

REVIEW

Open Access

# Instability of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> trinucleotide repeats and DNA synthesis

Guoqi Liu\* and Michael Leffak\*

## Abstract

Expansion of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> trinucleotide repeat (TNR) microsatellite sequences is the cause of more than a dozen human neurodegenerative diseases. (CTG)<sub>n</sub> and (CAG)<sub>n</sub> repeats form imperfectly base paired hairpins that tend to expand in vivo in a length-dependent manner. Yeast, mouse and human models confirm that (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability increases with repeat number, and implicate both DNA replication and DNA damage response mechanisms in (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNR expansion and contraction. Mutation and knockdown models that abrogate the expression of individual genes might also mask more subtle, cumulative effects of multiple additional pathways on (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability in whole animals. The identification of second site genetic modifiers may help to explain the variability of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNR instability patterns between tissues and individuals, and offer opportunities for prognosis and treatment.

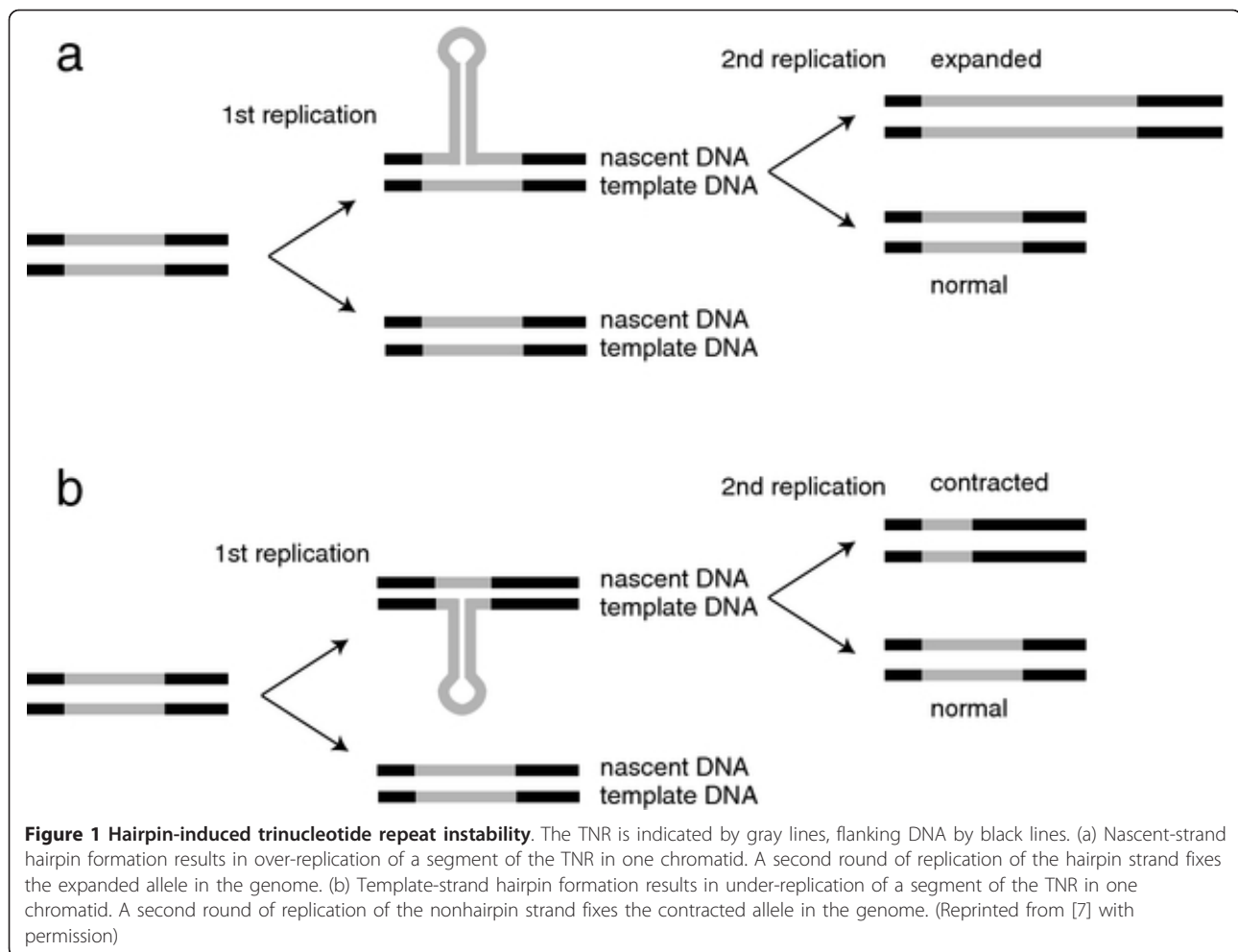
## Introduction

Expansion of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> trinucleotide repeat (TNR) sequences at distinct chromosomal loci is the mutation common to multiple neurological diseases including myotonic dystrophy type 1 (DM1), Huntington disease (HD), Huntington disease-like 2 (HDL2), dentatorubral-pallidoluysian atrophy (DRPLA), spinal and bulbar muscular atrophy (SBMA), and several forms of spinocerebellar ataxia (SCA). The polyglutamine diseases HD, DRPLA, SBMA, and SCA1, 3, 6, 7, 17 result from increases of (CAG)<sub>n</sub> repeats in the coding (nontemplate) strand for mRNA synthesis of the cognate genes ((CAG)<sub>n</sub> in RNA) to produce mutant polyglutamine proteins with toxic gain-of-function [1]. In contrast, (CTG)<sub>n</sub>•(CAG)<sub>n</sub> expansion at the DMPK 3' UTR alters the chromatin structure of the region, downregulates transcription of the locus and, as at the JPH3 gene produce poly-(CUG) pre-mRNAs respectively in DM1 and HDL2 patients that sequester the MBNL (CUG) binding proteins, leading to *trans*-dominant interference with the normal splicing of multiple RNAs. Finally, bidirectional transcription at the SCA8 locus can result in expression of both a polyglutamine protein and a (CUG)<sub>n</sub> expansion transcript, which may represent a toxic gain-of-function at both the protein and RNA levels.

Trinucleotide repeat expansion requires DNA synthesis, either during DNA replication or repair. The effects of replication origin proximity, replication polarity, and replication inhibition support replication-based models of TNR instability in mitotic cells [2-9]. Hairpin formation by DNA polymerase slippage is a likely mechanism for changes in TNR repeat length [10-12]. Hairpin structure formation by DNA polymerase slippage at (CTG)<sub>n</sub>•(CAG)<sub>n</sub> sequences has been well documented in vitro [13,14] and can result in either insertion or deletion mutations. However, hairpins have also been postulated to arise during replication fork reversal and postreplication repair [2,15,16], Okazaki fragment maturation [17-19], base excision repair [20], nucleotide excision repair [21-26] or repair of structures induced by R-loop formation during transcription [25,27]. Current models of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability during replication or repair envision that hairpin formation on the newly synthesized DNA strand leads to TNR expansion if the hairpin is sufficiently long-lived to serve as template in a subsequent round of replication. Conversely, stable hairpin formation in the leading or lagging template strand would lead to contraction of the repeat in the next round of replication (Figure 1).

The salient observation that TNR instability in humans and mice can occur in postmitotic cells argues that repair mechanisms, instead of replication origin-dependent mitotic DNA replication, are involved in TNR instability in these tissues [2,5,28-30]. In this vein,

\* Correspondence: guoqi.liu@yahoo.com; michael.leffak@wright.edu  
Department of Biochemistry and Molecular Biology, Boonshoft School of Medicine, Wright State University, Dayton, OH 45435, USA



it has been proposed that the process of transcription stimulates TNR instability due to the formation of hairpin or other non-B DNA structures in the single stranded nontemplate DNA, or in the template strand upon RNA displacement. These structures may be targets for DNA repair processes such as transcription-coupled repair, nucleotide excision repair, mismatch repair, or double-stranded DNA break repair [24,27,31].

Following extensive linkage analysis in myotonic dystrophy families [32-34], in 1992 several laboratories reported that expansion of the  $(CTG)_n \cdot (CAG)_n$  repeat region in the 3' untranslated region of the dystrophin myotonia protein kinase gene was highly correlated with the occurrence of congenital DM [35-37]. Strong correlations also exist between  $(CTG)_n \cdot (CAG)_n$  repeat length and the occurrence of Huntington disease [38,39], although second site modifier genes and epigenetic mechanisms play a significant role in the appearance of HD symptoms. In general, unaffected individuals display fewer than 30  $(CTG)_n \cdot (CAG)_n$  repeats at the DM1 or HD locus. Trinucleotide repeat (TNR) tracts in

the range of 30-40 repeats are termed premutation alleles (DM1) or intermediate alleles of incomplete penetrance (HD), while TNRs of 42 or more repeats have been associated with complete penetrance of HD [40] and increased expansion frequency during intergenerational transfer or somatic development in DM1 families [2]. The phenomenon of 'genetic anticipation' is a hallmark of the  $(CTG)_n \cdot (CAG)_n$  TNR instability disorders, in which an increase in the number of microsatellite repeats is correlated with an earlier age of onset and heightened severity of the disease in successive generations. Genetic anticipation reflects the bias towards expansion over contraction of long  $(CTG)_n \cdot (CAG)_n$  tracts, and may be explained by the greater tendency of extended repeats to adopt non-B form DNA structures prone to progressive expansion.

$(CTG)_n \cdot (CAG)_n$  expansion can have pathological effects on local chromatin structure and gene expression, as well as dominant negative effects on RNA metabolism and protein function [41-44]. This review will focus primarily on the structure and instability of  $(CTG)$

$n \cdot (\text{CAG})_n$  trinucleotide repeat sequences in eucaryotic cells. For further background on the metabolism of  $(\text{CTG})_n \cdot (\text{CAG})_n$  sequences in bacterial cells, the reader is referred to several excellent research articles and reviews [5,6,45-51].

#### **$(\text{CTG})_n \cdot (\text{CAG})_n$ hairpins in vitro**

NMR, melting and chemical modification analyses confirm that both  $(\text{CTG})_n$  and  $(\text{CAG})_n$  oligonucleotides as short as 6-10 repeats can form stable hairpin structures with mismatched base pairs [21,52-55]. Although short  $(\text{CTG})_n$  or  $(\text{CAG})_n$  hairpins represent stable structures, they rapidly convert to duplex DNA in the presence of the complementary oligonucleotide through loop-loop and stem-stem interactions without prior denaturation [52]. Considered with the greater thermodynamic stability of duplex vs. cruciform DNA, and the inhibition of cruciform formation by single base mismatches [56,57], these observations suggest that unless otherwise stabilized e.g. by protein binding or high negative superhelicity [2,58], apposed hairpin structures formed in vivo by transcription or replication fork regression to "chicken foot" structures would resolve to duplex DNA, and disfavor TNR expansions. Notably,  $(\text{CTG})_{25}$  or  $(\text{CAG})_{25}$  hairpins form duplex DNA approximately 5-fold more slowly than  $(\text{CTG})_{10}$  or  $(\text{CAG})_{10}$  hairpins in the presence of their complementary strands, despite similar thermal stabilities of the hairpins [55]. Thus, longer hairpins may also have longer half-lives in vivo.

Compared to nonrepetitive palindromic sequences, which would require half of the palindrome to become single stranded prior to hairpin formation, the free energy required to nucleate short hairpin formation may be provided by negative superhelicity [2,58]. The greater number of hairpin configurations in a repetitive TNR would be expected to increase the entropy of hairpin formation and decrease the free energy. Thus, in addition to the slower dissolution of longer hairpins by their complementary sequences, hairpin formation may occur more rapidly in longer repeats, competing against SSB binding.

To assess the mechanism of hairpin instability, Panigrahi et al. used in vitro replication of  $(\text{CTG})_{79} \cdot (\text{CAG})_{79}$  repeats driven by the SV40 T antigen (T-ag) helicase. The remaining enzymatic machinery of DNA synthesis is endogenous to the host cell. Plasmids in which the  $(\text{CAG})_n$  sequence was the lagging strand template showed an expansion bias, while plasmids containing the opposite TNR replication orientation ( $(\text{CTG})_n$  in the lagging strand template) displayed a preference for contraction [59]. To the extent that the replication fork driven by the strong T-ag helicase mimics the activity and interactions of the cellular Cdc45/MCM2-7/GINS replicative helicase [60,61] with replication fork stabilizing proteins, the effect of TNR orientation relative to the

replication origin on  $(\text{CTG})_n \cdot (\text{CAG})_n$  stability imply that unstable TNR structures can be processed to expansions or contractions depending on DNA replication polarity.

The question of why a contraction bias is observed in rapidly dividing eucaryotic cells [16] was addressed by Delagoutte et al. using a primer extension model of  $(\text{CTG})_n \cdot (\text{CAG})_n$  replication, in which replication by T4 DNA polymerase through short  $(\text{CAG})_n$  or  $(\text{CTG})_n$  TNRs was inhibited relative to polymerization through non-structure forming repeats.  $(\text{CAG})_n$  repeats blocked replication more efficiently than  $(\text{CTG})_n$  repeats, and this difference was eliminated by the addition of *E. coli* or T4 single strand binding SSB proteins [62]. Based on the preferential binding of SSB to lagging strand template DNA and the more efficient blockage of polymerization by the  $(\text{CAG})_n$  vs.  $(\text{CTG})_n$  template, the authors proposed a 'template-push' model in which the contraction bias of TNRs with  $(\text{CTG})_n$  in the lagging strand template is not the result of lagging strand structure formation while single stranded, but is the result of extrusion of the leading strand  $(\text{CAG})_n$  template and replication across the abasic bottom of the hairpin in order to maintain contact between DNA polymerase and replicative helicase [62]. Alternatively, transient release of the leading strand hairpin template from the stalled polymerase could allow hairpin slippage or migration in the 5' -> 3' direction away from the replication fork [13,63,64], and reestablishment of a functional primer-template junction.

Annealing of single stranded plasmid DNA to complementary strands containing excess  $(\text{CTG})_n$  or  $(\text{CAG})_n$  sequences, which formed hairpins as large as 25 repeats, yielded products that showed accurate repair in human cell extracts [56,65-68]. The requirement for PCNA, and the nick-dependence of accurate repair suggested that mismatch repair proteins (MMR) that function during replication in vivo might play a role in hairpin resolution in vitro. Indeed, in one study, the repair of short  $(\text{CTG})_{1-3}$  slip out structures was reported to increase with increasing concentrations of MutS $\beta$  (Msh2/Msh3) in cell extracts [69]. Nevertheless, in contrast to the apparent requirement for MMR proteins for  $(\text{CTG})_n \cdot (\text{CAG})_n$  expansions in transgenic mice [70-72], MMR proteins, or the nucleotide excision repair (NER) protein XPG, were not essential for repair of longer  $(\text{CTG})_{20-25}$  hairpins in cell extracts [66,67,69]. It remains possible that there are alternative pathways for hairpin removal. Additionally, the formation of stable hairpins in advance in these assays may have bypassed the contribution of MMR, NER or other pathways to hairpin repair.

#### **Yeast models of $(\text{CTG})_n \cdot (\text{CAG})_n$ instability**

Numerous genetic analyses have been performed in *S. cerevisiae* to characterize the effect of  $(\text{CTG})_n \cdot (\text{CAG})_n$

sequences on DNA replication and chromosome fragility, and to identify proteins that affect (CTG)<sub>n</sub>•(CAG)<sub>n</sub> trinucleotide repeat stability. Between these studies there are some disparities that are likely due to differences in the repeat sequence ((CTG)<sub>n</sub>•(CAG)<sub>n</sub>, (GAA)<sub>n</sub>•(TTC)<sub>n</sub>, (CGG)<sub>n</sub>•(CCG)<sub>n</sub>), its environment (plasmid vs. chromosome; leading vs. lagging strand replication polarity), number of repeats in the microsatellite tract, genetic background, and the sensitivity of the assay to small changes in repeat length [73-82]. Thus the frequency of expansions was approximately 500-fold greater when (CTG)<sub>25</sub> was in the lagging strand template than when (CAG)<sub>25</sub> was in the lagging strand template [17]. While the frequencies of (CTG)<sub>25</sub> or (CAG)<sub>25</sub> expansions, and (CTG)<sub>50</sub> or (CAG)<sub>50</sub> contractions, were unaffected in *msh2* mutants [17,83], (CTG)<sub>13</sub> expansion was stimulated by mutation of postreplication repair genes *rad18* (hRAD18, binding partner of hUBE2A/B), *rad5* (hSMARCA3), and PRR-specific alleles of *pol30* (hPCNA) [74].

When analyzed by two-dimensional gel electrophoresis (CTG)<sub>80</sub>•(CAG)<sub>80</sub> sequences showed only modest effects on replication fork progress, irrespective of replication polarity. In contrast, (CGG)<sub>40</sub>•(CCG)<sub>40</sub> repeats imposed strong blocks to fork progression [75]. Surprisingly, in an assay that used reversion to 5-fluoroorotic acid resistance (FOA<sup>R</sup>) to quantitate TNR expansions, comparable rates of repeat instability were found for (CGG)<sub>25</sub> or (CTG)<sub>25</sub> (lagging strand template) TNRs. In the presence of the *rfc1-1* mutation, which blocks PCNA loading and lagging strand Okazaki fragment synthesis, the expansion rates of (CGG)<sub>25</sub> and (CTG)<sub>25</sub> increased ~40-50 fold and ~2-3 fold respectively. One interpretation of this result is that inhibition of lagging strand synthesis can promote expansions in the leading strand nascent DNA. However, since (CGG)<sub>n</sub> and (CTG)<sub>n</sub> repeats in the lagging strand template characteristically show a strong bias towards contraction, this assay may not have revealed the full relationship between replication stalling and TNR instability. In a similar assay, expansion of (CAG)<sub>25</sub>•(CTG)<sub>25</sub> was increased ~100 fold when (CAG) was in the lagging vs. leading strand template, and a *ra27Δ* mutant in the Okazaki flap endonuclease (hFEN-1) enhanced (CTG)<sub>n</sub>•(CAG)<sub>n</sub> expansion an additional 100 fold, irrespective of replication orientation [84].

Bhattacharya and Lahue [85] reported that (CTG)<sub>13</sub> (lagging strand template) expansion was markedly (~40 fold) increased in *srs2* helicase mutants, while (CTG)<sub>25</sub> expansion was increased ~5 fold in the same cells, and these rates were minimally affected by mutation of the RecQ helicase *sgs1* or either of the homologous recombination proteins *rad51* or *rad52*, arguing against unequal sister chromatid exchange as a mechanism of expansion, consistent with the absence of exchange of markers

flanking expanded alleles in human patients [86,87]. These results differed from those of Kerrest et al. who reported that the fragility of yeast artificial chromosomes (YACs) containing longer (CTG)<sub>70</sub>•(CAG)<sub>70</sub> TNRs, which are above the expansion threshold, increased significantly in *sgs1Δ* or *srs2Δ* helicase mutants. Deletion of the homologous recombination protein genes mitigated the effect of the *srs2Δ* mutation in either orientation, and decreased the effect of the *sgs1Δ* mutation in the (CTG)<sub>70</sub> orientation, but exacerbated the effect of the *sgs1Δ* mutation in the (CAG)<sub>70</sub> orientation [15]. A simple relationship between YAC fragility, TNR length and replication polarity was difficult to ascertain for these mutants, leading the authors to suggest that multiple pathways coexist that involve Srs2, Sgs1, Rad51 and Rad52 in the repair of replication fork damage due to hairpin forming sequences of different lengths and orientations.

In a screen for mutants that affected (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability, disruption of the replication fork stabilization complex protein genes *Mrc1*, *Tof1* or *Csm3* selectively enhanced contractions of a (CAG)<sub>20</sub>-URA3 (lagging strand template) reporter, independent of replication checkpoint and DNA damage checkpoint factors [76], and control experiments demonstrated that mutation of the same fork stabilization complex proteins did not affect the stability of a non-structure forming (CTA)<sub>n</sub> repeat. In contrast, mutation of the fork stabilization complex proteins or the DNA damage checkpoint proteins *Ddc1*, *Rad9*, *Rad17*, *Mec1*, *Ddc2*, *Rad24*, *Mec3*, *Rad53* or *Chk1* led to increased expansion of a (CAG)<sub>13</sub>-URA3 reporter. These results suggest that *Mrc1*, *Tof1* and *Csm3* may maintain TNR length through coupling of the DNA polymerase and replicative helicase to prevent the formation of hairpins, whereas the DNA damage checkpoint is involved in stabilization of the replisome after the formation of hairpin structures.

Assays using longer TNRs (85-155 repeats) in *mrc1*, *rad9*, *mec1*, *ddc2*, *rad17*, *rad24*, *chk1*, or *rad53* mutant strains found elevated chromosome breakage due to expanded (CAG)<sub>n</sub>•(CTG)<sub>n</sub> tracts and increased instability (primarily contractions) of a (CAG)<sub>n</sub> (lagging strand template) reporter [77,78]. The inherently greater instability of long (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeats in wild type strains may have masked the effects of some checkpoint mutants, nevertheless, these studies indicate that distinct protein complexes respond to forms of DNA replicative stress that differ in size or geometry, and underscore the correlation between noncanonical (CTG)<sub>n</sub>•(CAG)<sub>n</sub> structures, checkpoint activation and chromosome breakage [88,89]. Indeed, Sundararajan et al. have recently shown that long (CTG)<sub>n</sub>•(CAG)<sub>n</sub> tracts induce chromosomal double strand breaks in yeast, and that the *Mre11*/*Rad50*/*Xrs2* complex is necessary for blocking chromosome fragility and inhibiting (CAG)<sub>70</sub> TNR instability (expansion and



contraction) by both homologous recombination and NHEJ pathways [90].

The homologous recombination protein Rad52 was also required to protect the  $(CAG)_{70}$ -URA3 reporter from length instability and chromosome breakage in the presence of mutations in the alternative clamp loader Ctf18-Dcc1-Ctf8-RFC (Ctf18-RFC) [91]. Previously thought to promote PCNA loading and unloading during replication fork navigation through sister chromatid cohesion (SCC) complexes [92,93], Gellon et al. showed that Ctf18-RFC is required for TNR stability independent of its role in SCC, in parallel to a pathway involving the Mrc1 protein which couples the leading strand polymerase  $\epsilon$  and the replicative helicase at the replication fork, and acts in signaling during the intra-S phase checkpoint and the DNA damage response [94-96]. On the replication fork lagging strand, Ctf4 collaborates with MCM10 to link DNA polymerase  $\alpha$  to the MCM2-7 helicase [97-99]. Like the *mrc1* mutant, a *ctf4* deletion mutant is associated with chromosomal instability, *ctf4 rad52* double mutants grow poorly and produce a high percentage of inviable cells [100], and *ctf4 mrc1* mutants are inviable [101].

Taken together these studies in yeast suggest a cellular fail-safe strategy of overlapping pathways to (i) prevent the formation of stable hairpin structures by maintaining the rate of replisome movement and coupling of leading and lagging strand polymerases to the replicative helicase, (ii) restore hairpin structures to duplex DNA by repair helicases, and (iii) recruit postreplication repair machinery to excise hairpins.

#### Mouse models of $(CTG)_n \cdot (CAG)_n$ instability

Murine models of several trinucleotide expansion diseases including Huntington disease, myotonic dystrophy type 1, Fragile X syndrome, and Friedrich's ataxia have been generated by random integration of pathological length repeat tracts or knock-in at homologous genetic sites, and have reproduced many, though not all, phenotypes of the associated disease. These models typically show tissue-specific, expansion-biased patterns of instability similar to those in humans, including expansions in germ cells, early embryos and adults. Although intergenerational  $(CTG)_n \cdot (CAG)_n$  expansion is typically smaller in transgenic mice than humans [16,102], some recent studies have reported relatively large expansions during parent-to-offspring transmission [103,104]. Among the *cis*-acting modulators of TNR instability in murine systems are the sequence and length of the TNR [105], the presence of human flanking DNA [27,106-108], the chromosomal integration site [109,110], chromatin structure [41,43,111-115], and replication polarity [4].

Relevant to studies of the relationship between the DNA replication and  $(CTG)_n \cdot (CAG)_n$  stability is a

comparison of origin activity at the DMPK locus in human cells and transgenic mice [4]. In this work, two origins were mapped upstream and downstream of the DMPK  $(CTG)_n \cdot (CAG)_n$  repeat in both control and DM1 human fibroblasts. Transgenic mice bearing a single copy of a ~45-kb genomic region of the expanded DM1 locus containing  $(CTG)_{>300} \cdot (CAG)_{>300}$  repeats showed high levels of intergenerational and somatic repeat instability [116]. The transcriptional activity of the DM1 locus and tissue-specific patterns of instability were similar to those of DM1 individuals. Unlike in humans, however, when origin activity (abundance of nascent DNA) was quantitated over the ~45-kb human DM1 transgene from pancreatic cells of mice bearing either  $>300$  (DM328) or  $20$  (DM20) [117] repeats, neither the upstream nor the downstream origin was active in DM20 mice, and only the upstream origin was inactive in DM328 mice, [4]. Thus, the nuclear environment of the transgene may also modulate its replication origin activity and downstream effects on TNR stability.

Conversely, pathological length  $(CTG)_n \cdot (CAG)_n$  transgenes could induce local heterochromatinization and position effect variegation (PEV) upon integration. With overexpression of the heterochromatin organizing protein HP1 $\beta$ , PEV increased only in transgenes containing the TNRs [43]. These data suggest that the integration site and the transgene may each effect biological pressure for or against integration of a DNA fragment at a particular genomic site. While the influences of chromosome environment are manifold, *cis*-effects of the integration site on TNR instability are generally diminished with increasing length of the microsatellite repeat tract and the human flanking DNA [27,107].

Transcription is a possible *cis*-acting modifier that could lead to tissue-specific TNR instability, although the constitutive expression of the associated disease genes in humans argues against this model, and no correlation was observed between instability and stable mRNA levels in DM1 [118], HD [119] or SCA7  $(CTG)_{92} \cdot (CAG)_{92}$  transgenic mice [115]. Nevertheless, secondary attributes of transcription, e.g. DNA supercoiling, histone modification, DNA repair induced by the formation of RNA-DNA hybrid loops [25,120,121] may indirectly account for the activation of ATR or ATM pathways [122-124], and transcription induced  $(CTG)_n \cdot (CAG)_n$  contraction has been reported in non-murine systems [26,125-129].

Specific *trans*-acting factors that have been implicated in  $(CTG)_n \cdot (CAG)_n$  instability based on crosses between transgenic TNR mice and mice defective in DNA mismatch repair or base excision repair [20,71,72,102,130,131]. In bacteria, MMR relies primarily on three protein complexes MutS, MutL and MutH [132]. The MutS dimer recognizes the mismatch and enlists a MutL dimer that then recruits the MutH endonuclease to initiate nick-directed repair.

In eukaryotes there are at least six homologs to the MutS and MutL proteins [133,134]. The major MutS homologue is MSH2, which can heterodimerize with either MSH6 to form MutS $\alpha$  which binds to single base mismatches, or with MSH3 to form MutS $\beta$  that recognizes short insertion/deletion loops. The PMS2/MLH1 (MutL $\alpha$ ) heterodimer interacts with mismatches recognized by MutS $\alpha$  or MutS $\beta$  to trigger downstream excision and resynthesis reactions.

The murine mismatch repair genes MSH2 and MSH3 are essential for germinal and somatic expansion, and PMS2 is required for somatic instability, of long (> 84 repeat) (CTG) $_n$ •(CAG) $_n$  TNRs in transgenic mice [16,70-72,102,116,130,131,135,136]. In vitro, MutS $\beta$  binding to (CAG) $_n$  hairpins nominally reduced its ATPase activity, suggesting that MutS $\beta$  might mask hairpin structures from repair [137,138]. However, a later study reported that the ATPase activity of MutS $\beta$  bound to (CAG) $_n$  hairpins was similar to that of the enzyme bound to nonhairpin duplex DNA, and that (CAG) $_n$  hairpin binding of MutS $\beta$  did not change its catalytic efficiency (kcat/Km) [139]. Thus, the precise mechanism by which the MMR system is involved in TNR instability in vivo remains unresolved.

When transgenic mice carrying the HD (CAG) $_n$  repeat were crossed with mice lacking the base excision repair (BER) glycosylase OGG1 (7,8-dihydro-8-oxoguanine DNA glycosylase), which is responsible for the removal of 7,8-dihydro-8-oxoguanine (8-oxoG) (the most common oxidized base in DNA), age-dependent somatic expansion was largely suppressed [20]. In an in vitro model of base excision repair, incision of 8-oxo-G within a (CAG) $_n$  tract by APE1 and extension by the major BER polymerase, pol $\beta$ , resulted in expansion of the repeat tract [20,140], leading to the hypothesis of a toxic oxidation cycle in which hairpin loops form during long-patch repair of bases damaged by reactive oxygen species. The hairpin may be protected from the endonuclease activity of FEN1 by MutS $\beta$  [141], or FEN1 may promote the ligation of hairpin-containing flaps [140]. Repeated cycles of oxidation, repair and expansion would promote progressive age-dependent expansion. The presence of additional glycosylases that can act on 8-oxo-G (and other oxidized bases) implies a unique function for OGG1 in (CAG) $_n$  expansion during BER, which is not yet understood [142]. Similar experiments in which DM328 mice ((CTG) $_{>300}$ •(CAG) $_{>300}$  repeats) were crossed with mice deficient in Rad52, Rad54 or DNA-PKcs showed little effect on TNR stability, arguing that homologous recombination or nonhomologous end joining are not involved in expansion or contraction [135].

A difference between yeast and murine systems is the reported absence of an effect of Fen1 loss on TNR instability in DM1 knock-in transgenic mice [143]. Fen1 knockdown (80-90%) did not affect the stability of the

HD (CAG) $_{27}$ •(CTG) $_{27}$  repeats in human cells, but Fen1 haploinsufficiency did induce expansion of transgenic (CAG) $_{120}$ •(CTG) $_{120}$  repeats [144]. Constitutively low levels of Fen1 have also been implicated in inducing (CAG) $_n$ •(CTG) $_n$  instability in the striatum vs. cerebellum of HD mice [145].

The similarity of tissue-specific patterns of somatic mosaicism in multiple mouse models also suggest the influence of additional tissue-specific factors affecting TNR stability [104,110,118,119,146]. Further, it has been proposed that sex-specific trans-acting factors are responsible for differences in intergenerational instability between murine and human systems [103]. A recent genomic study compared the phenotype of tissue-specific patterns of (CAG) $_n$ •(CTG) $_n$  instability of *Hdh*<sup>Q111</sup> (HD homolog) transgenic mice with microarray analysis of gene expression [147]. The collective expression signature of a group of 150 genes was highly correlated with tissue instability, although no single gene expression pattern was absolutely predictive of instability. Instability indices were highest in nondividing striatum (highly affected in HD) and liver cells, and lowest in testis and umbilical cord. Comparison of *Hdh*<sup>Q111/111</sup> and *Hdh*<sup>+/+</sup> littermates showed that the instability of normal or mutant striata was significantly higher than the instability index of cerebellum. The authors concluded that mutant and wild type striata have similar tendencies towards TNR expansion, but the HD (CAG) $_n$ •(CTG) $_n$  microsatellite does not expand in normal striatum because it is not of sufficient length to be susceptible to additional processes involved in expansion. Possibly, transient or short-lived fluctuations in protein function or DNA structure occur frequently in specific tissues to increase the susceptibility of long TNRs to expansion.

In contrast to previous reports that MMR and BER proteins contribute to (CTG) $_n$ •(CAG) $_n$  expansion in mice, expression levels of DNA repair genes including MSH2, MSH3, and OGG1 did not correlate with the tissue specificity of somatic instability [147]. To address the caveat that steady state RNA levels may not reflect changes in protein abundance, the authors confirmed by immunoblot that Cbp and MSH2 protein levels were indistinguishable in *Hdh*<sup>+/+</sup> and *Hdh*<sup>Q111/+</sup> mice. However, further inspection of the data revealed that 63 of 74 genes whose downregulation showed weak-to-medium range Pearson coefficient correlation to TNR instability are involved in DNA metabolism. The authors concluded that pathways including cell cycle, metabolism and neurotransmission act in combination to generate tissue-specific patterns of instability, and that multiple tissue factors reflect the level of somatic instability in different tissues. Components of any of these pathways may represent second site genetic modifiers that contribute to the tissue- and cell type-specific

variation of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNR instability observed in mice and humans.

#### Human models of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability

During human and mouse development, DM1 tracts tend to expand in premeiotic spermatogonia, and large alleles subsequently contract during later stages of spermatogenesis and early in male development. In females, large expansions can be observed in nondividing oocytes, and full mutations are inherited almost exclusively from the mother. Thus, it has been proposed that two mechanisms of instability apply to (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeats: as in oocytes, expansions occur by DNA repair, while contractions characteristic of male development are the result of DNA replication [142]. In both male and female DM1 patients (and transgenic mouse models), (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNR tracts also show a significant level of somatic instability that increases with age in a tissue-specific manner. In possible support of a dual mechanism model for expansions and contractions, DM328 transgenic mice made deficient in DNA ligase I displayed reduced (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability upon maternal transmission, but showed no effect on paternal transmission or somatic instability [148]. Although studies such as these are valuable in identifying candidate genes affecting (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability, non-human model systems arguably do not recapitulate all aspects of microsatellite expansion disease in human cells due to differences in chromatin structure, cell division and DNA replication rates, and cell type. Hence, several investigators have turned to analyses of patient-derived cells, human embryonic stem (hES) cells [149-151], and other human cell model systems [7].

Recent PCR and immunoblot studies reported that MMR (MSH2, MSH3, MSH6) gene expression and protein levels of VUB03\_DM1 and VUB19\_DM1 hES cells were as high as in MMR proficient HeLa cells and stable during culture in the undifferentiated state, when (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeat length increased significantly [151,152]. Following differentiation to osteoblast progenitor-like cells, sharp decreases in MSH2, MSH3 and MSH6 levels were correlated with stabilization of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeat lengths of the VUB03\_DM1 and VUB19\_DM1 hES cells [152]. These results imply that either the reduction in MMR protein expression, the decrease in cell proliferation during hES cell differentiation, or other *trans*-acting factors, may be related to (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNR stabilization.

In a human fibrosarcoma model that scores contraction of an intronic (CTG)<sub>95</sub>•(CAG)<sub>95</sub> repeat by cell survival under HPRT<sup>+</sup> selection (HAT medium), ~25-fold induction of transcription of randomly integrated HPRT cassettes increased contraction ~15-fold, to roughly 0.001% of cells. Transcription-induced contraction frequencies accumulated at the same rate in proliferating and confluent cells that differed by 10-fold in rates of

cell division. siRNA knockdown of proteins involved in mismatch repair (MSH2, MSH3), and transcription-coupled nucleotide excision repair (CSB, ERCC1, XPA, XPG, TFIIIS, BRCA1, BARD1) decreased the frequency of contractions 2- to 3-fold [25,26,126,153], indicating that these pathways likely play a role in TNR expansion in postmitotic cells. Supporting a role for BER in somatic alteration of (CAG)<sub>n</sub>•(CTG)<sub>n</sub> repeat length, tissue-specific decreases of (CAG)<sub>n</sub>•(CTG)<sub>n</sub> instability in the striatum, cerebral cortex and hippocampus of *Xpa*<sup>-/-</sup> SCA1 mice have recently been reported [154]. Nevertheless, a parsimonious mechanism for transcription-induced instability that addresses both supporting and conflicting evidence is not yet available [25].

The effect of DNA replication on the stability of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> sequences has been studied in SV40 origin plasmids replicating in COS-1 cells [155] or T-ag supplemented HeLa cell extracts [59]. While these systems do not duplicate the chromatin structure of genomic DNA, and replication does not utilize components of the ORC-dependent replisome that interact with the cellular DNA damage signaling and repair machinery, (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability in these plasmids was sensitive to TNR length, leading/lagging strand replication polarity, and distance to the viral replication origin. Hence, strong evidence supports both replication-dependent and replication-independent mechanisms of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability.

A pharmacological approach to reducing the length of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> TNRs was used by Hashem et al. in lymphoblast cell lines derived from DM1 patients (CTG)<sub>~770</sub>•(CAG)<sub>~770</sub> repeats [156]. Short term treatment with several DNA damaging drugs (ethylmethanesulfonate, mitomycin C, mitoxantrone, doxorubicin) led to the accumulation of smaller (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeat alleles in the cell population, often to fewer than 100 repeats. The rate of shift in the population profile indicated that the effects of the drugs were on the TNRs directly, rather than through mitotic selection. This result is significant given the tendency of the DM1 (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeats to expand rather than contract in patients and in culture. A similar study by Yang et al. showed that the replication inhibitors aphidicolin (which inhibits both leading and lagging strand DNA polymerases [157]) and emetine (which selectively blocks lagging strand Okazaki fragment synthesis [158]), but not mimosine (which induces DNA double strand breaks and arrests cells in late G1 phase [159,160]), increased the rate of (CTG)<sub>n</sub> expansion in DM fibroblast cells. In these experiments only the expanded DM1 allele ((CTG)<sub>~220</sub>•(CAG)<sub>~220</sub>) was altered, leaving the normal allele, (CTG)<sub>12</sub>•(CAG)<sub>12</sub> unaffected. Aphidicolin and emetine enhanced the magnitude of short expansions in almost 100% of cells approximately three-fold, while up to 25% of cells gained more than 120 repeats. Likewise, in kidney cells from *Dmt-D* transgenic mice carrying (CTG)



$160 \cdot (\text{CAG})_{160}$  repeats, Gomes-Pereira and Monckton [161] showed that prolonged exposure to the nucleoside analog and chain elongation inhibitor cytosine arabinoside, the intercalating mutagen ethidium bromide, the DNA methylation inhibitor 5-azacytidine, and aspirin, reduced the rate of repeat expansion, while exposure to caffeine, which uncouples DNA replication and repair from cell cycle checkpoints, increased the rate of expansion.

An alternative approach to the study of  $(\text{CTG})_n \cdot (\text{CAG})_n$  instability was taken by Liu et al., who constructed a clonal HeLa cell line containing a single FLP recombinase target site into which  $(\text{CTG})_n \cdot (\text{CAG})_n$  repeats of various lengths were integrated alongside the human c-myc replication origin in either replication orientation at the same ectopic chromosomal site [7]. In these HeLa/c-myc: $(\text{CTG})_n \cdot (\text{CAG})_n$  cell lines, the  $(\text{CTG})_n \cdot (\text{CAG})_n$  tracts displayed time-, replication polarity-, and repeat length-dependence of instability. Moreover, treatment of these cells with emetine, FEN1 siRNA or low dose aphidicolin rapidly (< 10 population doublings) and efficiently induced instability of the premutation length  $(\text{CTG})_{45} \cdot (\text{CAG})_{45}$  and disease-related  $(\text{CTG})_{102} \cdot (\text{CAG})_{102}$  TNRs, but not normal length  $(\text{CTG})_{12} \cdot (\text{CAG})_{12}$  TNRs. For all three treatments (low dose aphidicolin, emetine, siRNA) there was a bias towards contraction when  $(\text{CTG})_{102}$  was in the lagging strand template, and towards expansion when  $(\text{CAG})_{102}$  was in the lagging strand template. The presence of  $(\text{CAG})_n$  in the lagging strand template is the same replication polarity that has generated  $(\text{CTG})_n \cdot (\text{CAG})_n$  expansions in all other model systems [5,162]. Additional RNAi experiments using these cells (GL, ML, submitted) have confirmed the results of yeast studies in which mutation of Tof1 (human Timeless), Csm3 (human Tipin), or Mrc1 (human Claspin) dramatically increased similar patterns of  $(\text{CTG})_n \cdot (\text{CAG})_n$  instability.

In general, the treatment of cultured human or transgenic mouse cells with DNA damaging drugs or replication inhibitors demonstrates that environmental agents can modulate  $(\text{CTG})_n \cdot (\text{CAG})_n$  microsatellite instability, and that agents that cause acute DNA damage or repair are at least three orders of magnitude more efficient at inducing TNR instability than transcription-induced destabilization [153], although a unifying mechanism for explaining the observed changes has not emerged.

Towards clarifying the relationship between replication and expansion of the DMPK  $(\text{CTG})_n \cdot (\text{CAG})_n$  TNR, Cleary et al. analyzed origin activity across the DMPK locus in age-, tissue- and sex-matched human control and DM1 fibroblasts [4]. These experiments revealed two replication origins, upstream and downstream of the DMPK  $(\text{CTG})_n \cdot (\text{CAG})_n$  repeats in both control and DM1 cells. Our laboratory has independently confirmed the presence of origins upstream and downstream of the DMPK  $(\text{CTG})_n \cdot (\text{CAG})_n$  repeats in matched DM1 and non-DM1 cells. The

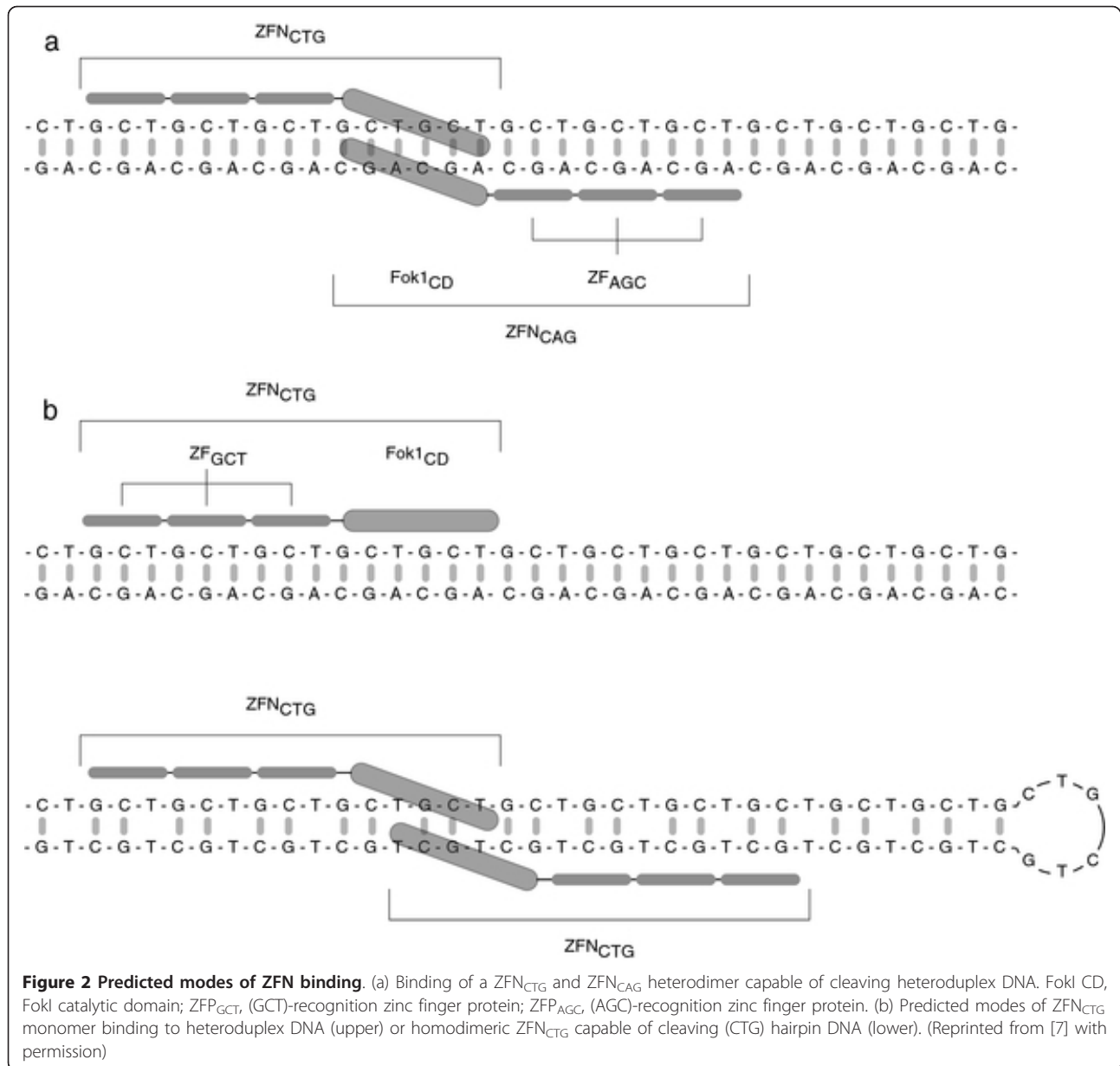
upstream origin coincides with that found by Cleary et al., while the downstream origin is approximately 2 kb closer to the TNR (GL, ML, submitted). Cleary et al. also mapped the activity of these origins in transgenic mice containing the ~45 kb DMPK locus from control (DM20,  $(\text{CTG})_{\sim 20} \cdot (\text{CAG})_{\sim 20}$ ) or DM1 (DM328,  $(\text{CTG})_{> 300} \cdot (\text{CAG})_{> 300}$ ) and found that only the origin downstream of the expanded  $(\text{CTG})_{> 300} \cdot (\text{CAG})_{> 300}$  TNR was active. It is thus possible that the downstream origin (which positions  $(\text{CAG})_n$  in the lagging strand template) is responsible for replication and expansion of the TNR in the transgenic DM328 cells. However, extension of this interpretation to human cells or other chromosomal environments is clouded by the observations that, in contrast to the transgenes, both upstream and downstream origins were equally active in human control and DM1 fibroblasts, while integration of the nonexpanded control (DM20) DMPK locus resulted in inactivation of both upstream and downstream origins in the transgenic mice.

As discussed above,  $(\text{CTG})_n \cdot (\text{CAG})_n$  instability is believed to result from the formation of hairpins in template strand DNA leading to contractions, or in newly synthesized DNA leading to expansions. Nevertheless, direct proof of hairpin formation in vivo has been lacking. To test for the presence of hairpins in vivo, synthetic zinc finger proteins (ZFPs) were engineered that specifically recognize either the  $(\text{CTG})_n$  strand or  $(\text{CAG})_n$  strand of the DMPK TNR, and fused to the FokI nuclease catalytic domain. The resulting zinc finger nucleases (ZFNs) dimerize only after zinc finger binding to their respective DNA substrate, which activates the nuclease catalytic domains. As diagrammed in Figure 2, heterodimerization of ZFN<sub>CTG</sub> and ZFN<sub>CAG</sub> is required to cleave Watson-Crick duplex DNA. However, hairpin DNA presents the same sequence ( $(\text{CTG})_n$  or  $(\text{CAG})_n$ ) on both legs of the stem, and can be cleaved by a ZFN<sub>CTG</sub> or ZFN<sub>CAG</sub> homodimer, respectively. Expression of the ZFNs in the HeLa/c-myc: $(\text{CTG})_n \cdot (\text{CAG})_n$  cell lines followed by PCR across the ectopic  $(\text{CTG})_n \cdot (\text{CAG})_n$  TNRs demonstrated directly that hairpins form in vivo on both leading strand and lagging strand templates. Moreover, ZFN cleavage was inhibited in serum-deprived nondividing cells, implying that hairpin formation in this system is replication-dependent [7].

## Conclusions

$(\text{CTG})_n$  and  $(\text{CAG})_n$  trinucleotide repeat sequences can form stable hairpins in vitro and in vivo, however, there is a facile transition of  $(\text{CTG})_n$  or  $(\text{CAG})_n$  hairpins to duplex in the presence of their complementary sequences in vitro. This suggests that other factors prolong the lifetime of  $(\text{CTG})_n$  and  $(\text{CAG})_n$  hairpins in vivo, among which may be MMR complexes, negative supercoiling behind replication or transcription forks, replication fork reversal, and protein, RNA, or leading





strand binding of the hairpin complement. In HeLa/c-myc:(CTG)<sub>n</sub>•(CAG)<sub>n</sub> cells in culture, the rapid and efficient cleavage of hairpins *in vivo* by sequence- and structure-specific synthetic zinc finger nucleases, compared to the relatively extended time required before the appearance of expansions or contractions, raises another alternative, namely that hairpins are common but short-lived *in vivo*, and rarely result in TNR instability unless DNA replication or repair is perturbed. The efficient and accurate repair of preformed hairpins in cell extracts is consistent with this notion.

The instability of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> repeats and the frequency of chromosome breakage are increased by

mutations in yeast replisome proteins. These findings strengthen the link between replication fork instability, hairpin formation, the intra-S phase checkpoint, and DNA damage responses. The similar phenotypes of mutations in yeast replisome proteins and knockdown of orthologous human proteins suggest that evolutionarily conserved pathways operate to stabilize replication forks and maximize the integrity of replication.

Not surprisingly, cell cycle and checkpoint pathways appear to play a role in murine (CTG)<sub>n</sub>•(CAG)<sub>n</sub> stability. An outstanding difference between transgenic mouse systems and the human *in vitro* repair systems is the apparent contribution of the MMR proteins to instability in

mice and the absence of their effect on in vitro repair. One possibility is that the preformation of stable hairpin substrates for in vitro repair may bypass an in vivo effect of chromatin structure, DNA metabolism, or MMR proteins.

Several fundamental questions concerning the mechanism of (CTG)<sub>n</sub>•(CAG)<sub>n</sub> instability remain to be addressed. For example, do contractions and expansions occur as consequences of the same process of replication, replication restart or postreplication repair? Do contractions and expansions occur in different phases of the mitotic cycle? Do contractions (or expansions) occur preferentially on the leading or lagging strand during replication? Are different pathways involved in the instability of various length TNRs? Does DNA damage promote hairpin formation? Which repair mechanisms are responsible for TNR instability in postmitotic cells? What is the mechanism of transcription-induced instability?

The use of yeast and transgenic mouse mutants, and RNAi to produce human cells and cell extracts deficient in specific functions promise to give insight into these questions, and thereby reveal second site genetic modifiers of TNR instability that can be used in prognosis and therapy.

#### Authors' contributions

GL and ML drafted and revised the manuscript. Both authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 7 December 2011 Accepted: 27 February 2012

Published: 27 February 2012

#### References

- Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, Hetherington C, Lawton M, Trotter Y, Leach H, Davies SW, Bates GP: **Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice.** *Cell* 1996, **87**:493-506.
- Mirkin SM: **Expandable DNA repeats and human disease.** *Nature* 2007, **447**:932-940.
- 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 USA* 2008, **105**:9936-9941.
- Cleary JD, Tome S, Lopez Castel A, Panigrahi GB, Foirey L, Hagerman KA, Sroka H, Chitayat D, Gourdon G, Pearson CE: **Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus.** *Nat Struct Mol Biol* 2010, **17**:1079-1087.
- Pearson CE, Edamura KN, Cleary JD: **Repeat instability: mechanisms of dynamic mutations.** *Nat Rev Genet* 2005, **6**:729-742.
- Cleary JD, Pearson CE: **Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability.** *Trends Genet* 2005, **21**:272-280.
- Liu G, Chen X, Bissler JJ, Sinden RR, Leffak M: **Replication-dependent instability at (CTG) × (CAG) repeat hairpins in human cells.** *Nat Chem Biol* 2010, **6**:652-659.
- Liu G, Bissler JJ, Sinden RR, Leffak M: **Unstable Spinocerebellar Ataxia Type 10 (ATTCT)\*(AGAAT) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells.** *MolCell Biol* 2007, **27**:7828-7838.
- Panaite PA, Gantelet E, Kraftsik R, Gourdon G, Kuntzer T, Barakat-Walter I: **Myotonic dystrophy transgenic mice exhibit pathologic abnormalities in diaphragm neuromuscular junctions and phrenic nerves.** *J Neuropathol Exp Neurol* 2008, **67**:763-772.
- Kunkel TA: **Nucleotide repeats. Slippery DNA and diseases.** *Nature* 1993, **365**:207-208.
- Schweitzer JK, Livingston DM: **The effect of DNA replication mutations on CAG tract stability in yeast.** *Genetics* 1999, **152**:953-963.
- Ireland MJ, Reinke SS, Livingston DM: **The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast.** *Genetics* 2000, **155**:1657-1665.
- Hartenstine MJ, Goodman MF, Petruska J: **Base stacking and even/odd behavior of hairpin loops in DNA triplet repeat slippage and expansion with DNA polymerase.** *J Biol Chem* 2000, **275**:18382-18390.
- Kunkel TA: **Misalignment-mediated DNA synthesis errors.** *Biochemistry* 1990, **29**:8003-8011.
- Kerrest A, Anand RP, Sundararajan R, Bermejo R, Liberi G, Dujon B, Freudenreich CH, Richard GF: **SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination.** *Nat Struct Mol Biol* 2009, **16**:159-167.
- Kovtun IV, Thornhill AR, McMurray CT: **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.
- 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 USA* 1998, **95**:12438-12443.
- Gordenin DA, Kunkel TA, Resnick MA: **Repeat expansion—all in a flap?** *Nat Genet* 1997, **16**:116-118.
- Callahan JL, Andrews KJ, Zakian VA, Freudenreich CH: **Mutations in yeast replication proteins that increase CAG/CTG expansions also increase repeat fragility.** *Mol Cell Biol* 2003, **23**:7849-7860.
- 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.
- Jarem DA, Wilson NR, Schermerhorn KM, Delaney S: **Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion.** *DNA Repair (Amst)* 2011, **10**:887-896.
- Lahue RS, Slater DL: **DNA repair and trinucleotide repeat instability.** *Front Biosci* 2003, **8**:s653-665.
- Hubert L Jr, Lin Y, Dion V, Wilson JH: **Topoisomerase 1 and Single-Strand Break Repair Modulate Transcription-Induced CAG Repeat Contraction in Human Cells.** *Mol Cell Biol* 2011, **31**:3105-3112.
- Lin Y, Dent SY, Wilson JH, Wells RD, Napierala M: **R loops stimulate genetic instability of CTG.CAG repeats.** *Proc Natl Acad Sci USA* 2010, **107**:692-697.
- Lin Y, Hubert L Jr, Wilson JH: **Transcription destabilizes triplet repeats.** *Mol Carcinog* 2009, **48**:350-361.
- Lin Y, Wilson JH: **Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair.** *Mol Cell Biol* 2007, **27**:6209-6217.
- Nakamori M, Pearson CE, Thornton CA: **Bidirectional transcription stimulates expansion and contraction of expanded (CTG)<sub>n</sub>(CAG)<sub>n</sub> repeats.** *Hum Mol Genet* 2011, **20**:580-588.
- Wang G, Vasquez KM: **Models for chromosomal replication-independent non-B DNA structure-induced genetic instability.** *Mol Carcinog* 2009, **48**:286-298.
- Wang G, Vasquez KM: **Non-B DNA structure-induced genetic instability.** *Mutat Res* 2006, **598**:103-119.
- Gonitel R, Moffitt H, Sathasivam K, Woodman B, Detloff PJ, Faull RL, Bates GP: **DNA instability in postmitotic neurons.** *Proc Natl Acad Sci USA* 2008, **105**:3467-3472.
- Parniewski P, Jaworski A, Wells RD, Bowater RP: **Length of CTG.CAG repeats determines the influence of mismatch repair on genetic instability.** *J Mol Biol* 2000, **299**:865-874.
- Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, Housman DE, Shaw DJ: **Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy.** *Nature* 1992, **355**:545-546.
- Aslanidis C, Jansen G, Amemiya C, Shutler G, Mahadevan M, Tsilifidis C, Chen C, Alleman J, Wormskamp NG, Vooijs M, et al: **Cloning of the essential myotonic dystrophy region and mapping of the putative defect.** *Nature* 1992, **355**:548-551.

34. Buxton J, Shelbourne P, Davies J, Jones C, Van Tongeren T, Aslanidis C, de Jong P, Jansen G, Anvret M, Riley B, et al: **Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy.** *Nature* 1992, **355**:547-548.
35. Mahadevan M, Tsilifidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, et al: **Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene.** *Science* 1992, **255**:1253-1255.
36. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, et al: **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* 1992, **69**:385.
37. Tsilifidis C, MacKenzie AE, Mettler G, Barcelo J, Korneluk RG: **Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy.** *Nat Genet* 1992, **1**:192-195.
38. Illarioshkin SN, Igarashi S, Onodera O, Markova ED, Nikolskaya NN, Tanaka H, Chabrashvili TZ, Insarova NG, Endo K, Ivanova-Smolenskaya IA, et al: **Trinucleotide repeat length and rate of progression of Huntington's disease.** *Ann Neurol* 1994, **36**:630-635.
39. Rosenblatt A, Brinkman RR, Liang KY, Almqvist EW, Margolis RL, Huang CY, Sherr M, Franz ML, Abbott MH, Hayden MR, Ross CA: **Familial influence on age of onset among siblings with Huntington disease.** *Am J Med Genet* 2001, **105**:399-403.
40. Brinkman RR, Mezei MM, Theilmann J, Almqvist E, Hayden MR: **The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size.** *Am J Hum Genet* 1997, **60**:1202-1210.
41. Cho DH, Tapscott SJ: **Myotonic dystrophy: emerging mechanisms for DM1 and DM2.** *Biochim Biophys Acta* 2007, **1772**:195-204.
42. Otten AD, Tapscott SJ: **Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure.** *Proc Natl Acad Sci USA* 1995, **92**:5465-5469.
43. Saveliev A, Everett C, Sharpe T, Webster Z, Festenstein R: **DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing.** *Nature* 2003, **422**:909-913.
44. Sicot G, Gourdon G, Gomes-Pereira M: **Myotonic dystrophy, when simple repeats reveal complex pathogenic entities: new findings and future challenges.** *Hum Mol Genet* 2011, **20**:R116-R123.
45. Kim SH, Pytlas MJ, Rosche WA, Sinden RR: **(CAG)<sup>n</sup>(CTG) repeats associated with neurodegenerative diseases are stable in the Escherichia coli chromosome.** *J Biol Chem* 2006, **281**:27950-27955.
46. Szwarcrocka ST, Staczek P, Parniewski P: **Chromosomal model for analysis of a long CTG/CAG tract stability in wild-type Escherichia coli and its nucleotide excision repair mutants.** *Can J Microbiol* 2007, **53**:860-868.
47. Blackwood JK, Okely EA, Zahra R, Eykelenboom JK, Leach DR: **DNA tandem repeat instability in the Escherichia coli chromosome is stimulated by mismatch repair at an adjacent CAG/CTG trinucleotide repeat.** *Proc Natl Acad Sci USA* 2010, **107**:22582-22586.
48. Zahra R, Blackwood JK, Sales J, Leach DR: **Proofreading and secondary structure processing determine the orientation dependence of CAG X CTG trinucleotide repeat instability in Escherichia coli.** *Genetics* 2007, **176**:27-41.
49. Napierala M, Bacolla A, Wells RD: **Increased negative superhelical density in vivo enhances the genetic instability of triplet repeat sequences.** *J Biol Chem* 2005, **280**:37366-37376.
50. Bichara M, Wagner J, Lambert IB: **Mechanisms of tandem repeat instability in bacteria.** *Mutat Res* 2006, **598**:144-163.
51. Mirkin SM: **DNA structures, repeat expansions and human hereditary disorders.** *Curr Opin Struct Biol* 2006, **16**:351-358.
52. Avila Figueroa A, Delaney S: **Mechanistic studies of hairpin to duplex conversion for trinucleotide repeat sequences.** *J Biol Chem* 2010, **285**:14648-14657.
53. Figueroa AA, Cattie D, Delaney S: **Structure of even/odd trinucleotide repeat sequences modulates persistence of non-B conformations and conversion to duplex.** *Biochemistry* 2011, **50**:4441-4450.
54. 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.
55. Gacy AM, McMurray CT: **Influence of hairpins on template reannealing at trinucleotide repeat duplexes: a model for slipped DNA.** *Biochemistry* 1998, **37**:9426-9434.
56. Zheng GX, Sinden RR: **Effect of base composition at the center of inverted repeated DNA sequences on cruciform transitions in DNA.** *J Biol Chem* 1988, **263**:5356-5361.
57. Panyutin IG, Hsieh P: **Formation of a single base mismatch impedes spontaneous DNA branch migration.** *J Mol Biol* 1993, **230**:413-424.
58. Cote AG, Lewis SM: **Mus81-dependent double-strand DNA breaks at in vivo-generated cruciform structures in S. cerevisiae.** *Mol Cell* 2008, **31**:800-812.
59. Panigrahi GB, Cleary JD, Pearson CE: **In vitro (CTG)<sup>n</sup>(CAG) expansions and deletions by human cell extracts.** *J Biol Chem* 2002, **277**:13926-13934.
60. Ilves I, Petojevic T, Pesavento JJ, Botchan MR: **Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins.** *Mol Cell* 2010, **37**:247-258.
61. Costa A, Ilves I, Tamberg N, Petojevic T, Nogales E, Botchan MR, Berger JM: **The structural basis for MCM2-7 helicase activation by GINS and Cdc45.** *Nat Struct Mol Biol* 2011, **18**:471-477.
62. Delagoutte E, Goellner GM, Guo J, Baldacci G, McMurray CT: **Single-stranded DNA-binding protein in vitro eliminates the orientation-dependent impediment to polymerase passage on CAG/CTG repeats.** *J Biol Chem* 2008, **283**:13341-13356.
63. Petruska J, Hartenstine MJ, Goodman MF: **Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease.** *J Biol Chem* 1998, **273**:5204-5210.
64. Leclercq S, Rivals E, Jarne P: **DNA slippage occurs at microsatellite loci without minimal threshold length in humans: a comparative genomic approach.** *Genome Biol Evol* 2010, **2**:325-335.
65. Panigrahi GB, Lau R, Montgomery SE, Leonard MR, Pearson CE: **Slipped (CTG)<sup>n</sup>(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair.** *Nat Struct Mol Biol* 2005, **12**:654-662.
66. Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L, Li GM: **Reconstitution of 5'-directed human mismatch repair in a purified system.** *Cell* 2005, **122**:693-705.
67. Hou C, Zhang T, Tian L, Huang J, Gu L, Li GM: **The Role of XPG in Processing (CAG)<sup>n</sup>(CTG)<sup>n</sup> DNA Hairpins.** *Cell Biosci* 2011, **1**:11.
68. Hou C, Chan NL, Gu L, Li GM: **Incision-dependent and error-free repair of (CAG)<sup>n</sup>(CTG)<sup>n</sup> hairpins in human cell extracts.** *Nat Struct Mol Biol* 2009, **16**:869-875.
69. Panigrahi GB, Slean MM, Simard JP, Gileadi O, Pearson CE: **Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired.** *Proc Natl Acad Sci USA* 2010, **107**:12593-12598.
70. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, Monckton DG: **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.
71. Manley K, Shirley TL, Flaherty L, Messer A: **Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice.** *Nat Genet* 1999, **23**:471-473.
72. Wheeler VC, Lebel LA, Vrbancac V, Teed A, te Riele H, MacDonald ME: **Mismatch repair gene Msh2 modifies the timing of early disease in Hdh (Q111) striatum.** *Hum Mol Genet* 2003, **12**:273-281.
73. Bhattacharyya S, Lahue RS: **Srs2 helicase of Saccharomyces cerevisiae selectively unwinds triplet repeat DNA.** *J Biol Chem* 2005, **280**:33311-33317.
74. Daee DL, Mertz T, Lahue RS: **Postreplication repair inhibits CAG/CTG repeat expansions in Saccharomyces cerevisiae.** *Mol Cell Biol* 2007, **27**:102-110.
75. 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.
76. Raziidlo DF, Lahue RS: **Mrc1, Tof1 and Csm3 inhibit CAG/CTG repeat instability by at least two mechanisms.** *DNA Repair (Amst)* 2008, **7**:633-640.
77. Freudenreich CH, Lahiri M: **Structure-forming CAG/CTG repeat sequences are sensitive to breakage in the absence of Mrc1 checkpoint function and S-phase checkpoint signaling: implications for trinucleotide repeat expansion diseases.** *Cell Cycle* 2004, **3**:1370-1374.
78. Lahiri M, Gustafson TL, Majors ER, Freudenreich CH: **Expanded CAG repeats activate the DNA damage checkpoint pathway.** *Mol Cell* 2004, **15**:287-293.
79. Yang J, Freudenreich CH: **Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner.** *Gene* 2007, **393**:110-115.

80. Krasilnikova MM, Mirkin SM: **Replication stalling at Friedreich's ataxia (GAA)<sub>n</sub> repeats in vivo.** *Mol Cell Biol* 2004, **24**:2286-2295.
81. 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.
82. Voineagu I, Surka CF, Shishkin AA, Krasilnikova MM, Mirkin SM: **Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility.** *Nat Struct Mol Biol* 2009, **16**:226-228.
83. Miret JJ, Pessoa-Brandao L, Lahue RS: **Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1997, **17**:3382-3387.
84. Spiro C, Pelletier R, Rolfmeier ML, Dixon MJ, Lahue RS, Gupta G, Park MS, Chen X, Mariappan SV, McMurray CT: **Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats.** *Mol Cell* 1999, **4**:1079-1085.
85. Bhattacharyya S, Lahue RS: ***Saccharomyces cerevisiae* Srs2 DNA helicase selectively blocks expansions of trinucleotide repeats.** *Mol Cell Biol* 2004, **24**:7324-7330.
86. Wells RD, Ashizawa T: *Genetic instabilities and neurological diseases.* 2 edition. Amsterdam; Boston: Elsevier; 2006.
87. Wolff RK, Plaetke R, Jeffreys AJ, White R: **Unequal crossingover between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci.** *Genomics* 1989, **5**:382-384.
88. Casper AM, Nghiem P, Airt MF, Glover TW: **ATR regulates fragile site stability.** *Cell* 2002, **111**:779-789.
89. Durkin SG, Glover TW: **Chromosome fragile sites.** *Annu Rev Genet* 2007, **41**:169-192.
90. Sundararajan R, Gellon L, Zunder RM, Freudenreich CH: **Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in *Saccharomyces cerevisiae*.** *Genetics* 2010, **184**:65-77.
91. Gellon L, Razidlo DF, Gleeson O, Verra L, Schulz D, Lahue RS, Freudenreich CH: **New functions of Ctf18-RFC in preserving genome stability outside its role in sister chromatid cohesion.** *PLoS Genet* 2011, **7**: e1001298.
92. Bermudez VP, Maniwa Y, Tappin I, Ozato K, Yokomori K, Hurwitz J: **The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA.** *Proc Natl Acad Sci USA* 2003, **100**:10237-10242.
93. Bylund GO, Burgers PM: **Replication protein A-directed unloading of PCNA by the Ctf18 cohesion establishment complex.** *Mol Cell Biol* 2005, **25**:5445-5455.
94. Lou H, Komata M, Katou Y, Guan Z, Reis CC, Budd M, Shirahige K, Campbell JL: **Mrc1 and DNA polymerase epsilon function together in linking DNA replication and the S phase checkpoint.** *Mol Cell* 2008, **32**:106-117.
95. Tourriere H, Pasero P: **Maintenance of fork integrity at damaged DNA and natural pause sites.** *DNA Repair (Amst)* 2007, **6**:900-913.
96. Tourriere H, Versini G, Cordon-Preciado V, Alabert C, Pasero P: **Mrc1 and Top1 promote replication fork progression and recovery independently of Rad53.** *Mol Cell* 2005, **19**:699-706.
97. Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, Calzada A, Labib K: **A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome.** *EMBO J* 2009, **28**:2992-3004.
98. Miles J, Formosa T: **Evidence that POB1, a *Saccharomyces cerevisiae* protein that binds to DNA polymerase alpha, acts in DNA metabolism in vivo.** *Mol Cell Biol* 1992, **12**:5724-5735.
99. Ricke RM, Bielinsky AK: **Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha.** *Mol Cell* 2004, **16**:173-185.
100. Kouprina N, Kroll E, Bannikov V, Bliskovsky V, Gizatullin R, Kirillov A, Shestopalov B, Zakharyev V, Hieter P, Spencer F, et al: **CTF4 (CHL15) mutants exhibit defective DNA metabolism in the yeast *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1992, **12**:5736-5747.
101. Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry R, Spencer FA: **S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion.** *Mol Biol Cell* 2004, **15**:1724-1735.
102. van den Broek WJ, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJ, Wieringa B: **Somatic expansion behaviour of the (CTG)<sub>n</sub> repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins.** *Hum Mol Genet* 2002, **11**:191-198.
103. 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.
104. Fortune MT, Vassilopoulos C, Coolbaugh MI, Siciliano MJ, Monckton DG: **Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability.** *Hum Mol Genet* 2000, **9**:439-445.
105. Usdin K: **Bending the rules: unusual nucleic acid structures and disease pathology in the repeat expansion diseases.** In *Genetic instabilities and neurological diseases.* 617-635 edition. Edited by: Wells RD, Ashizawa T. Amsterdam; Boston: Elsevier; 2006.
106. Shelbourne PF, Monckton DG: **Somatic mosaicism of expanded CAG/CTG repeats in humans and mice: and consequences.** In *Genetic instabilities and neurological diseases.* Edited by: Wells RD, Ashizawa T. Amsterdam; Boston: Elsevier; 2006:537-561.
107. Gomes-Pereira M, Foiry L, Gourdon G: **Transgenic mouse models of unstable trinucleotide repeats: towards an understanding of disease-associated repeat size mutation.** In *Genetic instabilities and neurological diseases.* 2 edition. Edited by: Wells RD, Ashizawa T. Amsterdam; Boston: Elsevier; 2006:563-586.
108. Savouret C, Junien C, Gourdon G: **Analysis of CTG repeats using DM1 model mice.** *Methods Mol Biol* 2004, **277**:185-197.
109. Kovtun IV, McMurray CT: **Models of repair underlying trinucleotide DNA expansion.** In *Genetic instabilities and neurological diseases.* 2 edition. Edited by: Wells RD, Ashizawa T. Amsterdam; Boston: Elsevier; 2006:679-690.
110. Seznec H, Agbulut O, Sergeant N, Savouret C, Ghestem A, Tabti N, Willer JC, Ourth L, Duros C, Brisson E, et al: **Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities.** *Hum Mol Genet* 2001, **10**:2717-2726.
111. Kumari D, Usdin K: **Chromatin remodeling in the noncoding repeat expansion diseases.** *J Biol Chem* 2009, **284**:7413-7417.
112. Kumari D, Biacsi RE, Usdin K: **Repeat expansion affects both transcription initiation and elongation in Friedreich ataxia cells.** *J Biol Chem* 2011, **286**:4209-4215.
113. Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ: **Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF.** *Mol Cell* 2005, **20**:483-489.
114. Libby RT, Hagerman KA, Pineda VW, Lau R, Cho DH, Baccam SL, Axford MM, Cleary JD, Moore JM, Sopher BL, et al: **CTCF cis-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination.** *PLoS Genet* 2008, **4**:e1000257.
115. Libby RT, Monckton DG, Fu YH, Martinez RA, McAbney JP, Lau R, Einum DD, Nichol K, Ware CB, Ptacek LJ, et al: **Genomic context drives SCA7 CAG repeat instability, while expressed SCA7 cDNAs are intergenerationally and somatically stable in transgenic mice.** *Hum Mol Genet* 2003, **12**:41-50.
116. Savouret C, Garcia-Cordier C, Megret J, te Riele H, Junien C, Gourdon G: **MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice.** *Mol Cell Biol* 2004, **24**:629-637.
117. Seznec H, Lia-Baldini AS, Duros C, Fouquet C, Lacroix C, Hofmann-Radvanyi H, Junien C, Gourdon G: **Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability.** *Hum Mol Genet* 2000, **9**:1185-1194.
118. Lia AS, Seznec H, Hofmann-Radvanyi H, Radvanyi F, Duros C, Saquet C, Blanche M, Junien C, Gourdon G: **Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities.** *Hum Mol Genet* 1998, **7**:1285-1291.
119. 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.
120. Wahba L, Amon JD, Koshland D, Vuica-Ross M: **RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability.** *Mol Cell* 2011, **44**:978-988.
121. Helmrich A, Ballarino M, Tora L: **Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes.** *Mol Cell* 2011, **44**:966-977.



122. Ljungman M: The transcription stress response. *Cell Cycle* 2007, **6**:2252-2257.
123. Entezam A, Usdin K: ATM and ATR protect the genome against two different types of tandem repeat instability in Fragile X premutation mice. *Nucleic Acids Res* 2009, **37**:6371-6377.
124. Entezam A, Usdin K: ATR protects the genome against CGG.CCG-repeat expansion in Fragile X premutation mice. *Nucleic Acids Res* 2008, **36**:1050-1056.
125. 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.
126. Lin Y, Wilson JH: Diverse effects of individual mismatch repair components on transcription-induced CAG repeat instability in human cells. *DNA Repair (Amst)* 2009, **8**:878-885.
127. Lin Y, Wilson JH: Transcription-induced DNA toxicity at trinucleotide repeats: double bubble is trouble. *Cell Cycle* 2011, **10**:611-618.
128. Reddy K, Tam M, Bowater RP, Barber M, Tomlinson M, Nichol Edamura K, Wang YH, Pearson CE: Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res* 2011, **39**:1749-1762.
129. Reddy S, Paul S: Cis effects of CTG expansion in myotonic dystrophy type 1. In *Genetic instabilities and neurological diseases*. 2 edition. Edited by: Wells RD, Ashizawa T. Amsterdam; Boston: Elsevier; 2006:55-78.
130. Tome S, Holt I, Edelmann W, Morris GE, Munnich A, Pearson CE, Gourdon G: MSH2 ATPase domain mutation affects CTG\*CAG repeat instability in transgenic mice. *PLoS Genet* 2009, **5**:e1000482.
131. Foirey 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 transgenic mice. *Hum Genet* 2006, **119**:520-526.
132. Modrich P, Lahue R: Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem* 1996, **65**:101-133.
133. Modrich P: Mechanisms in eukaryotic mismatch repair. *J Biol Chem* 2006, **281**:30305-30309.
134. Li GM: Mechanisms and functions of DNA mismatch repair. *Cell research* 2008, **18**:85-98.
135. Savouret C, Brisson E, Essers J, Kanaar R, Pastink A, te Riele H, Junien C, Gourdon G: CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J* 2003, **22**:2264-2273.
136. 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.
137. Owen BA, Yang Z, Lai M, Gajec 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.
138. McMurray CT: Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA Repair (Amst)* 2008, **7**:1121-1134.
139. Tian L, Hou C, Tian K, Holcomb NC, Gu L, Li GM: Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. *J Biol Chem* 2009, **284**:20452-20456.
140. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, McMurray CT, Wilson SH: Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *J Biol Chem* 2009, **284**:28352-28366.
141. Spiro C, Pelletier R, Rolfmeier ML, Dixon MJ, Lahue RS, Gupta G, Park MS, Chen X, Mariappan SV, McMurray CT: Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol Cell* 1999, **4**:1079-1085.
142. McMurray CT: Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* 2010, **11**:786-799.
143. van den Broek WJ, Nelen MR, van der Heijden GW, Wansink DG, Wieringa B: Fen1 does not control somatic hypermutability of the (CTG)(n)\*(CAG)(n) repeat in a knock-in mouse model for DM1. *FEBS Lett* 2006, **580**:5208-5214.
144. Spiro C, McMurray CT: Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. *Mol Cell Biol* 2003, **23**:6063-6074.
145. Goula AV, Berquist BR, Wilson DM, Wheeler VC, Trotter Y, Merienne K: Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet* 2009, **5**:e1000749.
146. Lloret A, Dragileva E, Teed A, Espinola J, Fossale E, Gillis T, Lopez E, Myers RH, MacDonald ME, Wheeler VC: Genetic background modifies nuclear mutant huntingtin accumulation and HD CAG repeat instability in Huntington's disease knock-in mice. *Hum Mol Genet* 2006, **15**:2015-2024.
147. Lee J-M, Zhang J, Su AI, Walker JR, Wiltshire T, Kang K, Dragileva E, Gillis T, Lopez ET, Boily M-J, et al: A novel approach to investigate tissue-specific trinucleotide repeat instability. *BMC Systems Biology* 2010, **4**:29.
148. Tome S, Panigrahi GB, Lopez Castel A, Foirey L, Melton DW, Gourdon G, Pearson CE: Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. *Hum Mol Genet* 2011, **20**:2131-2143.
149. Mateizel I, De Temmerman N, Ullmann U, Cauffman G, Sermon K, Van de Velde H, De Rycke M, Degreef E, Devroey P, Liebaers I, Van Steirteghem A: Derivation of human embryonic stem cell lines from embryos obtained after IVF and after PGD for monogenic disorders. *Hum Reprod* 2006, **21**:503-511.
150. Niclis JC, Trounson AO, Dottori M, Ellisdon AM, Bottomley SP, Verlinsky Y, Cram DS: Human embryonic stem cell models of Huntington disease. *Reprod Biomed Online* 2009, **19**:106-113.
151. De Temmerman N, Seneca S, Van Steirteghem A, Haentjens P, Van der Elst J, Liebaers I, Sermon KD: CTG repeat instability in a human embryonic stem cell line carrying the myotonic dystrophy type 1 mutation. *Mol Hum Reprod* 2008, **14**:405-412.
152. Seriola A, Spits C, Simard JP, Hilven P, Haentjens P, Pearson CE, Sermon K: Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. *Hum Mol Genet* 2011, **20**:176-185.
153. Lin Y, Dion V, Wilson JH: Transcription promotes contraction of CAG repeat tracts in human cells. *Nat Struct Mol Biol* 2006, **13**:179-180.
154. Hubert L Jr, 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.
155. Cleary JD, Nichol K, Wang YH, Pearson CE: Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat Genet* 2002, **31**:37-46.
156. Hashem VI, Pytlos MJ, Klysik EA, Tsuji K, Khajavi M, Ashizawa T, Sinden RR: Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. *Nucleic Acids Res* 2004, **32**:6334-6346.
157. Nethanel T, Kaufmann G: Two DNA polymerases may be required for synthesis of the lagging DNA strand of simian virus 40. *J Virol* 1990, **64**:5912-5918.
158. Burhans WC, Vassilev LT, Wu J, Sogo JM, Nallaseth FS, DePamphilis ML: Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. *EMBO J* 1991, **10**:4351-4360.
159. Mladenov E, Anachkova B: DNA breaks induction by mimosine. *Z Naturforsch C* 2003, **58**:732-735.
160. Szuts D, Krude T: Cell cycle arrest at the initiation step of human chromosomal DNA replication causes DNA damage. *J Cell Sci* 2004, **117**:4897-4908.
161. Gomes-Pereira M, Monckton DG: Chemically induced increases and decreases in the rate of expansion of a CAG\*CTG triplet repeat. *Nucleic Acids Res* 2004, **32**:2865-2872.
162. Mirkin SM, Smirnova EV: Positioned to expand. *Nat Genet* 2002, **31**:5-6.

doi:10.1186/2045-3701-2-7

Cite this article as: Liu and Leffak: Instability of (CTG)<sub>n</sub>\*(CAG)<sub>n</sub> trinucleotide repeats and DNA synthesis. *Cell & Bioscience* 2012 **2**:7.