	X		X	X	X	X	X
Pol ζ (B)	TLS and mutagenesis						X			
Pol β (X)	BER and meiotic recombination		X		X				X	
Pol λ (X)	V(D)J recombination: possibly end joining and BER		X		X					
Pol μ (X)	V(D)J recombination: possibly end joining				X					
TDT (X)	Immunoglobulin diversity at junctions of coding regions				X					
REV1 (Y)	TLS and mutagenesis, anchor for several DNA polymerases						X			
Pol η (Y)	Bypass of UV radiation-induced DNA adducts, especially CPDs						X			
Pol ι (Y)	Backup enzyme for bypass of UV radiation-induced DNA adducts and BER						X			
Pol κ (Y)	Bypass of bulky adducts, backup enzyme for NER						X			

^aBER, base excision repair; CPD, cyclobutene pyrimidine dimer; ICL interstrand crosslink; MMR, mismatch repair; NER, nucleotide excision repair; Pol, polymerase; TDT, terminal deoxynucleotidyltransferase; TLS, translesion DNA synthesis; UV, ultraviolet. Each polymerase is described in detail [70–72].