JBC RESEARCH ARTICLE

Stable G-quadruplex DNA structures promote replication-dependent genome instability

Received for publication, January 3, 2022, and in revised form, April 12, 2022 Published, Papers in Press, April 18, 2022,
https://doi.org/10.1016/j.jbc.2022.101947

S. Dean Rider Jr,[1,‡], Rujuta Yashodhan Gadgil[1,‡], David C. Hitch[1], French J. Damewood IV,[1], Nathen Zavada[1], Matilyn Shanahan[1], Venicia Alhawach[1], Resha Shrestha[1], Kazuo Shin-ya[2], and Michael Leffak[1,*]

From the [1]Department of Biochemistry and Molecular Biology, Wright State University, Dayton, Ohio, USA; [2]Biomedical Information Research Center, National Institute of Advanced Industrial Science and Technology, Koto-ku, Tokyo, Japan

Edited by Patrick Sung

G-quadruplex (G4)–prone structures are abundant in mammalian genomes, where they have been shown to influence DNA replication, transcription, and genome stability. In this article, we constructed cells with a single ectopic homopurine/homopyrimidine repeat tract derived from the polycystic kidney disease type 1 (PKD1) locus, which is capable of forming triplex (H3) and G4 DNA structures. We show that ligand stabilization of these G4 structures results in deletions of the G4 consensus sequence, as well as kilobase deletions spanning the G4 and ectopic sites. Furthermore, we show that DNA double-strand breaks at the ectopic site are dependent on the nuclease Mus81. Hypermutagenesis during sister chromatid repair extends several kilobases from the G4 site and breaks at the G4 site resulting in microhomology-mediated translocations. To determine whether H3 or G4 structures are responsible for homopurine/homopyrimidine tract instability, we derived constructs and cell lines from the PKD1 repeat, which can only form H3 or G4 structures. Under normal growth conditions, we found that G4 cell lines lost the G4 consensus sequence early during clonal outgrowth, whereas H3 cells showed DNA instability early during outgrowth but only lost reporter gene expression after prolonged growth. Thus, both the H3 and G4 non-B conformation DNAs exhibit genomic instability, but they respond differently to endogenous replication stress. Our results show that the outcomes of replication-dependent double-strand breaks at non–B-DNAs model the instability observed in microhomology-mediated break-induced replication (BIR). Marked variability in the frequency of mutagenesis during BIR suggests possible dynamic heterogeneity in the BIR replisome.

Microsatellite DNAs were originally identified as centrifugation peaks separate from bulk genomic DNA by virtue of their homogeneous base compositions (1). Most microsatellites comprise multiple tandem repeats of 1 to 9 nucleotide blocks. The tendency of microsatellite tracts to exhibit genomic instability by expansion, contraction, and recombination has been attributed to the predilection of these sequences to form non–B-DNA structures (2).

In yeast, expanded GAA (3, 4), CTG (5), and CGG (6) microsatellites are sites of DNA double-strand breaks (DSBs), chromosome fragility, and activation of the RAD9/ataxia telangiectasia and Rad3-related protein (ATR)/checkpoint kinase 1 (CHK1) DNA damage response (DDR) pathways (4, 7–9). Replication fork stalling at microsatellite non–B-DNA structures is one mechanism leading to single-ended DSBs and hypermutagenic break-induced replication (BIR) (10–14). An alternate mechanism for bypass of replication fork barriers by homologous recombination in *Schizosaccharomyces pombe* also leads to an error-prone pathway to circumvent replication barriers (15, 16).

In humans, microsatellite instability has been implicated in causing Huntington's disease ($(CAG)_n$) (17), the fragile X-related disorders, FXTAS, FXPOI, and FXS (18), myotonic dystrophy ($(CTG)_n$ and $(CCTG)_n$) (19, 20), and more than 40 other neuromuscular and developmental diseases, and cancers. The presence of microsatellite repeats at translocation break points and junctions has implicated microhomology-mediated BIR (MMBIR) in gross chromosomal rearrangement and nonrecurrent copy number variation (21–24). BIR is strongly indicated as an outcome of replication-dependent DSBs in human cells (25), and expanded CGG repeats have been shown to cause BIR in a murine cell culture system (11). Our laboratory has recently shown that Mus81 cleavage at the edge of an expanded CTG hairpin (26) causes replication-dependent BIR, hypermutation, and nonrandom chromosomal translocations (27).

Although it is not a strict tandem repeat, here we consider the asymmetric homopurine/homopyrimidine (Pu/Py) mirror repeat of the polycystic kidney disease type 1 (PKD1) gene intron 21 (IVS21) to behave as a microsatellite because of its homogeneous base composition, its replication polarity-dependent ability to stall replication forks through the formation of non–B-DNA structures, and its replication-dependent instability (27, 28). The Pu-rich strand of the PKD1 IVS21 (Pu/Py) repeat has the potential alternatively to form triplex (H3) DNA by Hoogsteen base pairing (29) or G-quadruplex (G4) DNA by Hoogsteen stabilization of G4 base stacking (30). The G4 consensus $(G)_{>2}(N)_{1-7}(G)_{>2}(N)_{1-7}(G)_{>2}(N)_{1-7}(G)_{>2}$

‡ These authors contributed equally to this work.
* For correspondence: Michael Leffak, Michael.leffak@wright.edu.

has been used to identify potential sites of G4 formation (31), although later work indicates that longer loops ($(N)_{\leq 30}$) can be readily accommodated into G4 structures (32).

Statistical calculations predict that approximately 200 G4 sequences with loops of $(N)_{\leq 15}$ will occur in a random nucleotide collection of 3.5×10^9 bp, while algorithms that account for dinucleotide preferences in human DNA predict >370,000 potential G4 structures in the genome, and G4-targeted antibody binding has indicated >700,000 sites capable of stable G4 formation in human B-lymphocyte DNA (33–38). G4 sequences have been implicated in the control of replication (39–41), transcription (30, 42), genomic looping (33), R-loop formation (43, 44), and genome instability (44). G4 structures are unwound by the FANCJ helicase *in vitro* (45), whereas depletion of FANCJ increases G4 formation *in vivo* (41, 46).

Our laboratory has shown that the PKD1 $(Pu/Py)_{88}$ repeat stalls replication forks when the Pu strand is the lagging strand template downstream from the SV40 replication origin *in vitro* and when the Pu strand is replicated as the lagging strand template from an ectopic c-myc replication origin *in vivo*. At the same ectopic site, inversion of the Pu/Py tract stalls replication when the c-myc origin is inactivated, and the Pu lagging strand template is replicated from a neighboring downstream origin (28).

Cells containing the ectopic Pu/Py repeat in the Pu lagging strand orientation ($(Pu)_{88}$) grow more slowly than cells containing the Pu/Py repeat at the same ectopic site in the Py lagging strand ($(Py)_{88}$) orientation (27). $(Pu)_{88}$ cells are also sensitive to the CHK1 inhibitor UCN-01 suggesting that the ATR/CHK1 DDR pathway is required for $(Pu)_{88}$ cell viability. Precise excision of the Pu/Py repeat from the ectopic site of $(Pu)_{88}$ cells using FLP recombinase eliminates the UCN-01 effect, demonstrating that the Pu/Py tract is directly responsible for UCN-01 sensitivity. Consistent with the requirement of the DDR for $(Pu)_{88}$ cell viability, ATR and RAD9 bind preferentially to the $(Pu)_{88}$ *versus* $(Py)_{88}$ cell ectopic site (28).

We have constructed cells in which spontaneously shortened $(Pu/Py)_{78}$ microsatellites are integrated by FLP recombinase at a single ectopic site alongside the core c-myc replication origin and flanked by dTomato and enhanced GFP (eGFP) reporter genes (26–28, 47). DNA DSBs occur with greater frequency in $(Pu)_{78}$ cells than $(Py)_{78}$ cells under nonperturbed conditions. To introduce replication stress in this system, we used the macrocyclic G4 ligand telomestatin (TMS) to stabilize G4s. TMS stacks tightly and specifically to the antiparallel basket conformation of G4s (48). Circular dichroism analyses show that TMS induces structural changes consistent with G4 formation in the PKD1 homoPu sequence but not homoPy, $(CTG)_n$, or $(CAG)_n$ sequences (27). At elevated dosage, TMS also inhibits telomerase activity and reduces the growth of certain tumor cell types (49–52). Here, we find that a low dose of TMS dramatically increases the frequency of ectopic site DSBs in $(Pu)_{78}$ cells but not in $(Py)_{78}$ cells.

Long-read sequencing of inverse PCR (iPCR) products show that TMS stabilization of the $(Pu)_{78}$ G4 structure results in precise deletion of the repeat tract in a population of cells and hypermutation in the flanking DNA. This instability resembles the sister chromatid mutagenesis occurring during BIR. In other $(Pu)_{78}$ cells, TMS treatment resulted in large deletions (>3–4 kb) of the ectopic site, hypermutagenesis, and nonallelic chromosomal translocations, similar to the effects of MMBIR. TMS also induced instability at G4 consensus sites flanking the ectopic Pu/Py repeat. Knockdown of the Mus81 resolvase in $(Pu)_{78}$ cells significantly decreased replication-dependent DNA DSBs at the ectopic site.

To determine whether the G4 structure or the H3 structure of the Pu/Py microsatellite was responsible for instability under nonperturbed conditions, we constructed clonal cell lines containing versions of the Pu/Py tract that could only form H3 or G4 structures. Our results indicate that stable H3 and G4 DNA constructs are fragile in the absence of exogenous replication stress. We conclude that H3 and G4 DNA structures that are stable under perturbed and nonperturbed conditions lead to genomic instability with characteristics of BIR.

Results

TMS induces DSBs at the $(Pu)_{78}$ ectopic site

The ectopic site constructs integrated by FLP recombinase are shown in Figure 1. All constructs were placed at the same ectopic site, alongside a copy of the 2.4 kb c-myc core replication origin (53–58). The ectopic site c-myc origin displays the chromatin structure, replication protein binding, and early S-phase replication firing pattern of the endogenous c-myc origin (26–28, 59, 60). The $(Pu)_{78}$ repeat is in the lagging strand template for replication from the c-myc origin. dTomato and eGFP reporter genes are situated to allow ectopic site DSBs to be detected by flow cytometry. Three Alu/IVS elements derived from the UBE2T locus provide targets for homology-dependent recombination (61).

$(Pu)_{78}$ and $(Py)_{78}$ cells (Fig. 1*A*) were untreated or treated with TMS and analyzed by flow cytometry (Fig. 2*A*). As measured by separation of the dTomato and eGFP reporter signals, TMS treatment had no visible effect on instability at the ectopic site in $(Py)_{78}$ cells (Fig. 2, *B* and *C*). Consistent with previous results (27), TMS treatment reproducibly induced a dramatic loss of dTomato fluorescence in $(Pu)_{78}$ cells (Fig. 2, *D* and *E*), in line with earlier circular dichroism analysis of the PKD1 tract treated with TMS, which confirmed the formation of G4 structures (27). We conclude that stabilization of G4 structures in the $(Pu)_{78}$ replication orientation leads to DSBs at the ectopic site.

These results also show that initiation at the ectopic c-myc origin is not stochastic (62), since initiations upstream and downstream of the origin would eliminate the asymmetric effect of Pu/Py replication.

Treatment of $(Pu)_{78}$ cells with TMS for approximately four cell division cycles (96 h) resulted in ca. 45% of cells losing the dTomato fluorescence. This suggested that not every cell division of the TMS-stabilized G4 resulted in the loss of the dTomato signal. When $(Pu)_{78}$ cells were incubated with TMS for 48 h to allow 1 to 2 cell cycles in the presence of the drug, we reproducibly observed a smaller, but statistically significant, 1.2% increase in the percentage of $eGFP^+$, $dTomato^-$ (“*green*”)

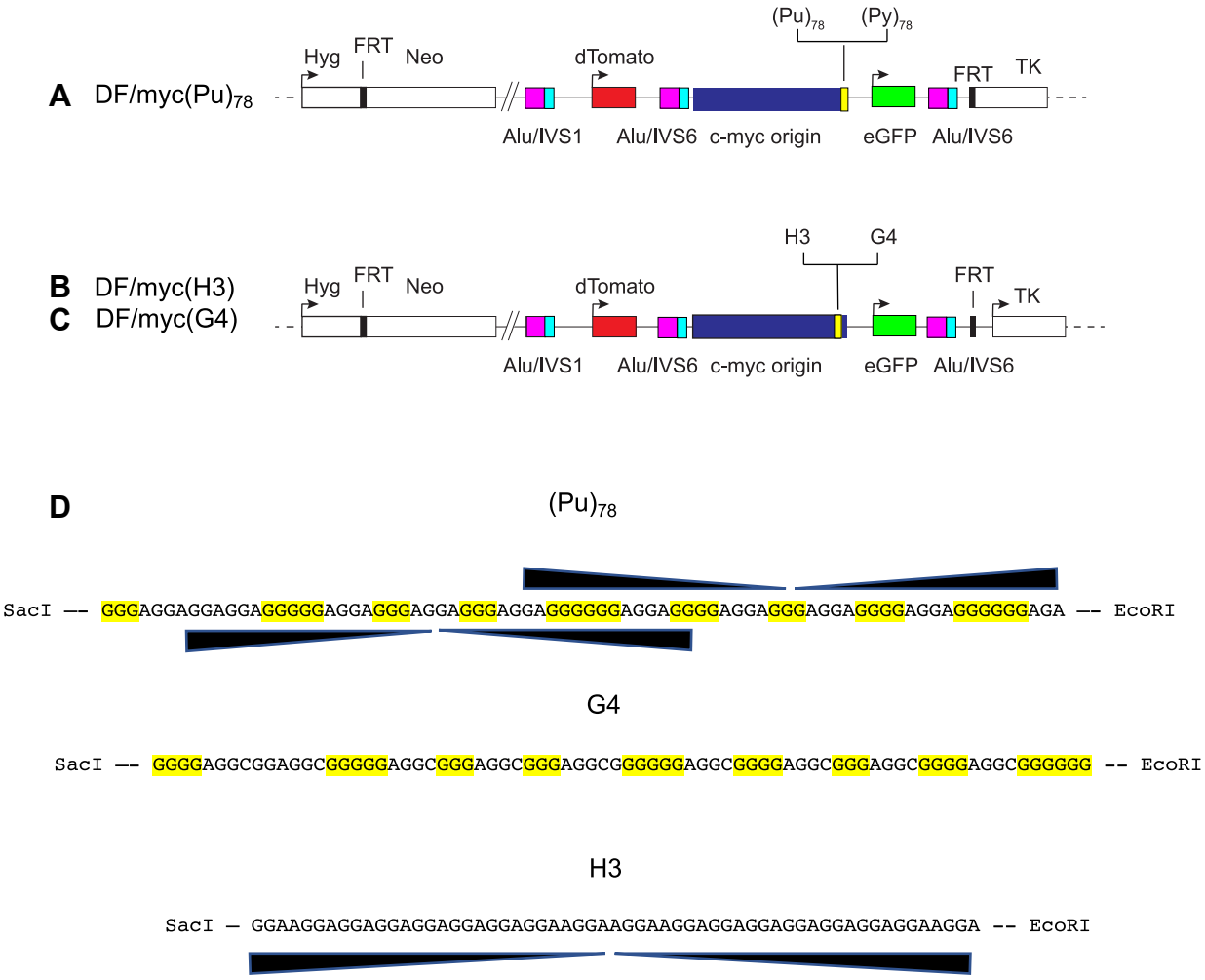

Figure 1. Ectopic site maps. *A*, schematic map of the $(Pu/Py)_{78}$ ectopic site integrant constructs. *B* and *C*, schematic maps of the H3 and G4 ectopic site integrant constructs. *D*, input DNA sequences. G4, G-quadruplex; H3, triplex; Pu/Py, homopurine/homopyrimidine.

cells (Figs. 2, *F*, *H* and *K* and S1, $p \leq 0.013$). These results indicate that multiple recombinations at the ectopic site G4 occur before loss of the dTomato signal. Strikingly, as in the case of the ectopic $(CTG)_{100}$ microsatellite (27), knockdown of Mus81 significantly reduced the percentage of green ($eGFP^+$, $dTomato^-$) cells in cultures treated or untreated with TMS (Fig. 2, *G* and *I–K*).

Single-molecule analysis of $(Pu)_{78}$ DSBs

To characterize the effects of G4-mediated instability, DNA was isolated from TMS-treated $(Pu)_{78}$ cells, digested with Xba1 and ligated at low DNA concentration to induce intramolecular circularization (63). Head-to-head PCR primers in the ectopic site were used to amplify the *circles* by iPCR (Fig. 3*A*), and the products were analyzed by PacBio long-read high-fidelity circular consensus sequencing (64).

Sequencing of the iPCR products (iPCR-Seq) showed that a population of cells had precisely deleted all or part of the G4 consensus at the $(Pu)_{78}$ insert (Fig. 3, *B* and *C*). In addition, a second larger population had deleted approximately 3 kb of DNA surrounding the G4 structure to break points at G4 consensus matches in the eGFP reporter gene (Fig. 3*A*). As in the case of cleavage by Mus81 alongside $(CTG)_{100}$ hairpins (27), recurrent break points were largely confined to specific sites around the G4 consensus sequences, whereas small insertions, deletions, and single base substitutions were abundant but dispersed in the flanking DNA. The hypermutagenesis revealed by this single-molecule analysis is reminiscent of sister chromatid BIR initiated at $(CTG)_{102}$ replication fork DSBs (27).

Should a translocation occur from the ectopic site, Xba1 cleavage of the acceptor chromosome would allow circularization, ligation, and identification of the break point junction (63). Thus, iPCR-Seq also revealed translocations to nonallelic sites (Fig. 3*D*) and complex chromosome rearrangements (Fig. 3*E*) in the TMS-treated $(Pu)_{78}$ cells, characteristic of unstable D-loops formed during BIR (14, 65–67). Recombination on both the upstream and downstream sides of the $(Pu)_{78}$ DSB supports the view that the $(Pu)_{78}$ DSB is double ended, similar to the DSBs at the $(CTG)_{100}$ ectopic site (27).

We note as well that the rate of mutation can change abruptly within a single translocation read (*e.g.*, r62783876, r54198608, r27918549) and between reads covering the same

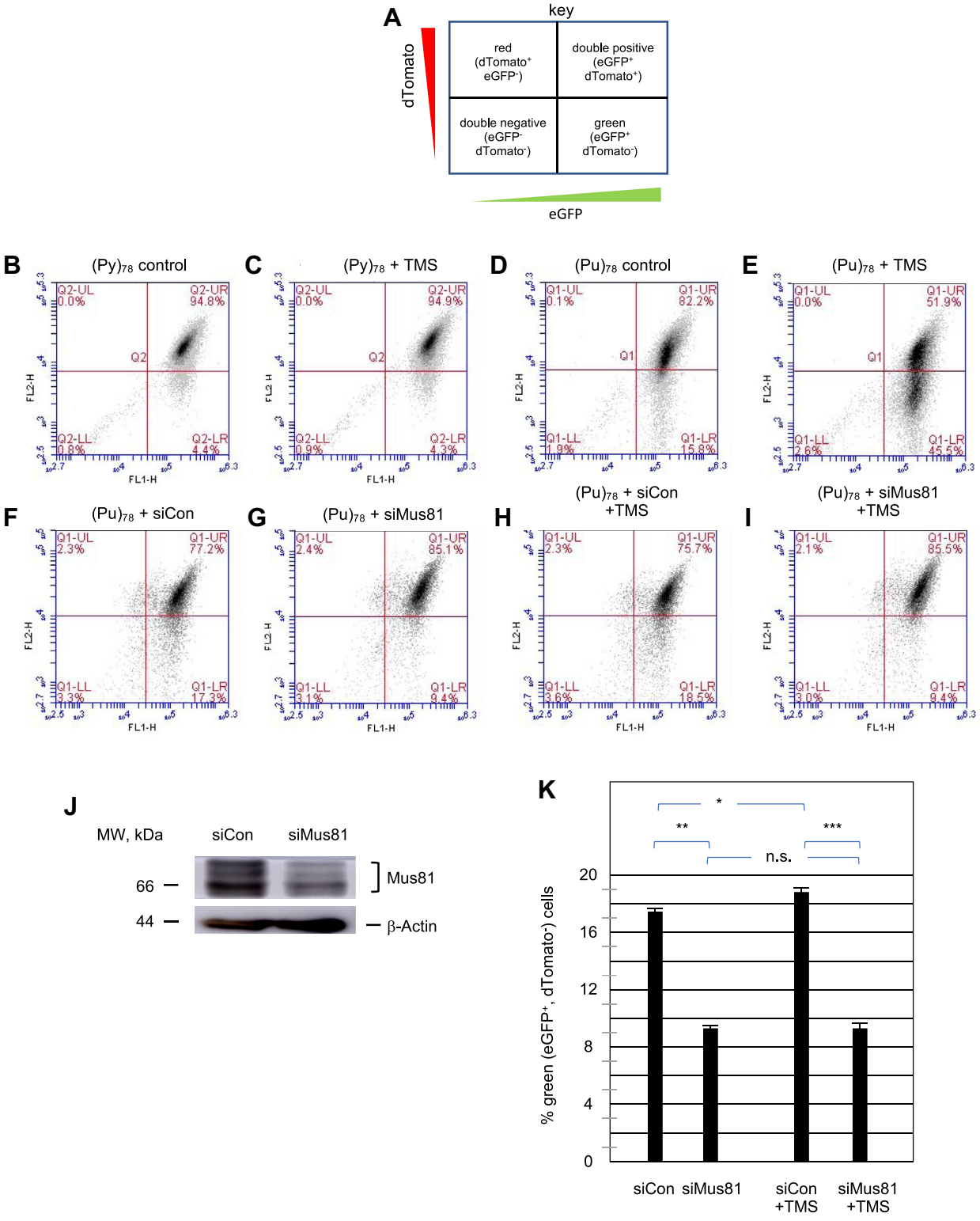

Figure 2. Orientation-dependent sensitivity of the $(Pu)_{78}$ structure to telomestatin (TMS). *A*, flow cytometry key. Flow cytometry of (*B*) $(Py)_{78}$ control cells. *C*, $(Py)_{78}$ cells treated with TMS. *D*, $(Pu)_{78}$ control cells. *E*, $(Pu)_{78}$ cells treated with TMS. *F*, $(Pu)_{78}$ cells treated with siCon. *G*, $(Pu)_{78}$ cells treated with siMus81. *H*, $(Pu)_{78}$ cells treated with siCon and TMS. *I*, $(Pu)_{78}$ cells treated with siMus81 and TMS. *J*, Western blot. *K*, effect of siMus81 on the level of *green* ($eGFP^+$, $dTomato^-$) cells, +/− SD, n = 3, two-tailed Student's *t* test. **p* = 0.013; ***p* = 0.001; ****p* = 0.0001. ns, not significant; Pu, purine; Py, pyrimidine.

sister chromatid sequence (*e.g.*, r41419292, r43713012) suggesting that the composition of the BIR replisome may be dynamic. Consistent with MMBIR, 3 to 7 bp microhomology was detected at six of seven translocation junctions (Table 1).

Mutagenesis of dTomato and eGFP flanking DNA

It is possible that loss of dTomato or eGFP signals could occur because of mutagenesis of the reporter genes. Sequencing of dTomato (Fig. 4, *A* and *B*) and eGFP (Fig. 5,

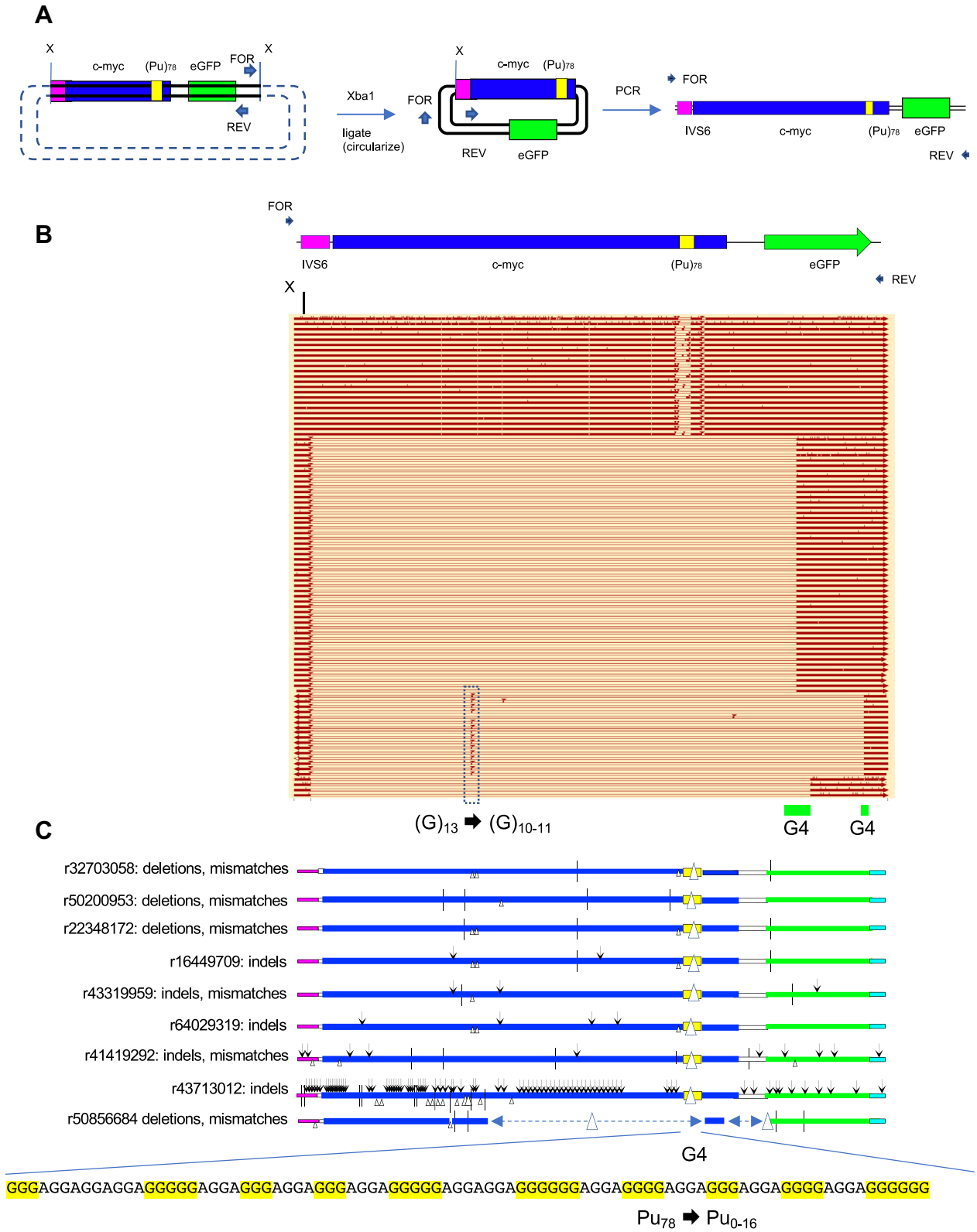

Figure 3. Inverse PCR (iPCR) analysis of $(Pu)_{78}$ instability. *A*, iPCR schematic. *B*, summary of PacBio sequencing results, aligned with an orientation map. *Open bars*, deletions; *arrowheads*, insertions; *vertical ticks*, mismatches; G4, G-quadruplex consensus sequences. *C*, expanded examples of mutagenesis results. *Arrowheads*, insertions; *triangles*, deletions; and *vertical ticks*, mismatches. The G4 sequence of the input $(Pu)_{78}$ cassette is shown below. *D*, chromosome translocations at the $(Pu)_{78}$ site. *E*, complex chromosome hopping at the ectopic site. Read designations (r) are indicated. *Dashed arrows* indicate the 5′ to 3′ polarity of the read. Pu, purine.

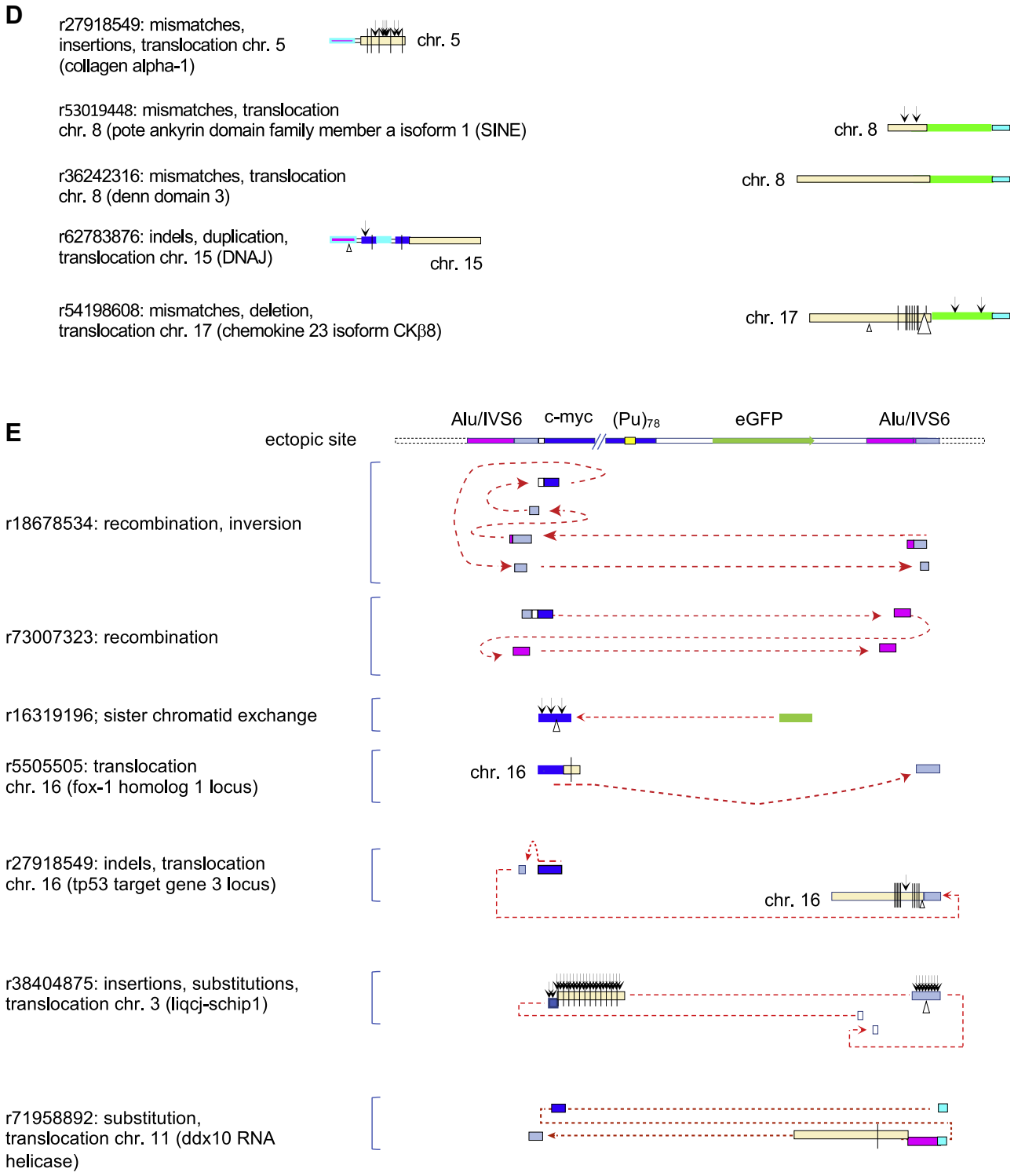

Figure 3. *Continued*

A and *B*) (standard) PCR products showed that hypermutagenesis extended into these reporter genes over a distance of more than 5 to 6 kb. In both dTomato and eGFP genes, G4 and mononucleotide repeats were hotspots for mutation during homology-mediated sister chromatid exchange. Although these results imply that dTomato and eGFP fluorescence should be lost because of mutagenesis, the preferential appearance of green cells suggests that the acentric dTomato end of the chromosome is less frequently recombined or more often lost. A similar result has been observed at ectopic site unstable $(CTG)_{100}$ repeats, where it was shown that BIR results in loss of dTomato fluorescence, whereas eGFP was lost by single-strand annealing repair of a clean I-Sce1 DSB (27).

Table 1
Break point junction homology

r27918549: mismatches, chromosome 5 (collagen alpha-1)	CCT
r53019448: chromosome 8 (POTE ankyrin domain family member A isoform 1) (SINE)	TCT
r36242316: chromosome 8 (DENN domain 3)	TGG
r62783876: chromosome 15 (DNAJ)	AGGCC
r27918549: chromosome 16 (SINE)	GCA
r5505505: chromosome 16 (Fox-1 homology region)	No homology
r54198608: chromosome 17 (chemokine 23 isoform CKb8)	AAAGCTT

Again, the frequency of mutation varies substantially between individual reads.

Instability of H3 and G4 repeats

The aforementioned results indicate that the PKD1 tract undergoes DSBs when the G4 forming sequence is stabilized by TMS. However, since both H3 and G4 structures are possible in this repeat under nonperturbed conditions,

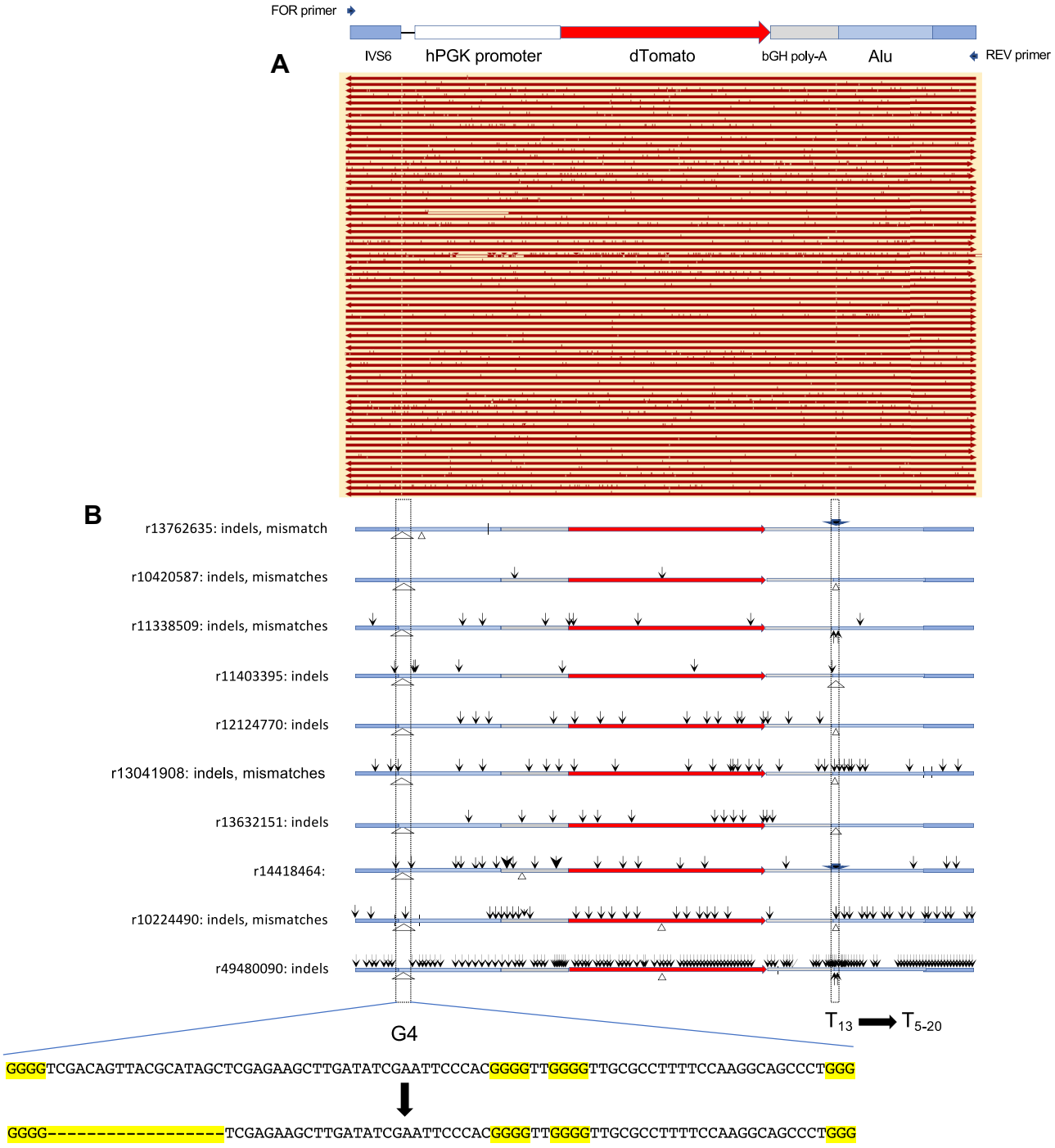

Figure 4. Mutagenesis of the dTomato region upstream of $(Pu)_{78}$. *A*, summary of PCR sequencing results, aligned with an orientation map. *B*, expanded examples of mutagenesis results. Pu, purine.

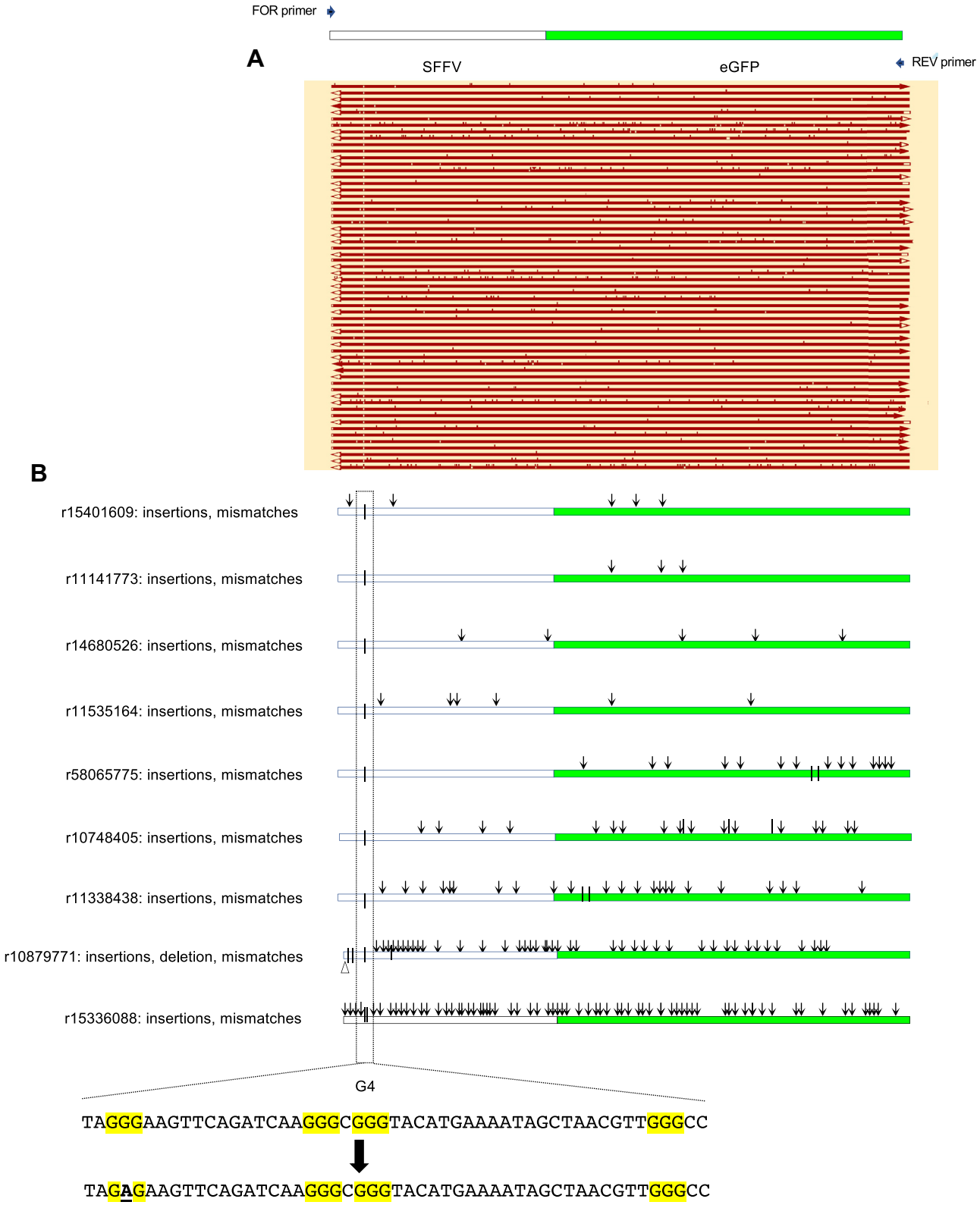

Figure 5. Mutagenesis of the eGFP region downstream of $(Pu)_{78}$. *A*, summary of PCR sequencing results, aligned with an orientation map. *B*, expanded examples of mutagenesis results. eGFP, enhanced GFP; Pu, purine.

we decided to construct cell lines containing derivatives of the PKD1 repeat, which could only form H3 or G4 structures (Fig. 1, *B* and *C*). The H3 and G4 sequences were cloned in place of the G4 prone sequence (Pu27) (30) in the natural c-myc origin to avoid spreading of the G4 structure.

The H3 and G4 cell lines revealed multiple differences. Two clones of the G4 repeat cell lines (#1 and #6) displayed different flow cytometry patterns under nonstressed conditions (Fig. 6, *A* and *C*). Surprisingly, G4 clone #1 did not lose dTomato or eGFP to produce green or double-negative cells following TMS treatment (Fig. 6*B*); similarly, G4 clone #6 did not lose eGFP to produce double-negative cells (Fig. 6*D*). As expected, the H3 cells did not show sensitivity to TMS (Fig. 6, *E* and *F*) presumably because they do not contain an ectopic G4 consensus.

DNA sequencing of the short PCR products (Fig. 7*A*) across the H3 and G4 inserts revealed that G4 clone #1 had substituted and deleted G4 consensus nucleotides, and that the entire G4 cassette had been deleted in G4 clone #6, while no mutations were detected in the H3 repeat (Fig. 7*B*). The absence of perturbations in the flow cytometry patterns of G4 clone #1 and the H3 clone suggested that these inserts might not have caused ectopic site instability. However, PCR analysis of the H3 clone, and small-pool PCR analysis of G4 clone #1 and G4 clone #6 (long PCR, Fig. 7*D*) showed that

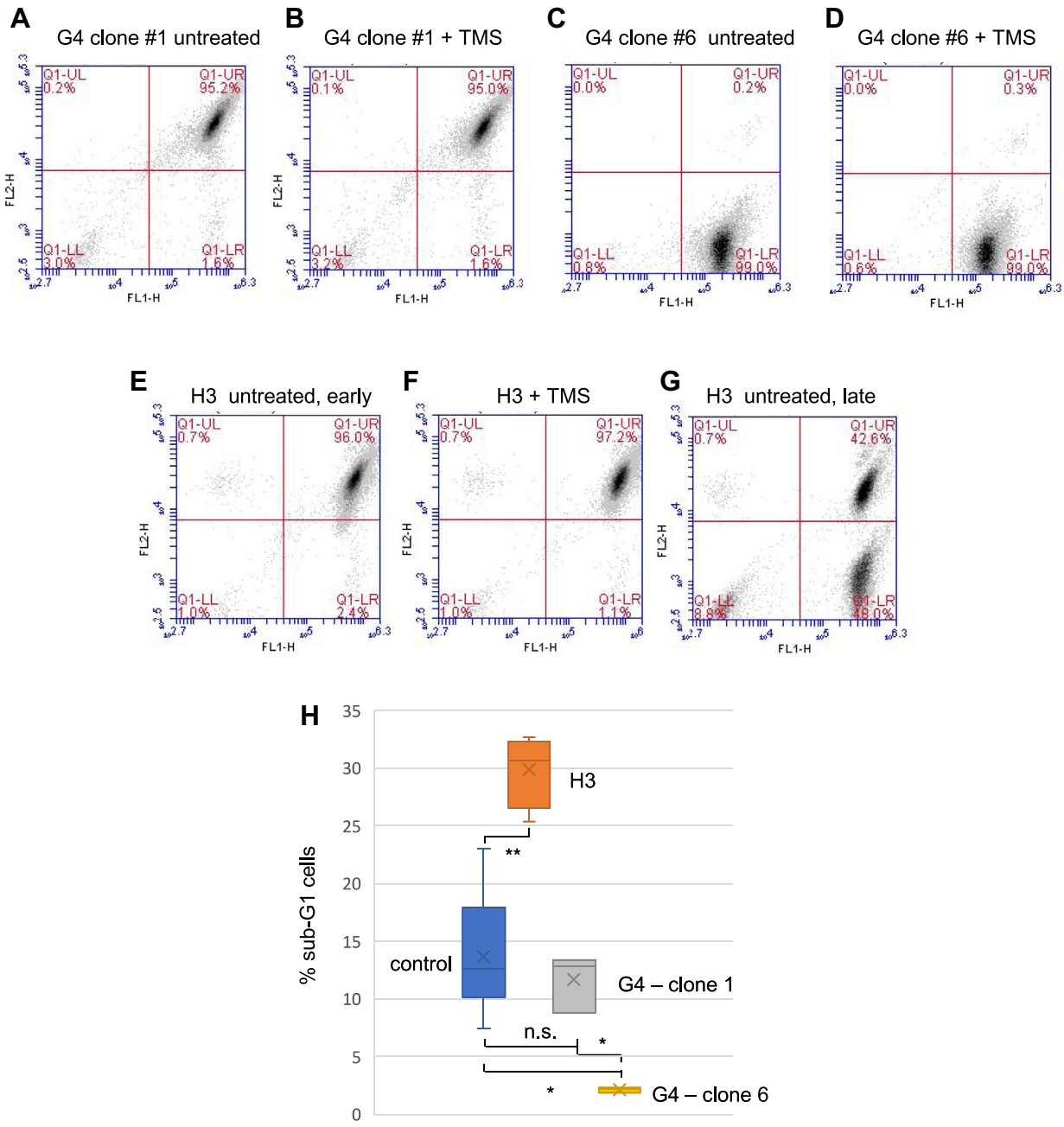

Figure 6. Insensitivity of H3 and G4 clones to TMS. Flow cytometry of (*A*) G4 clone #1 untreated cells; (*B*) G4 clone #1 cells treated with TMS; (*C*) G4 clone #6 untreated cells; (*D*) G4 clone #6 cells treated with TMS; (*E*) H3 untreated cells, early passage; (*F*) H3 cells treated with TMS; (*G*) H3 untreated cells, late passage; (*H*) percent sub-G1 cells in control (minus ectopic site non–B-DNA), H3, and G4 cells. $*p < 0.005$; $**p < 0.0002$. n ≥ 3. G4, G-quadruplex; H3, triplex; TMS, telomestatin.

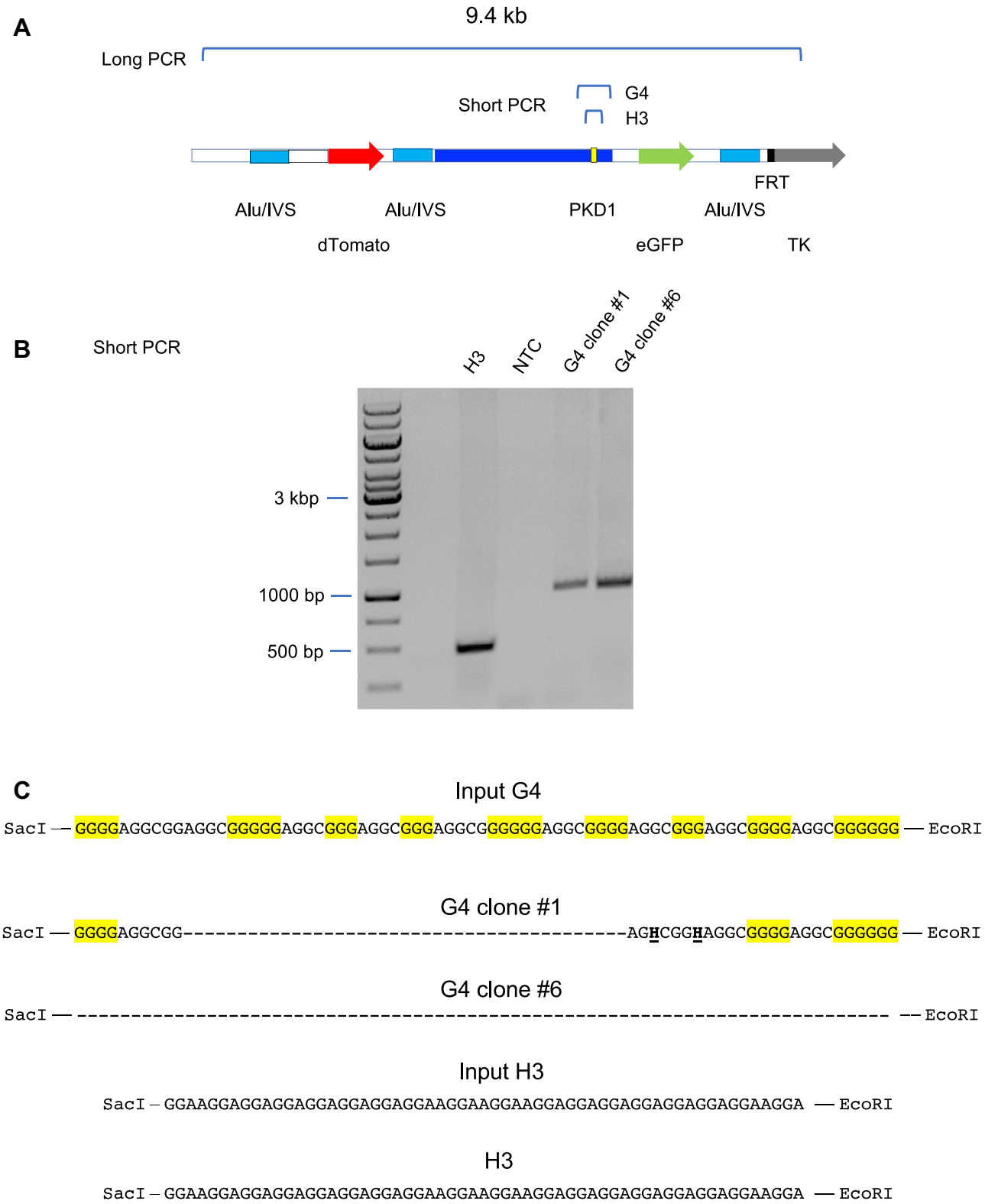

Figure 7. H3 cells and G4 ectopic sites are unstable under nonstressed growth conditions. *A*, map of short and long PCR templates. *B*, short PCR confirmation of H3 and G4 inserts. *C*, sequences of input and outcome H3 and G4 cassettes. H, non-G nucleotide. *D*, PCR analysis of H3, G4 recombination. *Arrows*, expected PCR product size (9.4 kb). G4, G-quadruplex; H3, triplex; NTC, no template control.

significant deletions had occurred at both the H3 and G4 ectopic sites during clonal outgrowth under nonperturbed conditions.

The long PCR results were confirmed to be specific for the ectopic site by the negative control amplification of genomic DNA of HeLa/406 acceptor cells containing

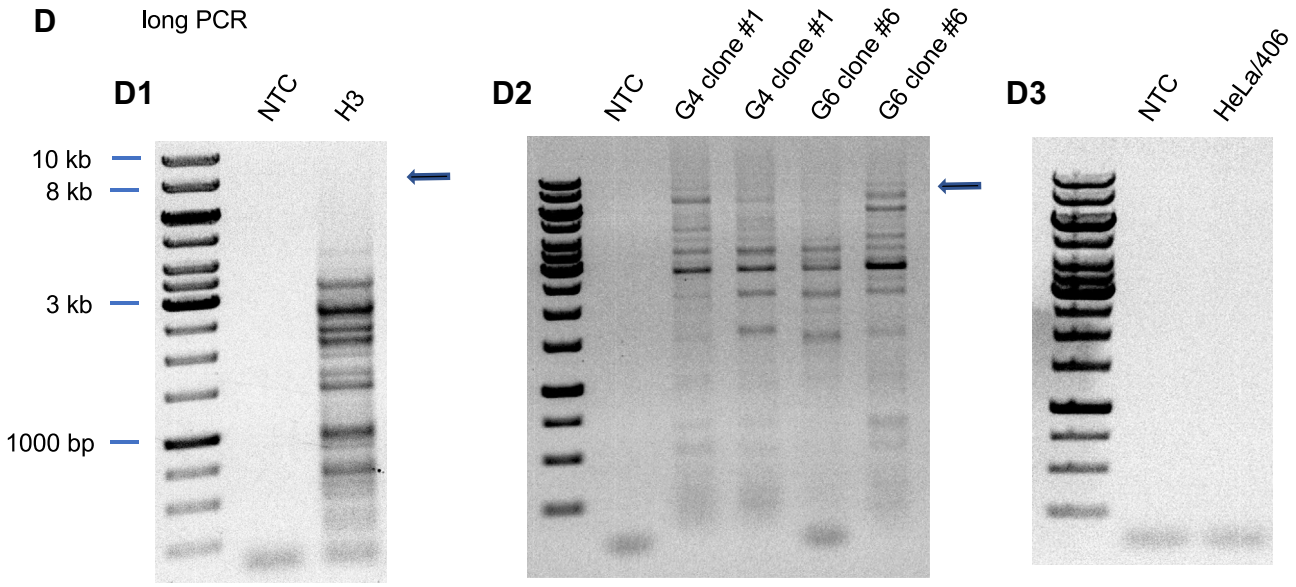

Figure 7. *Continued*

an empty ectopic site (Fig. 7*D*.3). We conclude that G4 clones #1 and #6 had undergone distinct mutations prior to clonal growth and that the insensitivity of these clones to TMS is due to the loss of the G4 consensus repeat.

The retention of the unmutagenized H3 insert was surprising in view of the iPCR results demonstrating rearrangements in the H3 ectopic site. However, quantitation by flow cytometry (Fig. 6*H*) showed that the H3 cells generated more sub-G1 cells than either the G4 clones or control cells that did not contain non–B-DNA at the ectopic site. In addition, prolonged growth (3 months) of the H3 cells led to the appearance of >40% of green (eGFP$^+$, dTomato$^-$) cells (Fig. 6*G*), consistent with the view that multiple DSBs may be required before loss of either reporter gene signal. We conclude that both the H3 and G4 constructs lead to genomic instability and that different non–B-DNAs respond differently to replication stress.

Discussion

G4-prone sequences are abundant in mammalian DNA and have been implicated in having positive or negative effects on replication, transcription, and genome stability. In this study, we show that stabilization of G4 structures in the PKD1 IVS21 Pu/Py repeat by the TMS ligand leads to replication-dependent DSBs, hypermutagenesis, and gross chromosomal rearrangements associated with unstable D-loops characteristic of BIR. In cells that proliferated after TMS treatment, the PKD1 G4 consensus sequence was eliminated either by precise excision or kilobase length deletion. In either situation, deletion break points were nonrandomly located at G4-forming sequences. We note as well that G4 consensus sequences and mononucleotide repeats in the flanking DNA were also hotspots for mutation and that deletions at G4 upstream or downstream flanking sites were oriented at consensus sequences with Pus in the lagging strand template. Mutagenesis extended for >2 to 3 kb upstream and downstream of the ectopic site break points, suggesting that fork restart by homology-dependent sister chromatid exchange could initiate at both ends of the DSB.

$(Pu)_{78}$ cells displayed greater instability than $(Py)_{78}$ cells under nonperturbed conditions (*cf.* Fig. 2, *B* and *D*, and (27)). Cells containing a construct designed only to form G4 structures were also unstable under nonperturbed conditions, and G4 clones #1 and #6 grown out from these populations displayed deletions of the G4 consensus sequences similar to what was seen with TMS-treated $(Pu)_{78}$ cells. We conclude that TMS ligand binding, or sequence optimization and spontaneous G4 formation, leads to stable G4 structures, and that mutation of the G4 consensus is responsible for the insensitivity of G4 clone #1 and G4 clone #6 cells to TMS and intrinsic replication stress. Similarly, an extended perfect mirror repeat H3-prone sequence appears to have increased the efficiency and instability of the H3 structure.

In contrast to the G4 cassette, we did not detect deletion of the H3 structure in early passage cells, despite dramatic recombinations at the H3 ectopic site. It is worth noting that the PCR results of Figure 7, *B* and *D* cannot be compared quantitively, since the short PCR is sufficiently sensitive to detect the H3 template retained in a small fraction of cells. Retention of the unaltered H3 cassette may occur because not all recombinations remove the ectopic microsatellite or because of structural differences in remodeling at G4 and H3 replication forks. The former explanation is supported by the observation that prolonged outgrowth is required to observe loss of the dTomato signal in H3 cells. Further next-generation sequencing analyses will address these questions.

In a broader sense, these results indicate that fork stalling at different non–B-DNAs can result in different forms of

mutagenesis. Thus, different forms of replication stress may result in heterogeneous patterns of instability at a single non–B-DNA, and a single form of replication stress may induce multiple patterns of instability at different non–B-DNAs.

Experimental procedures

Cell lines

HeLa/406 cells containing the FLP recombinase target site and construction of ectopic cell lines have been described previously (26, 27, 47, 59, 68–70). The PKD1 insert in (Pu/Py) cells was cloned from the pJB4 plasmid (71). H3 and G4 inserts were designed to conform to the requirements of H3 mirror-repeat symmetry and G4 consensus matches and synthesized by Integrated DNA Technologies.

HeLa/406 acceptor cells contain a single FLP recombinase target site (69, 72). DF/myc cell lines were derived by cotransfecting HeLa/406 cells with dual fluorescence donor plasmids (27, 61) and the FLP recombinase expression vector pOG44 (73). Cells were maintained on Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin, and 5% CO_2 at 37 °C. For TMS treatment, DF/myc(Pu)$_{78}$ cells were plated in 6-well plates at 30% confluency. Twenty-four hours later, they were treated with TMS at a final concentration of 0.5 μM. The treatment lasted for 4 days; this concentration of TMS did not significantly inhibit cell division, consistent with previous work (74). During the course of the treatment, TMS was replenished at 48 h. After the 4 day treatment ended, the cells were allowed to recover for 4 days. Following this, cells were harvested to extract genomic DNA for PCR, iPCR, or flow cytometry. Treatment and recovery of cells was in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin, and 5% CO_2 at 37 °C.

PCR and iPCR

iPCR was performed using genomic DNA from DF/myc(Pu)$_{78}$ cells. About 1.5 μg DNA was digested with XbaI (New England Biolabs) for 1 h at 37 °C. After a 20 min heat inactivation step at 65 °C, 30 μl of the reaction (0.4–1 ng/μl) was used for ligation with T4 DNA ligase (New England Biolabs) overnight at 4 °C. The ligation reaction products were cleaned using E.Z.N.A. Cycle Pure kits (Omega). iPCR primers used for amplification were as follows: Nhe primer forward-BC-1 (5′-AGCTAAGCTTGCCTTGAGTGCTTC-3′) and eGFP primer reverse-BC-1 (5′-AGCTGTCCATGCCGAGAGTGATC-3′). Q5 HotStart polymerase (New England Biolabs) was used per manufacturer's instructions with cycling conditions as follows: initial denaturation at 98 °C for 30 s. Touchdown PCR started with an annealing temperature of 66 °C for 15 s with a reduction of 0.5 °C for the first ten cycles. Subsequent cycles utilized 61 °C for 15 s. This was followed by an extension time of up to 5 min at each cycle. Final extension time was for 5 min followed by a hold at 4 °C.

Long PCR was performed using Lac—forward (5′-CTTCAAATCCGACCCGTAGA-3′) and TK—reverse-1 (5′-GGGTATCGACAGAGTGCCAG-3′) for H3, G4 clones #1 and #6, and HeLa/406 cells. Q5 HotStart polymerase was used per manufacturer's instructions for 50 μl reactions using 1 to 200 ng template. Cycling conditions were as follows: initial denaturation at 98 °C for 30 s. Touchdown PCR started with an annealing temperature of 66 °C for 15 s with a reduction of 0.5 °C for the first ten cycles. Subsequent cycles utilized 61 °C for 15 s. This was followed by an extension time of up to 5 min at each cycle. Final extension time was for 5 min and then a hold temperature of 4 °C. PCR products were electrophoresed on a 0.4% agarose gel (Invitrogen) to verify the size of the recombination products. Appropriate no template controls were also run.

Short PCR was performed using multiple cloning site (MCS)—forward (5′-GGTTCCCAAAGCAGAGGG-3′) and MCS—reverse (5′-TGGAAACATCTGATGGGTCTT-3′) primers for H3 cells and BstZ17I NEBuilder—forward (5′-CATAACGCGCTCTCCAAGTA-3′) and MCS—reverse (5′-TGGAAACATCTGATGGGTCTT-3′) for G4 clones #1 and #6. Cycling conditions were as aforementioned.

DNA sequence analysis

Amplicon EZ sequencing was performed on standard PCR products less than 1 kb, whereas Pacific Biosciences HiFi sequencing was performed on longer amplicons (iPCR, eGFP, and dTomato) at GeneWiz. Circular consensus sequences were trimmed, and contaminants were removed by screening reads for sequences matching the amplification primers or sequences immediately adjacent to the primer binding sites. Exact sequence duplicates were removed, and the remaining unique sequences were used for mapping. Trimmed Amplicon EZ sequences were used directly for mapping. BWA-MEM (Burrows-Wheeler Aligner Maximal Exact Match) (75) was used to map reads against their respective ectopic site reference sequences or the human genome (hg38). Chimeric and nonchimeric reads were identified within the resulting .sam file and examined manually using BLAST or Muscle in SnapGene (Insightful Science).

Data availability

Data will be shared upon request to michael.leffak@wright.edu.

Supporting information—This article contains supporting information.

Acknowledgments—We gratefully acknowledge the support of the National Institutes of Health (grant no.: GM122976; to M. L.). We are grateful to Michael Bottomley of the WSU Statistics Center for the calculation of genomic G4 probabilities and to Dale Schiff for assistance with figure preparation.

Author contributions—S. D. R. Jr, R. Y. G., D. C. H., F. J. D. IV, N. Z., M. S., V. A., R. S., K. S.-y., and M. L. conceptualization; S. D. R. Jr, R. Y. G., D. C. H., F. J. D. IV, N. Z., M. S., V. A., R. S., K. S.-y., and M. L. methodology; S. D. R. Jr, R. Y. G., and D. C. H. software; S. D. R. Jr, R. Y. G., D. C. H., and M. L. validation; S. D. R. Jr, R. Y. G., D. C. H., F. J. D. IV, N. Z., M. S., V. A., R. S., K. S.-y., and M. L.

writing–review and editing; M. L. project administration; M. L. funding acquisition.

Funding and additional information—V. A. and R. S. were supported by the Wright State University Biomedical Sciences PhD program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ATR, ataxia telangiectasia and Rad3-related protein; BIR, break-induced replication; CHK1, checkpoint kinase 1; DDR, DNA damage response; DSB, double-strand break; eGFP, enhanced GFP; G4, G-quadruplex; H3, triplex; iPCR, inverse PCR; MMBIR, microhomology-mediated BIR; PKD1, polycystic kidney disease type 1; Pu/Py, homopurine/homopyrimidine; TMS, telomestatin.

References

1. Kit, S. (1961) Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *J. Mol. Biol.* **3**, 711–716
2. Khristich, A. N., and Mirkin, S. M. (2020) On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170
3. Krasilnikova, M. M., and Mirkin, S. M. (2004) Replication stalling at Friedreich's ataxia (GAA)n repeats *in vivo*. *Mol. Cell. Biol.* **24**, 2286–2295
4. Spivakovsky-Gonzalez, E., Polleys, E. J., Masnovo, C., Cebrian, J., Molina-Vargas, A. M., Freudenreich, C. H., and Mirkin, S. M. (2021) Rad9-mediated checkpoint activation is responsible for elevated expansions of GAA repeats in CST-deficient yeast. *Genetics* **219**, iyab125
5. Sundararajan, R., and Freudenreich, C. H. (2011) Expanded CAG/CTG repeat DNA induces a checkpoint response that impacts cell proliferation in Saccharomyces cerevisiae. *PLoS Genet.* **7**, e1001339
6. Balakumaran, B. S., Freudenreich, C. H., and Zakian, V. A. (2000) CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in Saccharomyces cerevisiae. *Hum. Mol. Genet.* **9**, 93–100
7. Sundararajan, R., Gellon, L., Zunder, R. M., and Freudenreich, C. H. (2010) Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in Saccharomyces cerevisiae. *Genetics* **184**, 65–77
8. Usdin, K., House, N. C., and Freudenreich, C. H. (2015) Repeat instability during DNA repair: Insights from model systems. *Crit. Rev. Biochem. Mol. Biol.* **50**, 142–167
9. Yang, J., and Freudenreich, C. H. (2007) Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner. *Gene* **393**, 110–115
10. Sakofsky, C. J., and Malkova, A. (2017) Break induced replication in eukaryotes: Mechanisms, functions, and consequences. *Crit. Rev. Biochem. Mol. Biol.* **52**, 395–413
11. Kononenko, A. V., Ebersole, T., Vasquez, K. M., and Mirkin, S. M. (2018) Mechanisms of genetic instability caused by (CGG)n repeats in an experimental mammalian system. *Nat. Struct. Mol. Biol.* **25**, 669–676
12. Sakofsky, C. J., Roberts, S. A., Malc, E., Mieczkowski, P. A., Resnick, M. A., Gordenin, D. A., and Malkova, A. (2014) Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* **7**, 1640–1648
13. Malkova, A., and Haber, J. E. (2012) Mutations arising during repair of chromosome breaks. *Annu. Rev. Genet.* **46**, 455–473
14. Donnianni, R. A., Zhou, Z. X., Lujan, S. A., Al-Zain, A., Garcia, V., Glancy, E., Burkholder, A. B., Kunkel, T. A., and Symington, L. S. (2019) DNA polymerase delta synthesizes both strands during break-induced replication. *Mol. Cell* **76**, 371–381.e374
15. Naiman, K., Campillo-Funollet, E., Watson, A. T., Budden, A., Miyabe, I., and Carr, A. M. (2021) Replication dynamics of recombination-dependent replication forks. *Nat. Commun.* **12**, 923
16. Lambert, S., Watson, A., Sheedy, D. M., Martin, B., and Carr, A. M. (2005) Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. *Cell* **121**, 689–702
17. Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* **87**, 493–506
18. Gazy, I., Hayward, B., Potapova, S., Zhao, X., and Usdin, K. (2019) Double-strand break repair plays a role in repeat instability in a fragile X mouse model. *DNA Repair (Amst.)* **74**, 63–69
19. Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., Hunter, K., Stanton, V. P., Thirion, J. P., Hudson, T., Sohn, R., Zemelman, B., Snell, R. G., Rundle, S. A., Crow, S., *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* **68**, 799–808
20. Cho, D. H., and Tapscott, S. J. (2007) Myotonic dystrophy: Emerging mechanisms for DM1 and DM2. *Biochim. Biophys. Acta* **1772**, 195–204
21. Bahrambeigi, V., Song, X., Sperle, K., Beck, C. R., Hijazi, H., Grochowski, C. M., Gu, S., Seeman, P., Woodward, K. J., Carvalho, C. M. B., Hobson, G. M., and Lupski, J. R. (2019) Distinct patterns of complex rearrangements and a mutational signature of microhomeology are frequently observed in PLP1 copy number gain structural variants. *Genome Med.* **11**, 80
22. Verdin, H., D'Haene, B., Beysen, D., Novikova, Y., Menten, B., Sante, T., Lapunzina, P., Nevado, J., Carvalho, C. M., Lupski, J. R., and De Baere, E. (2013) Microhomology-mediated mechanisms underlie non-recurrent disease-causing microdeletions of the FOXL2 gene or its regulatory domain. *PLoS Genet.* **9**, e1003358
23. Zhang, F., Khajavi, M., Connolly, A. M., Towne, C. F., Batish, S. D., and Lupski, J. R. (2009) The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat. Genet.* **41**, 849–853
24. Bacolla, A., and Wells, R. D. (2009) Non-B DNA conformations as determinants of mutagenesis and human disease. *Mol. Carcinog.* **48**, 273–285
25. Costantino, L., Sotiriou, S. K., Rantala, J. K., Magin, S., Mladenov, E., Helleday, T., Haber, J. E., Iliakis, G., Kallioniemi, O. P., and Halazonetis, T. D. (2014) Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* **343**, 88–91
26. Liu, G., Chen, X., Bissler, J. J., Sinden, R. R., and Leffak, M. (2010) Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. *Nat. Chem. Biol.* **6**, 652–659
27. Gadgil, R. Y., Romer, E. J., Goodman, C. C., Rider, S. D., Jr., Damewood, F. J., Barthelemy, J. R., Shin-Ya, K., Hanenberg, H., and Leffak, M. (2020) Replication stress at microsatellites causes DNA double-strand breaks and break-induced replication. *J. Biol. Chem.* **295**, 15378–15397
28. Liu, G., Myers, S., Chen, X., Bissler, J. J., Sinden, R. R., and Leffak, M. (2012) Replication fork stalling and checkpoint activation by a PKD1 locus mirror repeat polypurine-polypyrimidine (Pu-Py) tract. *J. Biol. Chem.* **287**, 33412–33423
29. Belotserkovskii, B. P., De Silva, E., Tornaletti, S., Wang, G., Vasquez, K. M., and Hanawalt, P. C. (2007) A triplex-forming sequence from the human c-MYC promoter interferes with DNA transcription. *J. Biol. Chem.* **282**, 32433–32441
30. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11593–11598
31. Todd, A. K., Johnston, M., and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **33**, 2901–2907

32. Guedin, A., Gros, J., Alberti, P., and Mergny, J. L. (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* **38**, 7858–7868
33. Lyu, J., Shao, R., Kwong Yung, P. Y., and Elsasser, S. J. (2022) Genome-wide mapping of G-quadruplex structures with CUT&Tag. *Nucleic Acids Res.* **50**, e13
34. Henderson, A., Wu, Y., Huang, Y. C., Chavez, E. A., Platt, J., Johnson, F. B., Brosh, R. M., Jr., Sen, D., and Lansdorp, P. M. (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* **42**, 860–869
35. Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186
36. Lam, E. Y., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* **4**, 1796
37. Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33**, 877–881
38. Hansel-Hertsch, R., Di Antonio, M., and Balasubramanian, S. (2017) DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* **18**, 279–284
39. Valton, A. L., Hassan-Zadeh, V., Lema, I., Boggetto, N., Alberti, P., Saintome, C., Riou, J. F., and Prioleau, M. N. (2014) G4 motifs affect origin positioning and efficiency in two vertebrate replicators. *EMBO J.* **33**, 732–746
40. Prorok, P., Artufel, M., Aze, A., Coulombe, P., Peiffer, I., Lacroix, L., Guedin, A., Mergny, J. L., Damaschke, J., Schepers, A., Cayrou, C., Teulade-Fichou, M. P., Ballester, B., and Mechali, M. (2019) Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat. Commun.* **10**, 3274
41. Lee, W. T. C., Yin, Y., Morten, M. J., Tonzi, P., Gwo, P. P., Odermatt, D. C., Modesti, M., Cantor, S. B., Gari, K., Huang, T. T., and Rothenberg, E. (2021) Single-molecule imaging reveals replication fork coupled formation of G-quadruplex structures hinders local replication stress signaling. *Nat. Commun.* **12**, 2525
42. Yang, D., and Hurley, L. H. (2006) Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides Nucleotides Nucleic Acids* **25**, 951–968
43. Lokanga, R. A., Kumari, D., and Usdin, K. (2021) Common threads: Aphidicolin-inducible and folate-sensitive fragile sites in the human genome. *Front. Genet.* **12**, 708860
44. De Magis, A., Manzo, S. G., Russo, M., Marinello, J., Morigi, R., Sordet, O., and Capranico, G. (2019) DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 816–825
45. Gupta, R., Sharma, S., Sommers, J. A., Jin, Z., Cantor, S. B., and Brosh, R. M., Jr. (2005) Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer. *J. Biol. Chem.* **280**, 25450–25460
46. Brosh, R. M., Jr., and Cantor, S. B. (2014) Molecular and cellular functions of the FANCJ DNA helicase defective in cancer and in Fanconi anemia. *Front. Genet.* **5**, 372
47. Chen, X., Liu, G., and Leffak, M. (2013) Activation of a human chromosomal replication origin by protein tethering. *Nucleic Acids Res.* **41**, 6460–6474
48. Kim, M. Y., Vankayalapati, H., Shin-Ya, K., Wierzba, K., and Hurley, L. H. (2002) Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular g-quadruplex. *J. Am. Chem. Soc.* **124**, 2098–2099
49. Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J. H., and Ohyashiki, K. (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: In vitro and *in vivo* studies in acute leukemia. *Oncogene* **25**, 5719–5725
50. Shalaby, T., von Bueren, A. O., Hurlimann, M. L., Fiaschetti, G., Castelletti, D., Masayuki, T., Nagasawa, K., Arcaro, A., Jelesarov, I., Shin-ya, K., and Grotzer, M. (2010) Disabling c-Myc in childhood medulloblastoma and atypical teratoid/rhabdoid tumor cells by the potent G-quadruplex interactive agent S2T1-6OTD. *Mol. Cancer Ther.* **9**, 167–179
51. Miyazaki, T., Pan, Y., Joshi, K., Purohit, D., Hu, B., Demir, H., Mazumder, S., Okabe, S., Yamori, T., Viapiano, M., Shin-ya, K., Seimiya, H., and Nakano, I. (2012) Telomestatin impairs glioma stem cell survival and growth through the disruption of telomeric G-quadruplex and inhibition of the proto-oncogene, c-Myb. *Clin. Cancer Res.* **18**, 1268–1280
52. Nakamura, T., Okabe, S., Yoshida, H., Iida, K., Ma, Y., Sasaki, S., Yamori, T., Shin-Ya, K., Nakano, I., Nagasawa, K., and Seimiya, H. (2017) Targeting glioma stem cells *in vivo* by a G-quadruplex-stabilizing synthetic macrocyclic hexaoxazole. *Sci. Rep.* **7**, 3605
53. McWhinney, C., and Leffak, M. (1988) Episomal persistence of a plasmid containing human c-myc DNA. In: Stillman, B., Kelly, T., eds. *Cancer Cells*, CSH Laboratory Press, New York, NY: 467–471
54. McWhinney, C., and Leffak, M. (1990) Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-myc gene. *Nucleic Acids Res.* **18**, 1233–1242
55. McWhinney, C., Waltz, S. E., and Leffak, M. (1995) Cis-acting effects of sequences within 2.4-kb upstream of the human c-myc gene on autonomous plasmid replication in HeLa cells. *DNA Cell Biol.* **14**, 565–579
56. Tao, L., Dong, Z., Leffak, M., Zannis-Hadjopoulos, M., and Price, G. (2000) Major DNA replication initiation sites in the c-myc locus in human cells. *J. Cell. Biochem.* **78**, 442–457
57. Trivedi, A., Waltz, S. E., Kamath, S., and Leffak, M. (1998) Multiple initiations in the c-myc replication origin independent of chromosomal location. *DNA Cell Biol.* **17**, 885–896
58. Waltz, S. E., Trivedi, A. A., and Leffak, M. (1996) DNA replication initiates non-randomly at multiple sites near the c-myc gene in HeLa cells. *Nucleic Acids Res.* **24**, 1887–1894
59. Liu, G., Bissler, J. J., Sinden, R. R., and Leffak, M. (2007) 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. *Mol. Cell. Biol.* **27**, 7828–7838
60. Liu, G., Chen, X., and Leffak, M. (2013) Oligodeoxynucleotide binding to (CTG). (CAG) microsatellite repeats inhibits replication fork stalling, hairpin formation, and genome instability. *Mol. Cell. Biol.* **33**, 571–581
61. Lewis, T. W., Barthelemy, J. R., Virts, E. L., Kennedy, F. M., Gadgil, R. Y., Wiek, C., Linka, R. M., Zhang, F., Andreassen, P. R., Hanenberg, H., and Leffak, M. (2019) Deficiency of the Fanconi anemia E2 ubiqitin conjugase UBE2T only partially abrogates Alu-mediated recombination in a new model of homology dependent recombination. *Nucleic Acids Res.* **47**, 3503–3520
62. Wang, W., Klein, K. N., Proesmans, K., Yang, H., Marchal, C., Zhu, X., Borrman, T., Hastie, A., Weng, Z., Bechhoefer, J., Chen, C. L., Gilbert, D. M., and Rhind, N. (2021) Genome-wide mapping of human DNA replication by optical replication mapping supports a stochastic model of eukaryotic replication. *Mol. Cell* **81**, 2975–2988.e6
63. Barthelemy, J., Hanenberg, H., and Leffak, M. (2016) FANCJ is essential to maintain microsatellite structure genome-wide during replication stress. *Nucleic Acids Res.* **44**, 6803–6816
64. Hiatt, S. M., Lawlor, J. M. J., Handley, L. H., Ramaker, R. C., Rogers, B. B., Partridge, E. C., Boston, L. B., Williams, M., Plott, C. B., Jenkins, J., Gray, D. E., Holt, J. M., Bowling, K. M., Bebin, E. M., Grimwood, J., *et al.* (2021) Long-read genome sequencing for the molecular diagnosis of neurodevelopmental disorders. *HGG Adv.* **2**, 100023
65. Donnianni, R. A., and Symington, L. S. (2013) Break-induced replication occurs by conservative DNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13475–13480
66. Kramara, J., Osia, B., and Malkova, A. (2017) Break-induced replication: An unhealthy choice for stress relief? *Nat. Struct. Mol. Biol.* **24**, 11–12
67. Kramara, J., Osia, B., and Malkova, A. (2018) Break-induced replication: The where, the why, and the how. *Trends Genet.* **34**, 518–531
68. Liu, G., and Leffak, M. (2012) Instability of (CTG)n*(CAG)n trinucleotide repeats and DNA synthesis. *Cell Biosci.* **2**, 7
69. Guo, J., Gu, L., Leffak, M., and Li, G. M. (2016) MutSbeta promotes trinucleotide repeat expansion by recruiting DNA polymerase beta to nascent (CAG)n or (CTG)n hairpins for error-prone DNA synthesis. *Cell Res.* **26**, 775–786

70. Barthelemy, J., Hanenberg, H., and Leffak, M. (2017) FANCJ is essential to maintain microsatellite structure genome-wide during replication stress. *Nucleic Acids Res.* **45**, 509–511
71. Patel, H. P., Lu, L., Blaszak, R. T., and Bissler, J. J. (2004) PKD1 intron 21: Triplex DNA formation and effect on replication. *Nucleic Acids Res.* **32**, 1460–1468
72. Liu, G., Malott, M., and Leffak, M. (2003) Multiple functional elements comprise a mammalian chromosomal replicator. *Mol. Cell. Biol.* **23**, 1832–1842
73. O'Gorman, S., Fox, D. T., and Wahl, G. M. (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* **251**, 1351–1355
74. Tahara, H., Shin-Ya, K., Seimiya, H., Yamada, H., Tsuruo, T., and Ide, T. (2006) G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3' telomeric overhang in cancer cells. *Oncogene* **25**, 1955–1966
75. Li, H., and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595