

Mismatch repair is a double-edged sword in the battle against microsatellite instability

Carson J. Miller  and Karen Usdin 

Laboratory of Cell and Molecular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA

Review

Cite this article: Miller CJ, Usdin K (2022). Mismatch repair is a double-edged sword in the battle against microsatellite instability. *Expert Reviews in Molecular Medicine* **24**, e32, 1–8. <https://doi.org/10.1017/erm.2022.16>

Received: 7 January 2022
Revised: 26 April 2022
Accepted: 27 April 2022

Key words:

Mismatch Repair (MMR); Repeat expansion Diseases; Microsatellite instability (MSI); Short tandem repeats (STRs)

Author for correspondence:

Karen Usdin, E-mail: karenu@nih.gov

Abstract

Roughly 3% of the human genome consists of microsatellites or tracts of short tandem repeats (STRs). These STRs are often unstable, undergoing high-frequency expansions (increases) or contractions (decreases) in the number of repeat units. Some microsatellite instability (MSI) is seen at multiple STRs within a single cell and is associated with certain types of cancer. A second form of MSI is characterised by expansion of a single gene-specific STR and such expansions are responsible for a group of 40+ human genetic disorders known as the repeat expansion diseases (REDs). While the mismatch repair (MMR) pathway prevents genome-wide MSI, emerging evidence suggests that some MMR factors are directly involved in generating expansions in the REDs. Thus, MMR suppresses some forms of expansion while some MMR factors promote expansion in other contexts. This review will cover what is known about the paradoxical effect of MMR on microsatellite expansion in mammalian cells.

Introduction

Tracts containing tandem arrays of short perfect- or near-perfect repeat units are common in the human genome (Ref. 1). These short tandem repeats (STRs), or microsatellites, consist of repeat units that are generally ~1–6 nucleotides long. The STRs are found in promoters, exons, introns, as well as in intergenic regions where they can impact gene expression in a myriad of different ways, including affecting promoter activity, RNA polymerase processivity, splicing, translation rates and protein function (Ref. 2). Many STRs are polymorphic, giving rise to expansions, or increases in the number of repeat units; as well as contractions, or loss of repeat units. Such STRs are sometimes referred to as variable number tandem repeats (VNTRs). STRs can be a significant source of human genetic variation and the instability of some of these tracts can have biological consequences because of their intrinsic effects on gene expression (Refs 2, 3). In addition, many of these sequences form secondary structures that are thought to make them difficult to replicate (Refs 4–11). This can result in the generation of chromosome abnormalities of different kinds (reviewed in Ref. 12).

Two major classes of human STR expansions are known: the first class is associated with genome-wide microsatellite instability (MSI), while the second class is associated with expansions of a specific microsatellite. Genome-wide MSI is associated with a predisposition to certain cancers including haematological malignancies as well as certain colon, urothelial, hepatobiliary, pancreatic, bladder, kidney, prostate, endometrial, ovarian and breast cancers (Refs 13–20). In contrast, the locus-specific expansions define the repeat expansion disorders (REDs), a group of 40+ human genetic disorders that are primarily neurological or neurodevelopmental in nature. Diseases in this group include Huntington's disease (HD), caused by expansion of CAG/CTG-STR in the first exon of the huntingtin (*HTT*) gene (Ref. 21); Friedreich ataxia (FRDA), caused by a GAA/TTC-STR in intron 1 of the frataxin (*FXN*) gene (Ref. 22); myotonic dystrophy type 1 (DM1), caused by expansion of a CTG/CAG-STR in the 3' UTR of the *DMPK* gene (Refs 23, 24); and the Fragile X-related disorders (FXDs), resulting from expansion of a CGG/CCG-STR in the 5' UTR of the *FMR1* gene (Refs 25–28). While it was initially thought that STR expansions in the REDs occurred by a mechanism similar to cancer-associated MSI, emerging evidence suggests these two types of STR expansions have completely different molecular mechanisms. Genome-wide MSI results from errors arising during DNA synthesis that normally would be repaired by the MMR machinery, that is, MMR factors all act anti-mutagenically at these loci to suppress expansions. In contrast, STR expansions in the REDs actually require certain components of the MMR machinery, that is, these MMR factors can also act pro-mutagenically. While work in model systems suggests that other mechanisms of STR expansion may be possible (e.g. Ref. 29), this review will focus on what is currently known about STR expansions arising from either the pro- or anti-mutagenic roles of MMR.

Cancer-associated MSI

During DNA replication two types of errors can be introduced into DNA: mismatches and insertions or deletions (INDELs). Mismatches arise from insertion of the incorrect base in the daughter strand by the DNA polymerase. Most of these mismatches are removed by the

© National Institutes of Health, 2022. This is a work of the US Government and is not subject to copyright protection within the United States. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

proofreading function of the polymerase, but those that escape this proofreading will cause point mutations if the daughter strand is replicated before the mismatch is removed (Refs 30, 31). INDELS are thought to result primarily from the dissociation of the DNA polymerase from the template thus creating an opportunity for two strands of DNA to slip relative to one another. Dissociation might be exacerbated by an encounter with impediments to replication fork progression such as those formed by unusual DNA structures, strongly bound proteins or collisions with a transcription complex, while mispriming may be favoured when one of the DNA strands forms a stable secondary structure (Ref. 32). When strand-slippage occurs within a STR, out-of-register reannealing can occur with priming from the slipped position as illustrated in Figure 1. This results in either a loop out of the template strand or a loop out on the nascent strand depending on whether reannealing occurs 5' or 3' on the template. Failure to remove the loop

out leads to expansions if the loop out is in the nascent strand and contractions if it is in the template strand (Refs 33–36).

Repair of these replication errors is carried out primarily by the MMR machinery that travels behind the replication complex. In eukaryotes, recognition of mismatches and INDELS during MMR is accomplished by either of two MutS complexes, both of which are heterodimers of homologous proteins: MSH2/MSH6 in the MutS α complex and MSH2/MSH3 in the MutS β complex. MutS α is primarily involved in the recognition of mismatches and 1 base INDELS (Refs 37–39). MutS β , on the other hand, is involved primarily in the repair of larger loops (Refs 40–42). (A third MutS complex found in mammals, MutS γ , is a MSH4/MSH5 heterodimer that functions almost exclusively in meiotic crossover resolution (Ref. 43).) After lesion binding, MutS α and MutS β recruit a member of the MutL family of proteins. Mammals have three different MutL

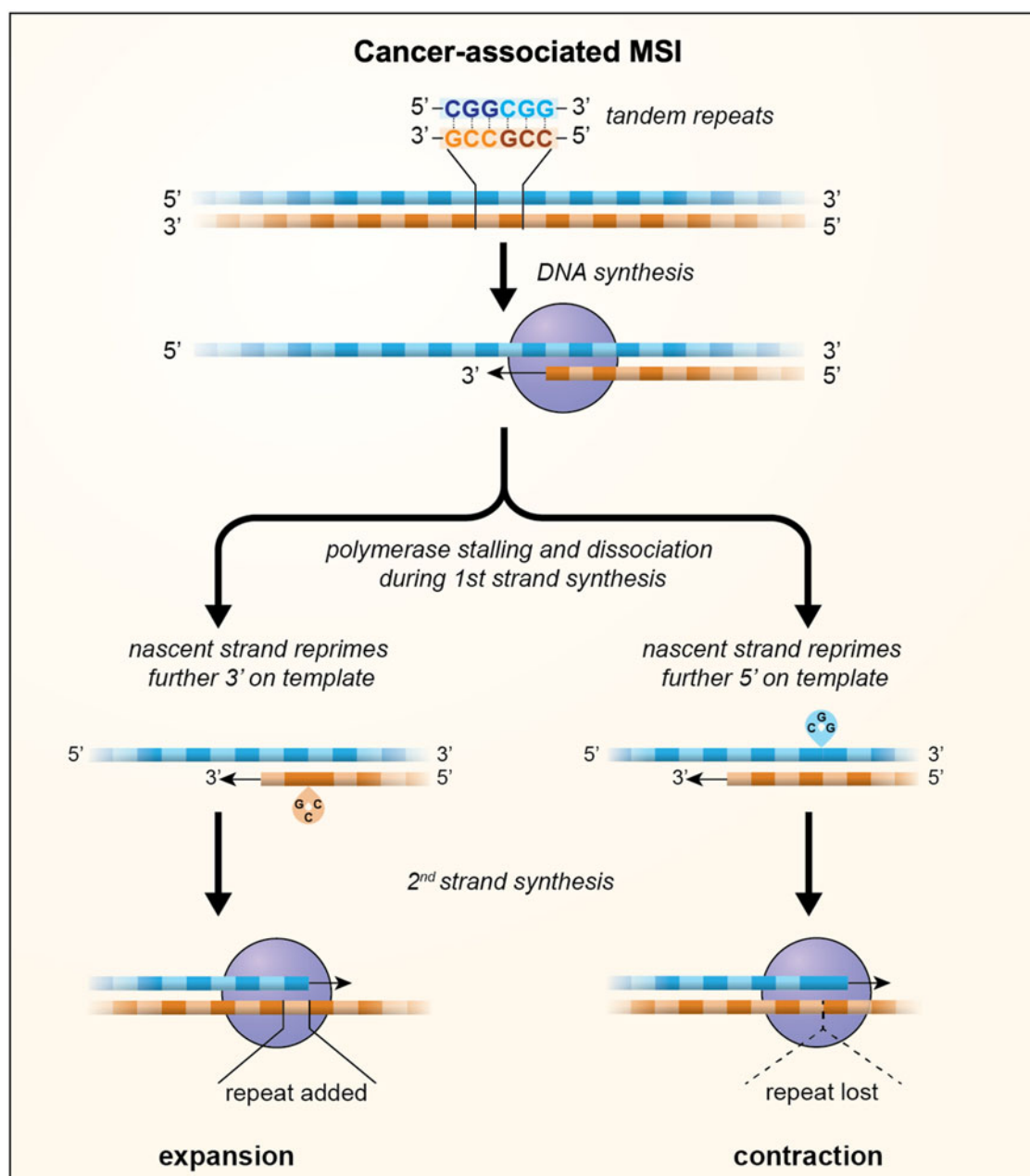


Fig. 1. A model for MSI in which loop outs generated by strand-slippage of the nascent strand during DNA replication escape MMR and result in the incorporation of a small number of additional units, indicated by the green boxes, or the loss of a small number of repeat units, indicated by the dotted triangle. Whether repeats are gained or lost depends on whether repriming occurs further 3' on the template resulting in nascent strand loop outs or further 5' on the template resulting in loop outs being formed on the template strand.

complexes. Like the MutS complexes, each of the MutL complexes are heterodimers; consisting of MLH1 bound to either PMS2, MLH3 or PMS1 to form MutL α , MutL γ or MutL β , respectively (Refs 44–46). MutL α is the most abundant of the MutL complexes and is responsible for most MMR. MutL γ plays a minor role in MMR, primarily in cooperation with MutS β (Refs 47, 48). MutL α and MutL γ are nucleases that introduce nicks into the MMR template, a critical step in the repair process. The role of MutL β in MMR is unclear (Refs 49, 50). While PMS1 is more abundant than MLH3, it lacks the DQHA(X)₂E(X)₄E nuclease motif present in both PMS2 and MLH3 and its loss is not associated with increased MSI in mice (Ref. 49). Multiple lines of evidence suggest that after initial mismatch binding by MutS α , additional MutS α complexes load onto the DNA, followed by recruitment of multiple MutL α proteins (Refs 51, 52). A similar situation may apply to MutS β -directed recruitment of MutL γ , since formation of MutL γ polymers on the mismatch template has also been shown to be important for proper MutL γ -mediated repair (Ref. 53). Excision of the nicked strand is carried out by a 5' to 3' exonuclease such as exonuclease 1 (EXO1) (Refs 54, 55) or FAN1 (Ref. 56). Strand-displacement synthesis by Pol δ can also remove the nicked strand. This is followed by repair synthesis by Pol δ , with sealing of the remaining nick by DNA ligase I to complete the repair process.

While this process is relatively efficient, strand-slippage occurs so frequently that some MSI occurs even in the presence of the normal MMR machinery. Mean rates of $\sim 10^{-5}$ to 10^{-7} MSI events per locus per cell generation have been reported in human cells with functional MMR, orders of magnitude higher than the mutation rate seen in unique sequence (Refs 57–59). Loss of MMR results in rates of MSI that can be 2–3 orders of magnitude higher (Ref. 60). The wide variation in mutation rates of different STRs is related in part to the size of the repeat units, their sequence composition and the size and purity of the repeat tract (Refs 61, 62). The likelihood of instability at a specific microsatellite is also related to the normal target of the dysfunctional MMR gene. So, mutations in MSH2, MLH1 and PMS2 increase instability of microsatellites containing mononucleotide, dinucleotide and tetranucleotide repeat units; MSH6 mutations affect microsatellites with mononucleotide and some dinucleotide repeat units; and MSH3 mutations affect dinucleotide and tetranucleotide containing repeat units, but not ones consisting of mononucleotide repeat units.

More than 90% of MSI events involve the gain or loss of a single repeat unit with a very limited number of mutations involving multiple units (Ref. 63). MSI often exhibits an expansion bias (Refs 64–67). This bias is reduced at very large microsatellites (Refs 64, 65), perhaps reflecting the formation of stable secondary structures and the resultant difficulties associated with replication of the region. This could result in a dependency on proteins such as the Werner's syndrome helicase (WRN) to remove the secondary structure thus allowing replication to proceed (Ref. 68). It has also been suggested that these structures promote error-prone DNA synthesis resulting in mutations that affect the purity of the repeat tract (Ref. 10). This in turn would reduce the likelihood of further expansion. While most MSI events of this kind involve a single repeat unit, MSI with an expansion bias could over time result in the large microsatellites that accumulate in cancer cell lines such as HCT116 and KM12 that lack MLH1 (as well as MSH3 in the case of HCT116) (Ref. 68).

MSI in the REDs

In contrast, studies in RED patient cohorts using genome-wide association (GWA) or the testing of candidate MMR gene polymorphisms suggest that functional MMR components are

required for some, if not all, STR expansions that cause the REDs (Refs 69–74). This is consistent with evidence from mouse and human cell models of a number of these disorders that shows a requirement for MutS β and MutL γ (reviewed in Ref. 75). Canonical MMR *per se* is unlikely to be responsible for these expansions since mutations in EXO1 and FAN1, 5'-3' exonucleases that act downstream of the MutS and MutL proteins in normal MMR, protect against expansion in mouse or human tissue culture models (Refs 73, 76–79), and GWA studies data are consistent with a protective role for FAN1 in reducing expansions in humans (Refs 69, 73). Furthermore, Lig4, the ligase required for non-homologous end joining (NHEJ), a form of double-strand break repair (DSBR), also protects against expansion in a mouse model of the FXDs (Ref. 80). This suggests that NHEJ competes with the expansion pathway for access to a common DSB intermediate.

As with cancer-associated MSI, the extent of expansion in the REDs is related in part to the length and purity of the repeat tract (Refs 81–88). Mathematical modelling of human somatic expansions and empirical observations of both germline and somatic expansions over time in mice support the idea that most expansion events involve the addition of 1–2 repeat units (Refs 89, 90). As with expansions arising in MMR-deficient cells, this can result in the production of much larger alleles over time. However, in some cell types the MMR factor-dependent expansions occur at frequencies that are orders of magnitude higher than the MSI occurring in the absence of MMR at the same locus (Ref. 91).

The mechanism responsible for this high-frequency expansion process is not fully understood. Clues to what this process may be include the fact that expansions can occur in post-mitotic cells such as oocytes and neurons (Refs 78, 92–95). Thus, these events can be independent of chromosomal replication. Furthermore, the fact that the STR in the X-linked *FMRI* gene that causes the FXDs only expands when it is on the active X chromosome indicates that STR expansion requires transcription or transcriptionally-competent chromatin (Ref. 96). A role for oxidative damage is suggested by the fact that the loss of OGG1 and NEIL1, DNA glycosylases involved in the base excision repair of oxidative damage, decreases the expansion frequency in a HD mouse model (Refs 97, 98) and exogenous sources of oxidative stress increase the expansion frequency in some mouse and tissue culture models (Refs 99, 100). However, antioxidants only have a minimal effect on expansion (Refs 101, 102) and even in the absence of OGG1 and NEIL1 many expansions are still seen (Refs 97, 98). Thus, endogenous oxidative stress may not be the only trigger for the expansion process or even the most important one.

Although MutL γ is the least abundant of all the MutL complexes, its nuclease activity is required for repeat expansion (Refs 103–105). Thus, expansion either involves a substrate that is bound preferentially by MutS β /MutL γ or MutL γ cleavage is uniquely able to generate an intermediate that can be processed to generate an expansion. Interestingly, MutL γ has been shown to cut the DNA strand opposite to the mismatch *in vitro* (Ref. 106) and MutL γ is required during meiosis for the resolution of Holliday junctions (HJs) (Ref. 107). Loop outs formed within the STR by both DNA strands might resemble such a four-way junction. Such structures could potentially arise any time the repeat tract was unpaired since out-of-register reannealing could occur particularly if one or both strands formed stable secondary structures as many STRs do (reviewed in Refs 75, 108). Cleavage of the opposite strand at each of the loop outs could then result in a staggered DSB. Interestingly, we have shown that EXO1, which plays a structural role in determining the orientation of cleavage of HJs (Ref. 47), also plays a nuclease-independent role in protecting against repeat expansion (Ref. 77). This raises the possibility that cleavage of the expansion intermediate may result in a DSB that

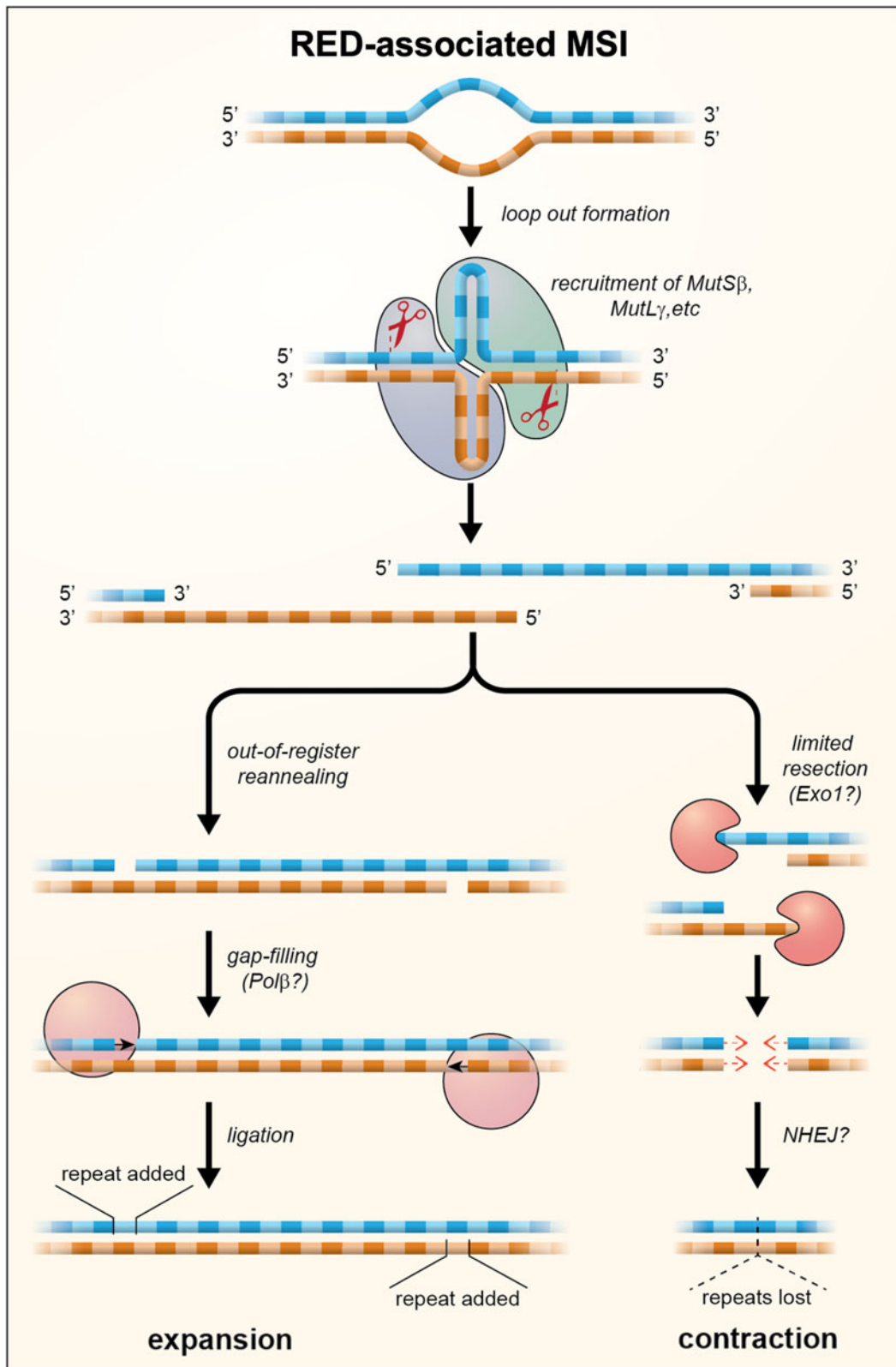


Fig. 2. A model for repeat expansions and contraction in the REDs in which loop outs are formed on one or both strands during transcription or at other times that the DNA was unpaired. These loop outs are then bound by MutS and MutL proteins. Cleavage by MutL γ results in the formation of a staggered DSB with 5' overhangs. Out-of-register reannealing of the DSB can produce a substrate for simple gap filling which results in the addition of repeat units. Exonucleolytic processing of the DSB can result in products with shorter 5' overhangs or blunt ends. These products may then be processed, perhaps by NHEJ or gap-filling, to generate the loss of repeat units as indicated by the dotted triangle. The extent of contraction would depend on the amount of exonucleolytic cleavage that occurs prior to repair.

is prone to expand and one that is not. A model for repeat expansion that accommodates these observations is shown in Figure 2. In this model expansions arise when out-of-register reannealing of the DSB occurs. This leaves a gap of a small number of repeat units that

is then repaired by gap-filling. The net effect is that a small number of repeats are added to the repaired allele.

In addition to MutS β and MutL γ , MutL β has also been shown to be required for expansion in embryonic stem cells from a

mouse FXD model (Ref. 87). An active role for MutL β in generating expansions in HD is also suggested by the fact that *PMS1* variants predicted to be deleterious are most frequently associated with a later age at onset/less severe phenotype in HD cohorts (Ref. 73). However, since MutL β lacks a nuclease and is not required for MMR, how it contributes to expansions is unclear. This is not the only unresolved issue; although MutS α contributes to expansions in FXD and FRDA mice and FRDA iPSCs (Refs 31, 109, 110), little if any effect of the loss of MSH6 was seen in DM1 mouse model (Ref. 111) or in a human cell model system of FRDA (Ref. 112). Furthermore, reducing MutL α levels also has different effects in various systems. In a mouse model of the FXDs, MutL α is required for expansions in ESCs (Ref. 87), while in a mouse model of DM1 loss of MutL α only resulted in a 50% decrease in expansions (Ref. 113). Reduced PMS2 caused no change in the expansion frequency in a human cell model of FRDA (Ref. 100), whereas a mouse model of FRDA lacking *Pms2* showed an increase in expansions (Ref. 109). The latter is consistent with the observation that a missense mutation in *PMS2* correlates with an earlier age at onset in HD (Ref. 69). These differences do not necessarily mean that the expansion mechanisms in these diseases are fundamentally different. Since multiple MutS and MutL complexes are involved in binding to a mismatch, a case can be made that MutS α and MutL α are able to act in an auxiliary capacity to promote expansions when the essential factors, such as MutS β and MutL γ , and perhaps MutL β , are limiting (Ref. 75).

Contractions are also seen in the REDs-associated STRs and their mouse models, although their aetiology is less well understood. A bimodal distribution of contractions is seen in the germline of a mouse model of the FXDs, with some contractions involving the loss of just 1–2 repeat units whilst others involve the loss of much larger numbers of repeat units (Ref. 114). These larger contractions are sometimes difficult to discern in somatic cells because, unlike expansions, the contraction products do not seem to fall into a single size class. Curiously, loss of MSH3 results in a decrease not only in expansions, but also in the number of large contractions that are observed, with the decrease in these events being associated with a corresponding increase in the number of small contractions (Ref. 114). One model consistent with these observations is that large contractions represent a second possible outcome of the events that give rise to expansions, with contractions arising from DSBs that undergo some exonucleolytic cleavage prior to DSBR as illustrated in Figure 2.

Concluding remarks

Thus, MMR factors can both suppress and promote MSI. Which MSI pathway predominates is likely to depend on a variety of cell-type specific factors including the frequency of cell division and the relative levels of factors that promote or suppress each type of MSI. For example, while MMR is important for preventing genome-wide MSI in the colon, as evidenced by the high frequency of MSI-high colonic tumours in individuals with germline mutations in MMR proteins, MSI-high tumours originating in neurons are rare (Refs 115–117). The low level of MSI-high tumours in neurons may reflect in part the fact that neurons are non-dividing and thus likely to rarely generate the substrates for the MMR pathway. The high tumour incidence in the colon might reflect the consequences of exposure to dietary mutagens in rapidly dividing cells. In contrast, repeat expansion in neurons, particularly striatal neurons, occurs at high frequency in both mouse models of REDs and in REDs patients (Refs 118–120). The high frequency of STR expansion in neurons of REDs patients may reflect the high levels of factors such as MutS β , OGG1 and NEIL1 that promote expansion, and low levels of

proteins such as EXO1, that protect against them (Ref. 77). The high levels of transcription of the affected genes in neurons may also contribute to the incidence of these expansions, by increasing the opportunity for formation of the substrates upon which the expansion process acts.

The paradoxical effect of MMR proteins on MSI is particularly apparent at the *Fmr1* locus in embryonic stem cells from a mouse model of the FXDs. Consistent with the role of functional MMR components in generating expansions, a high rate of expansions is seen at this locus in MMR-proficient cells derived from these mice. For example, most alleles with ~280 repeat units gained an additional 19 repeats over a 52-day period in cells wildtype for MLH3 (Ref. 87). In contrast, in cells with a similar repeat number that lacked MLH3, the modal size of the allele actually decreased by one repeat over the same period. Thus, the same microsatellite expands rapidly in MMR-proficient cells and contracts more slowly in MMR-deficient ones. The fact that MMR-deficiency results in contractions at this locus rather than expansions serves to emphasise the fundamentally different events occurring at this locus.

In the case of MSI-high tumours, the expanded microsatellites themselves might be a source of vulnerability that could be exploited for therapeutic purposes. Since expansion results in a dependence on DNA helicases such as WRN (Ref. 68), it may be possible to selectively eliminate the cancer cells using a synthetic lethal approach that targets these enzymes (Ref. 121). In the case of many REDs, a growing body of evidence suggests that somatic expansion of the disease-associated STR significantly worsens the age at onset and/or disease severity (Refs 69, 70, 72, 73). Since most of these diseases are severely life-limiting and lack any effective treatment or cure, there is an interest in exploring ways to reduce this MSI in somatic cells. This approach has additional appeal in that any success in this regard would be relevant to multiple REDs. Of course, given the requirement of many MMR factors for protecting against cancer, targeting these factors to reduce expansion poses a challenge. However, MSH3 and MLH3 are not major players in MMR and may thus be acceptable targets particularly for those diseases with a high early mortality. For example, tail vein injection of a splice switching oligonucleotide that favours the production of an MLH3 isoform lacking the nuclease domain has already been shown to reduce expansion in some peripheral tissues of a mouse model of HD (Ref. 101). Since the absence of PMS1 is not associated with tumour susceptibility or any other obvious phenotype in mice (Ref. 49), depletion of this factor may be an even more attractive approach. While delivery of therapeutics to deep brain regions such as the striatum is not trivial, implanted intracerebroventricular devices have been used successfully for decades to deliver chemotherapeutic agents to treat central nervous system malignancies (Ref. 122). This experience could perhaps be leveraged for the ongoing delivery of MMR-targeting molecules to treat REDs.

Financial support. This work was made possible by funding from the Intramural Program of the National Institute of Diabetes, Kidney and Digestive Diseases to KU (DK057808).

Conflict of interest. None.

References

1. Subramanian S *et al.* (2003) SSRD: simple sequence repeats database of the human genome. *Comparative and Functional Genomics* **4**, 342–345.
2. Usdin K (2008) The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. *Genome Research* **18**, 1011–1019.
3. Marshall JN *et al.* (2021) Variable number tandem repeats – their emerging role in sickness and health. *Experimental Biology and Medicine* (Maywood) **246**, 1368–1376.

4. **Usdin K and Woodford KJ** (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Research* **23**, 4202–4209.
5. **Yudkin D et al.** (2014) Chromosome fragility and the abnormal replication of the FMR1 locus in fragile X syndrome. *Human Molecular Genetics* **23**, 2940–2952.
6. **Voineagu I et al.** (2009) Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility. *Nature Structural & Molecular Biology* **16**, 226–228.
7. **Krasilnikova MM and Mirkin SM** (2004) Replication stalling at Friedreich's ataxia (GAA)_n repeats in vivo. *Molecular and Cellular Biology* **24**, 2286–2295.
8. **Follonier C et al.** (2013) Friedreich's ataxia-associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nature Structural & Molecular Biology* **20**, 486–494.
9. **Thys RG and Wang YH** (2015) DNA Replication dynamics of the GGGGCC repeat of the C9orf72 gene. *Journal of Biological Chemistry* **290**, 28953–28962.
10. **Murat P, Guilbaud G and Sale JE** (2020) DNA polymerase stalling at structured DNA constrains the expansion of short tandem repeats. *Genome Biology* **21**, 209.
11. **Shastri N et al.** (2018) Genome-wide identification of structure-forming repeats as principal sites of fork collapse upon ATR inhibition. *Molecular Cell* **72**, 222–238, e11.
12. **Lokanga RA, Kumari D and Usdin K** (2021) Common threads: aphidicolin-inducible and folate-sensitive fragile sites in the human genome. *Frontiers in Genetics* **12**, 708860.
13. **Grindedal EM et al.** (2009) Germ-line mutations in mismatch repair genes associated with prostate cancer. *Cancer Epidemiology, Biomarkers & Prevention* **18**, 2460–2467.
14. **Grindedal EM et al.** (2009) High risk of endometrial cancer in colorectal cancer kindred is pathognomonic for MMR-mutation carriers. *Familial Cancer* **8**, 145–151.
15. **Peltomaki P et al.** (1993) Microsatellite instability is associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome. *Cancer Research* **53**, 5853–5855.
16. **Ionov Y et al.** (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363**, 558–561.
17. **Thibodeau SN, Bren G and Schaid D** (1993) Microsatellite instability in cancer of the proximal colon. *Science* **260**, 816–819.
18. **Wimmer K and Kratz CP** (2010) Constitutional mismatch repair-deficiency syndrome. *Haematologica* **95**, 699–701.
19. **Buerki N et al.** (2012) Evidence for breast cancer as an integral part of Lynch syndrome. *Genes Chromosomes & Cancer* **51**, 83–91.
20. **Win AK et al.** (2012) Risks of primary extracolonic cancers following colorectal cancer in lynch syndrome. *Journal of the National Cancer Institute* **104**, 1363–1372.
21. **The Huntington's Disease Collaborative Research Group** (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–983.
22. **Campuzano V et al.** (1996) Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271**, 1423–1427.
23. **Brook JD et al.** (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **69**, 385.
24. **Mahadevan M et al.** (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**, 1253–1255.
25. **Verkerk AJ et al.** (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914.
26. **Yu S et al.** (1991) Fragile X genotype characterized by an unstable region of DNA. *Science* **252**, 1179–1181.
27. **Allingham-Hawkins DJ et al.** (1999) Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study – preliminary data. *American Journal of Medical Genetics* **83**, 322–325.
28. **Brunberg JA et al.** (2002) Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *AJNR. American Journal of Neuroradiology* **23**, 1757–1766.
29. **Kononenko AV et al.** (2018) Mechanisms of genetic instability caused by (CGG)_n repeats in an experimental mammalian system. *Nature Structural & Molecular Biology* **25**, 669–676.
30. **Kunkel TA and Erie DA** (2015) Eukaryotic mismatch repair in relation to DNA replication. *Annual Review of Genetics* **49**, 291–313.
31. **Lujan SA et al.** (2012) Mismatch repair balances leading and lagging strand DNA replication fidelity. *PLoS Genetics* **8**, e1003016.
32. **Gadgil R et al.** (2017) Replication stalling and DNA microsatellite instability. *Biophysical Chemistry* **225**, 38–48.
33. **Ellegren H** (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* **5**, 435–445.
34. **Schlötterer C and Tautz D** (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**, 211–215.
35. **Viguera E, Canceill D and Ehrlich SD** (2001) Replication slippage involves DNA polymerase pausing and dissociation. *EMBO Journal* **20**, 2587–2595.
36. **Streisinger G et al.** (1966) Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harbor Symposia on Quantitative Biology* **31**, 77–84.
37. **Drummond JT et al.** (1995) Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science* **268**, 1909–1912.
38. **Palombo F et al.** (1995) GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science* **268**, 1912–1914.
39. **Warren JJ et al.** (2007) Structure of the human MutSalpha DNA lesion recognition complex. *Molecular Cell* **26**, 579–592.
40. **Gupta S, Gellert M and Yang W** (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nature Structural & Molecular Biology* **19**, 72–78.
41. **Habraken Y et al.** (1996) Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Current Biology* **6**, 1185–1187.
42. **Palombo F et al.** (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology* **6**, 1181–1184.
43. **Snowden T et al.** (2004) hMSH4-hMSH5 recognizes Holliday junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Molecular Cell* **15**, 437–451.
44. **Flores-Rozas H and Kolodner RD** (1998) The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proceedings of the National Academy of Sciences of the USA* **95**, 12404–9.
45. **Kadyrov FA et al.** (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297–308.
46. **Pluciennik A et al.** (2010) PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proceedings of the National Academy of Sciences of the USA* **107**, 16066–16071.
47. **Cannavo E et al.** (2020) Regulation of the MLH1-MLH3 endonuclease in meiosis. *Nature* **586**, 618–622.
48. **Roesner LM et al.** (2013) Stable expression of MutLγ in human cells reveals no specific response to mismatched DNA, but distinct recruitment to damage sites. *Journal of Cellular Biochemistry* **114**, 2405–2414.
49. **Prolla TA et al.** (1998) Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nature Genetics* **18**, 276–279.
50. **Räschle M et al.** (1999) Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. *Journal of Biological Chemistry* **274**, 32368–32375.
51. **Bradford KC et al.** (2020) Dynamic human MutSalpha-MutLalpha complexes compact mismatched DNA. *Proceedings of the National Academy of Sciences of the USA* **117**, 16302–16312.
52. **Hombauer H et al.** (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* **147**, 1040–1053.
53. **Manhart CM et al.** (2017) The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans. *PLoS Biology* **15**, e2001164.
54. **Wei K et al.** (2003) Inactivation of exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes & Development* **17**, 603–614.
55. **Genschel J, Bazemore LR and Modrich P** (2002) Human exonuclease I is required for 5' and 3' mismatch repair. *Journal of Biological Chemistry* **277**, 13302–13311.

56. **Kratz K et al.** (2021) FANCD2-associated nuclease 1 partially compensates for the lack of exonuclease 1 in mismatch repair. *Molecular and Cellular Biology* **41**, e0030321.
57. **Hile SE, Yan G and Eckert KA** (2000) Somatic mutation rates and specificities at TC/AG and GT/CA microsatellite sequences in nontumorigenic human lymphoblastoid cells. *Cancer Research* **60**, 1698–1703.
58. **Eckert KA, Yan G and Hile SE** (2002) Mutation rate and specificity analysis of tetranucleotide microsatellite DNA alleles in somatic human cells. *Molecular Carcinogenesis* **34**, 140–150.
59. **Hatch SB and Farber RA** (2004) Mutation rates in the complex microsatellite MYCL1 and related simple repeats in cultured human cells. *Mutation Research* **545**, 117–126.
60. **Christopher J et al.** (2019) Quantifying microsatellite mutation rates from intestinal stem cell dynamics in Msh2-deficient murine epithelium. *Genetics* **212**, 655–665.
61. **Eckert KA and Hile SE** (2009) Every microsatellite is different: intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. *Molecular Carcinogenesis* **48**, 379–388.
62. **Campregher C et al.** (2010) The nucleotide composition of microsatellites impacts both replication fidelity and mismatch repair in human colorectal cells. *Human Molecular Genetics* **19**, 2648–2657.
63. **Brinkmann B et al.** (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American Journal of Human Genetics* **62**, 1408–1415.
64. **Ellegren H** (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genetics* **24**, 400–402.
65. **Sun JX et al.** (2012) A direct characterization of human mutation based on microsatellites. *Nature Genetics* **44**, 1161–1165.
66. **Wierdl M, Dominska M and Petes TD** (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* **146**, 769–779.
67. **Twerdi CD, Boyer JC and Farber RA** (1999) Relative rates of insertion and deletion mutations in a microsatellite sequence in cultured cells. *Proceedings of the National Academy of Sciences of the USA* **96**, 2875–2879.
68. **van Wietmarschen N et al.** (2020) Repeat expansions confer WRN dependence in microsatellite-unstable cancers. *Nature* **586**, 292–298.
69. **Genetic Modifiers of Huntington's Disease Consortium** (2019) CAG repeat not polyglutamine length determines timing of Huntington's disease onset. *Cell* **178**, 887–900.
70. **Bettencourt C et al.** (2016) DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Annals of Neurology* **79**, 983–990.
71. **Lee JM et al.** (2017) A modifier of Huntington's disease onset at the MLH1 locus. *Human Molecular Genetics* **26**, 3859–3867.
72. **Flower M et al.** (2019) MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain* **142**, 1876–1886.
73. **McAllister B et al.** (2022) Exome sequencing of individuals with Huntington's disease implicates FAN1 nuclease activity in slowing CAG expansion and disease onset. *Nature Neuroscience* **25**, 446–457.
74. **Hwang YH et al.** (2022) Both cis and trans-acting genetic factors drive somatic instability in female carriers of the FMR1 premutation. *Scientific Reports* **12**, 10419.
75. **Zhao X et al.** (2021) Modifiers of somatic repeat instability in mouse models of Friedreich ataxia and the fragile X-related disorders: implications for the mechanism of somatic expansion in Huntington's disease. *Journal of Huntington's Disease* **10**, 149–163.
76. **Zhao XN and Usdin K** (2018) FAN1 protects against repeat expansions in a Fragile X mouse model. *DNA Repair (Amst)* **69**, 1–5.
77. **Zhao X et al.** (2018) MutLgamma promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. *PLoS Genetics* **14**, e1007719.
78. **Zhao X, Lu H and Usdin K** (2021) FAN1's protection against CGG repeat expansion requires its nuclease activity and is FANCD2-independent. *Nucleic Acids Research* **49**, 11643–11652.
79. **Goold R et al.** (2021) FAN1 controls mismatch repair complex assembly via MLH1 retention to stabilize CAG repeat expansion in Huntington's disease. *Cell Reports* **36**, 109649.
80. **Gazy I et al.** (2019) Double-strand break repair plays a role in repeat instability in a fragile X mouse model. *DNA Repair (Amst)* **74**, 63–69.
81. **Wright GEB et al.** (2019) Length of uninterrupted CAG, independent of polyglutamine size, results in increased somatic instability, hastening onset of Huntington disease. *American Journal of Human Genetics* **104**, 1116–1126.
82. **Findlay Black H et al.** (2020) Frequency of the loss of CAA interruption in the HTT CAG tract and implications for Huntington disease in the reduced penetrance range. *Genetics in Medicine* **22**, 2108–2113.
83. **Wright GEB et al.** (2020) Interrupting sequence variants and age of onset in Huntington's disease: clinical implications and emerging therapies. *The Lancet. Neurology* **19**, 930–939.
84. **Cumming SA et al.** (2018) De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *European Journal of Human Genetics* **26**, 1635–1647.
85. **Overend G et al.** (2019) Allele length of the DMPK CTG repeat is a predictor of progressive myotonic dystrophy type 1 phenotypes. *Human Molecular Genetics* **28**, 2245–2254.
86. **Morales F et al.** (2020) Longitudinal increases in somatic mosaicism of the expanded CTG repeat in myotonic dystrophy type 1 are associated with variation in age-at-onset. *Human Molecular Genetics* **29**, 2496–2507.
87. **Latham GJ et al.** (2014) The role of AGG interruptions in fragile X repeat expansions: a twenty-year perspective. *Frontiers in Genetics* **5**, 244.
88. **Nolin SL et al.** (2015) Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. *Genetics in Medicine* **17**, 358–364.
89. **Møllersen L et al.** (2010) Continuous and periodic expansion of CAG repeats in Huntington's disease R6/1 mice. *PLoS Genetics* **6**, e1001242.
90. **Zhao XN and Usdin K** (2018) Timing of expansion of fragile X premutation alleles during intergenerational transmission in a mouse model of the fragile X-related disorders. *Frontiers in Genetics* **9**, 314.
91. **Miller CJ et al.** (2020) All three mammalian MutL complexes are required for repeat expansion in a mouse cell model of the Fragile X-related disorders. *PLoS Genetics* **16**, e1008902.
92. **Yrigollen CM et al.** (2014) AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. *Journal of Neurodevelopmental Disorders* **6**, 24.
93. **Kovalenko M et al.** (2012) 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* **7**, e44273.
94. **Kennedy L and Shelbourne PF** (2000) Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Human Molecular Genetics* **9**, 2539–2544.
95. **Swami M et al.** (2009) Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Human Molecular Genetics* **18**, 3039–3047.
96. **Lokanga AR et al.** (2014) X inactivation plays a major role in the gender bias in somatic expansion in a mouse model of the fragile X-related disorders: implications for the mechanism of repeat expansion. *Human Molecular Genetics* **23**, 4985–4994.
97. **Kovtun IV et al.** (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* **447**, 447–452.
98. **Møllersen L et al.** (2012) Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. *Human Molecular Genetics* **21**, 4939–4947.
99. **Entezam A et al.** (2010) Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. *Human Mutation* **31**, 611–616.
100. **Jonson I et al.** (2013) Oxidative stress causes DNA triplet expansion in Huntington's disease mouse embryonic stem cells. *Stem Cell Research* **11**, 1264–1271.
101. **Møllersen L et al.** (2016) Effects of anthocyanins on CAG repeat instability and behaviour in Huntington's disease R6/1 mice. *PLoS Currents* **8**.
102. **Gomes-Pereira M and Monckton DG** (2020) Chronic exposure to cadmium and antioxidants does not affect the dynamics of expanded CAG*CTG trinucleotide repeats in a mouse cell culture system of unstable DNA. *Frontiers in Cellular Neuroscience* **14**, 606331.
103. **Hayward BE, Steinbach PJ and Usdin K** (2020) A point mutation in the nuclease domain of MLH3 eliminates repeat expansions in a mouse stem cell model of the Fragile X-related disorders. *Nucleic Acids Research* **48**, 7856–7863.
104. **Halabi A, Fuselier KTB and Grabczyk E** (2018) GAA*TTT repeat expansion in human cells is mediated by mismatch repair complex MutLgamma and depends upon the endonuclease domain in MLH3 isoform one. *Nucleic Acids Research* **46**, 4022–4032.
105. **Roy JCL et al.** (2021) Somatic CAG expansion in Huntington's disease is dependent on the MLH3 endonuclease domain, which can be excluded via splice redirection. *Nucleic Acids Research* **49**, 3907–3918.

106. **Kadyrova LY et al.** (2020) Human MutLgamma, the MLH1-MLH3 heterodimer, is an endonuclease that promotes DNA expansion. *Proceedings of the National Academy of Sciences of the USA* **117**, 3535–3542.
107. **Zakharovich K et al.** (2012) Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* **149**, 334–347.
108. **Mirkin SM** (2006) DNA structures, repeat expansions and human hereditary disorders. *Current Opinion in Structural Biology* **16**, 351–358.
109. **Bourn RL et al.** (2012) Pms2 suppresses large expansions of the (GAA.TTC)_n sequence in neuronal tissues. *PLoS ONE* **7**, e47085.
110. **Zhao XN et al.** (2016) A MutSbeta-dependent contribution of MutSalpha to repeat expansions in fragile X premutation mice? *PLoS Genetics* **12**, e1006190.
111. **Foiry L et al.** (2006) Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. *Human Genetics* **119**, 520–526.
112. **Halabi A et al.** (2012) DNA mismatch repair complex MutSbeta promotes GAA.TTC repeat expansion in human cells. *Journal of Biological Chemistry* **287**, 29958–29967.
113. **Gomes-Pereira M et al.** (2004) Pms2 is a genetic enhancer of trinucleotide CAG/CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Human Molecular Genetics* **13**, 1815–1825.
114. **Zhao XN et al.** (2015) Mutsbeta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Human Molecular Genetics* **24**, 7087–7096.
115. **Aarnio M et al.** (1999) Cancer risk in mutation carriers of DNA-mismatch-repair genes. *International Journal of Cancer* **81**, 214–218.
116. **Therkildsen C et al.** (2015) Glioblastomas, astrocytomas and oligodendrogliomas linked to Lynch syndrome. *European Journal of Neurology* **22**, 717–724.
117. **Guerrini-Rousseau L et al.** (2019) Constitutional mismatch repair deficiency-associated brain tumors: report from the European C4CMMRD consortium. *Neurooncology Advances* **1**, vdz033.
118. **Shelbourne PF et al.** (2007) Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Human Molecular Genetics* **16**, 1133–1142.
119. **Lee JM et al.** (2011) Quantification of age-dependent somatic CAG repeat instability in Hdh CAG knock-in mice reveals different expansion dynamics in striatum and liver. *PLoS ONE* **6**, e23647.
120. **Gonitel R et al.** (2008) DNA instability in postmitotic neurons. *Proceedings of the National Academy of Sciences of the USA* **105**, 3467–3472.
121. **Chan EM et al.** (2019) WRN helicase is a synthetic lethal target in microsatellite unstable cancers. *Nature* **568**, 551–556.
122. **Triarico S et al.** (2019) Improving the brain delivery of chemotherapeutic drugs in childhood brain tumors. *Cancers (Basel)* **11**, 824.