RNA transcripts regulate G-quadruplex landscapes through G-loop formation

Sato, K.; Lyu, J.; Berg, J. van den; Braat, D.; Cruz, V.M.; Luzón, C.N.; ... ; Knipscheer, P.

MOLECULAR BIOLOGY

RNA transcripts regulate G-quadruplex landscapes through G-loop formation

Koichi Sato[1,2]*, Jing Lyu[3,4], Jeroen van den Berg[1], Diana Braat[1,2], Victoria M. Cruz[1,2], Carmen Navarro Luzón[3,4], Joost Schimmel[2], Clara Esteban-Jurado[2], Maëlys Alemany[1], Jan Dreyer[5], Aiko Hendrikx[1], Francesca Mattiroli[5], Alexander van Oudenaarden[1], Marcel Tijsterman[2,6], Simon J. Elsässer[3,4], Puck Knipscheer[1,2]*

G-quadruplexes (G4s) are prevalent DNA structures that regulate transcription but also threaten genome stability. How G4 dynamics are controlled remains poorly understood. Here, we report that RNA transcripts govern G4 landscapes through coordinated G-loop assembly and disassembly. G-loop assembly involves activation of the ATM and ATR kinases, followed by homology-directed invasion of RNA opposite the G4 strand mediated by BRCA2 and RAD51. Disassembly of the G-loop resolves the G4 structure through DHX36-FANCJ–mediated G4 unwinding, which triggers nucleolytic incision and subsequent hybrid strand renewal by DNA synthesis. Inhibition of G-loop disassembly causes global G4 and R-loop accumulation, leading to transcriptome dysregulation, replication stress, and genome instability. These findings establish an intricate G-loop assembly-disassembly mechanism that controls G4 landscapes and is essential for cellular homeostasis and survival.

Eukaryotic transcription is a highly regulated process that is required to ensure cell identity. G-quadruplexes (G4s), four-stranded alternative DNA structures formed at G-rich genomic sequences, are thought to be transcriptional regulators enriched in active regulatory regions (*1*). Although mammalian genomes contain more than half a million DNA sequences that can adopt a G4 structure (potential G4-forming sequences, PQSs) (*2*), only a limited number of G4 structures form, and specific G4 subsets have been detected in different cell types (*3*, *4*). Misregulation of G4 positioning disrupts gene expression and embryonic differentiation (*5*, *6*) and has been linked to neurodegenerative disease, cancer, and accelerated aging (*7*, *8*). Furthermore, spontaneous G4s can arise during transcription and DNA replication (*9*, *10*). Such inadvertent G4s, as well as persistent regulatory G4s, cause DNA damage and are associated with genomic deletions and rearrangements (*11*–*13*), further underscoring the importance of G4 structure regulation.

Currently, we lack knowledge on how G4s are controlled to prevent transcription dysregulation and genome instability. G4s are often found in proximity to R-loops, three-stranded structures with an RNA-DNA hybrid and a displaced DNA strand (*14*). These R-loops have been proposed to reinforce the transcriptional regulatory functions of G4s by stabilizing the structure through an RNA-DNA hybrid opposite the G4 strand (*15*). However, this structure, called a G-loop, has not been verified directly beyond bacteria (*9*, *16*). By contrast, R-loops are also induced in the vicinity of ectopically induced G4s through unknown mechanisms and result in the rapid resolution of both structures (*17*). How proximal R-loops control the fate of G4s and whether G-loops are required for this process remain unknown.

G-loop assembly through RNA invasion

To explore the mechanism that couples G4s and R-loops, we used *Xenopus laevis* egg extracts that recapitulate physiological DNA replication, repair, and stress responses (*18*). In cells, the number of PQSs vastly exceeds the number of G4 structures, most likely because G4 formation requires temporal duplex DNA unwinding (*19*). We thus generated a pG4^{BOT} plasmid that contains a canonical G4 structure (G4^{G3N}) opposite a G4-stabilizing noncomplementary sequence, resembling a physiological G4-forming site (fig. S1, A and B). This template was incubated in nucleoplasmic extract (NPE) (*20*), isolated, and subjected to DNA-RNA immunoprecipitation and quantitative polymerase chain reaction (DRIP-qPCR) (fig. S1C). A marked R-loop signal was observed, which peaked at 60 min and declined to the basal level by 120 min, comparable to the half-life time of R-loops in human cells (*21*) (Fig. 1A). Treatment of products with RNaseH, which degrades RNA-DNA hybrids, abolished R-loop signals, validating the signal specificity (fig. S1, D and E). The R-loop signal was lost when the G4 motif was absent (pPolyT) or situated in a duplex region that prevented G4 structure formation (pdsG4^{BOT}) (Fig. 1A and fig. S1F). Moreover, placing a G4 structure on both strands (pG4^{DBL}) prevented R-loop formation, indicating RNA-DNA hybrid formation on the unstructured strand across from the G4 to form a G-loop. R-loop accumulation at the G4 region was also observed for a G4 structure from the telomeric sequence pTEL^{BOT} (fig. S1, A, B, and G to I). This telomeric G4 is less thermodynamically stable compared with G4^{G3N} (*22*), indicating that G-loop assembly is a general response to G4 structure formation regardless of G4 stability. The invading RNA is derived from oocyte transcripts (fig. S1J), because pretreatment of the extract with RNaseA prevented R-loop accumulation, whereas transcription inhibition by RNA polymerase II depletion and α-amanitin treatment had no effect (Fig. 1B and fig. S1, K and L). Chromatin association of RNA can be posttranscriptionally regulated by ribonucleoproteins (*23*, *24*). Consistent with this, depletion of the hnRNPA1 complex, a complex interacting with RNA transcripts involved in R-loop formation (*24*, *25*), prevented hybrid formation, and this was reversed by the addition of the hnRNPA1 complex (fig. S1, M and N). Thus, ribonucleoproteins promote RNA invasion into the non-G4 strand across from a G4 structure to assemble a G-loop (Fig. 2I).

Homology-directed G-loop assembly

Posttranscriptional R-loop formation can be mediated by the RAD51 recombinase (*24*), and, similar to G4 structures, both RAD51 and its loader BRCA2 are enriched in regulatory genomic regions through unknown mechanisms (*26*, *27*). Depletion of RAD51 and BRCA2 prevented R-loop accumulation in pG4^{BOT} (Fig. 1, C and D). Consistent with a role in G-loop formation, BRCA2 and RAD51 were rapidly recruited to the plasmid (Fig. 1, E and F) and specifically located at the G4 region (fig. S2, A to C). This was followed by hnRNPA1 accumulation, allowing G-loop formation (Fig. 1F). RAD51 recruitment required a single-stranded DNA (ssDNA) region, because it was lost when G4 structures were placed on both strands (fig. S2E). Furthermore, recruitment of both RAD51 and hnRNPA1 was prevented upon depletion of BRCA2 (fig. S2D). This suggests that BRCA2-mediated RAD51 loading on the displaced G4-opposing strand promotes homology-directed RNA invasion.

Analogous to its role in recombination activation (*28*), the accumulation of BRCA2 and RAD51 at pG4^{BOT} and subsequent G-loop assembly were promoted by two DNA damage response kinases, ATR and ATM (Fig. 1G and fig. S2, C and F). Rapid ATR activation was detected,

[1]Oncode Institute, Hubrecht Institute-KNAW & University Medical Center Utrecht, Utrecht, Netherlands. [2]Department of Human Genetics, Leiden University Medical Center, Leiden, Netherlands. [3]Science for Life Laboratory, Division of Genome Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. [4]Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, Stockholm, Sweden. [5]Hubrecht Institute-KNAW & University Medical Center Utrecht, Utrecht, Netherlands. [6]Institute of Biology Leiden, Leiden University, Leiden, Netherlands. ***Corresponding author. Email: k.sato@hubrecht.eu (K.S.); p.knipscheer@hubrecht.eu (P.K.)**

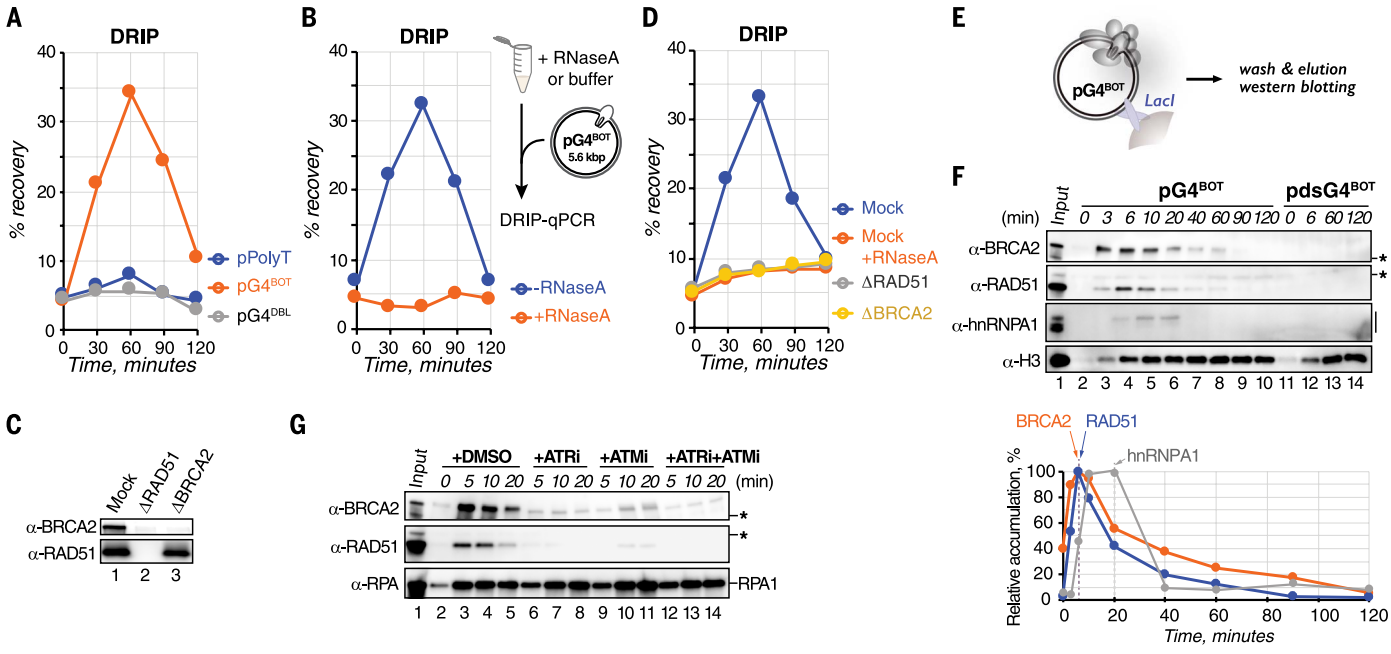

Fig. 1. Mechanism of G-loop establishment. (**A**) pPolyT, pG4^{BOT}, and pG4^{DBL} were incubated with pQuant in NPE, and products were analyzed by DRIP-qPCR with primers for the *G4* locus and the *pQuant* locus (for background signal measurement). After background subtraction, DRIP recovery percentages compared with input signals were plotted. (**B**) pG4^{BOT} was incubated in NPE that was pretreated with RNaseA or buffer, and products were analyzed as in (A). (**C**) Mock-, RAD51-, and BRCA2-depleted NPEs were analyzed by Western blot with RAD51 and BRCA2 antibodies. (**D**) pG4^{BOT} was incubated in the NPEs described in (C), and products were analyzed as in (A). Where indicated, NPE was pretreated with RNaseA. Co-depletion of BRCA2 with RAD51 complicated rescue experiments with recombinant proteins, so the involvement of BRCA2 and RAD51 in the G-loop assembly was not unequivocally established. (**E**) Scheme of the plasmid pull-down assay. In this assay, pG4^{BOT} was incubated in NPE, pulled down with streptavidin beads coated with biotinylated LacI, and bound proteins (oval) were analyzed by Western blot. (**F**) pG4^{BOT} and pdsG4^{BOT} were incubated in NPE and subjected to plasmid pull-down assay. Proteins were analyzed by Western blot with BRCA2, RAD51, hnRNPA1, and histone H3 antibodies (top). The band intensity of each protein was quantified, and the relative intensities compared with the highest signal among the conditions were plotted (bottom). (**G**) pG4^{BOT} was incubated in NPE supplemented with dimethylsulfoxide (DMSO), AZ20 (ATRi), and/or KU-55933 (ATMi); subjected to plasmid pull-down assay; and analyzed with BRCA2, RAD51, and RPA antibodies.

as measured by CHK1 phosphorylation (fig. S2G). ATR activation correlated with RPA binding to the unstructured strand and did not depend on BRCA2 and RAD51 (fig. S2, G to K). By contrast, RPA binding did not require ATR activation (Fig. 1G). Concurrent with RPA binding, both ATR and phosphorylated TOPBP1, a key ATR activator (*29*, *30*), accumulated on pG4^{BOT} (fig. S2, L and M). Therefore, G4-induced ATR activation to promote G-loop assembly resembles canonical checkpoint activation, in which ssDNA-bound RPA activates ATR through interaction with phosphorylated TOPBP1, stimulating recruitment of recombination factors (Fig. 2I).

G4 unwinding initiates G-loop disassembly

To investigate the fate of G-loops, we incubated pG4^{BOT} in egg extract and excised an 81–base pair (bp) fragment containing the G4 (fig. S3A). Native polyacrylamide gel electrophoresis (PAGE) analysis showed the initial noncomplementary bubble structure and an accumulating, faster-migrating, 81-bp linear duplex product (fig. S3, B to D). This structural conversion to duplex DNA must involve DNA synthesis past one of the strands of the G4-containing region (fig. S3A). Consistently, we observed active nucleotide incorporation in the G4 region (fig. S3, E and F). Differential digestion of the top and bottom strands showed preferential synthesis of the G4-complementary strand (Fig. 2, A and B, and fig. S3, C and D). This strand-specific DNA synthesis was blocked upon RAD51 and BRCA2 depletion, which was partly rescued by preloading of RAD51, but not the homologous pairing–deficient mutant (fig. S3, G to I). Thus, homology-mediated G-loop assembly promotes DNA synthesis. Similar G4 conversion was observed for a telomeric G4 (fig. S3J). Other DNA secondary structure motifs also induced R-loop formation and DNA synthesis, suggesting that this may be a general response to DNA secondary structures (fig. S3, K and L). Because DNA synthesis requires a free 3′ end, we hypothesized that this process may involve local nucleolytic incisions. Sequencing gel analysis of linearized reaction products indeed detected incisions specifically on the non-G4 strand one to eight nucleotides from the 3′ ssDNA–double-stranded DNA (dsDNA) junction, which depended on G-loop assembly (Fig. 2, C and D, and fig. S4A). On the basis of the substrate structure, these incisions are likely to be mediated by the XPF-ERCC1 or MUS81-EME1 endonuclease complexes (*31*). Depletion of XPF, but not MUS81, abrogated the incisions, which was rescued by the addition of recombinant XPF-ERCC1 (Fig. 2, E and F). Furthermore, XPF recruitment to the G4 region was dependent on G-loop assembly (fig. S4, B and C). Thus, G-loop assembly triggers the recruitment of XPF-ERCC1 to promote site-specific cleavage of the hybrid strand, leading to resynthesis of the G4-complementary strand.

Because XPF-ERCC1 acts most efficiently on ssDNA-dsDNA junctions with unstructured arms (*32*), we hypothesized that this incision requires prior G4 unfolding. To investigate this, we depleted two DNA helicases, DHX36 and FANCJ, which are also required for DNA replication–coupled G4 unwinding in *X. laevis* egg extracts (*33*). Loss of both helicases abolished G4 resolution and incision without affecting G-loop assembly and XPF-ERCC1 recruitment (Fig. 2, G and H, and fig. S4, B to H). The depletion also prevented G4 conversion, and these defects were rescued by recombinant wild-type (WT) DHX36 and FANCJ but not their ATPase-dead mutants (fig. S4, I and J). Moreover, recombinant XPF-ERCC1 catalytic activity is inhibited by a G4 structure on the 5′ ssDNA flap of a splayed arm model substrate (fig. S4K). G4 unwinding

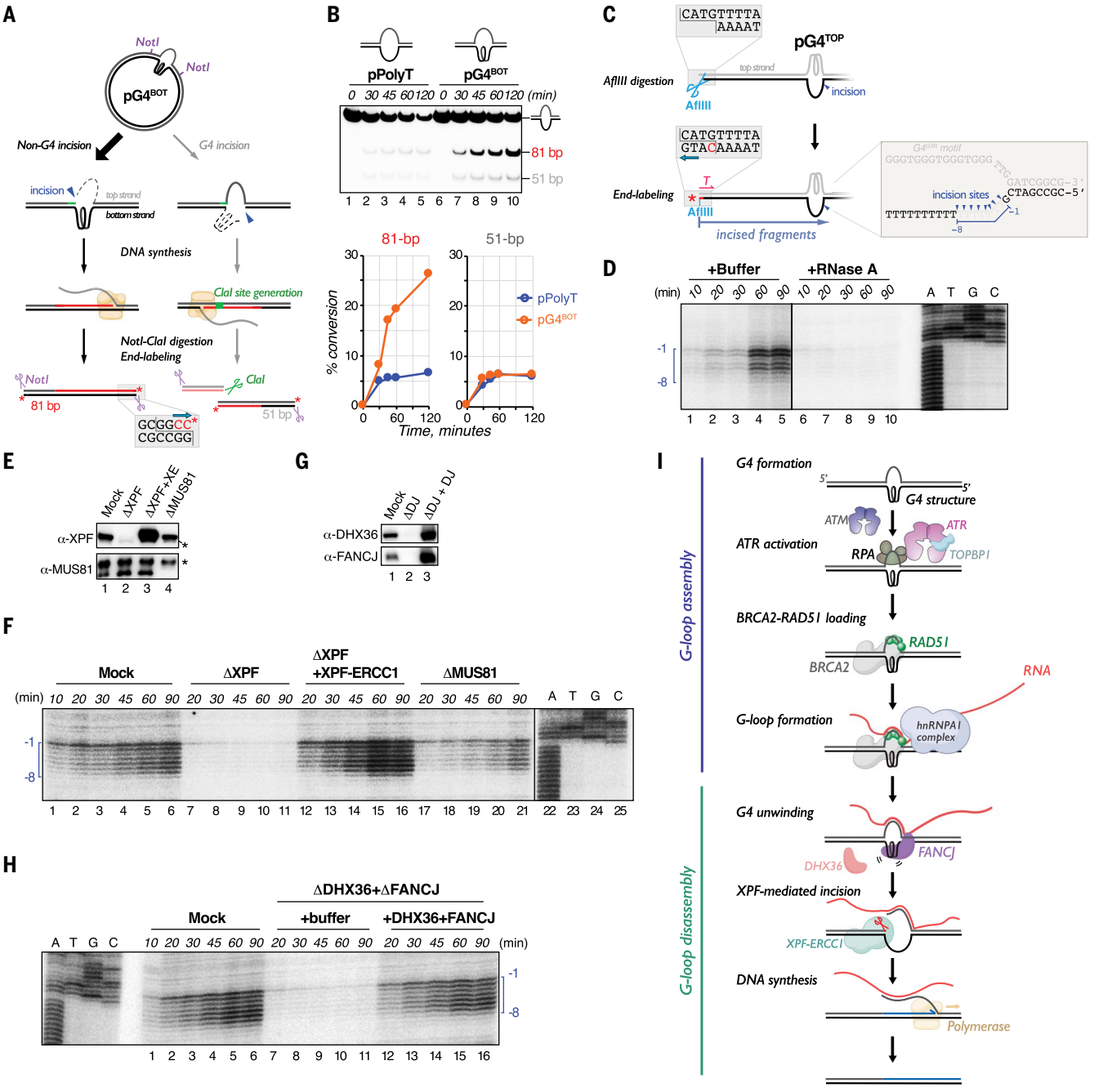

Fig. 2. Mechanism of G-loop disassembly. (**A**) Schematic of the G4 conversion assay. The G4-forming region of pG4^{BOT} is converted to complementary dsDNA in NPE. This conversion must involve DNA synthesis past at least one of the strands of the G4-containing region, which requires a free 3′ end that is likely generated by incision (blue arrowhead). The products were digested with NotI and ClaI and end-labeled (asterisk) by fill-in of the 5′-overhangs with Sequenase DNA polymerase (blue arrow) in the presence of ^{32}P-α-dCTP (see also fig. S3A). ClaI specifically incises the product when the G4 strand is resynthesized (see also fig. S1A). (**B**) pPolyT and pG4^{BOT} were incubated in NPE. At various times, DNA was isolated, digested with NotI and ClaI, end-labeled by fill-in of the 5′-overhangs with ^{32}P-α-dCTP, separated by native PAGE, and visualized by autoradiography (top). The 81- and 51-bp fragments were quantified, and the conversion percentage was calculated by comparing their intensity against the total intensity of all fragments at time point 0. Conversion was plotted against time (bottom). (**C**) Schematic representation of incision assay. After incubation of pG4^{TOP} in NPE, products were digested with AflIII and end-labeled by fill-in of the 5′-overhang with Sequenase DNA polymerase in the presence of ^{32}P-α-dCTP (blue arrow). The region surrounding the incision sites is enlarged and incision positions observed in (D) are indicated (blue arrowheads) (right). Incision products are numbered −8 to −1, where the −1 product corresponds to a product in which incision takes place between the first and the second nucleotide upstream of the ssDNA-dsDNA junction. (**D**) pG4^{TOP} was incubated in NPE that was pretreated with RNaseA or buffer. At various times, DNA was isolated, digested with AflIII, end-labeled by fill-in of the 5′-overhang with ^{32}P-α-dCTP, separated by denaturing PAGE, and visualized by autoradiography. A sequencing ladder derived from extension of primer *T* annealed to pG4^{TOP} (C) was used as a size marker. Incision products are indicated with a blue bracket. (**E**) Mock-, XPF-, and MUS81-depleted NPEs supplemented with buffer or XPF-ERCC1 were analyzed by Western blot with XPF and MUS81 antibodies (bottom). The asterisks represent nonspecific bands. (**F**) pG4^{TOP} was incubated in the NPEs described in (E), and products were analyzed as in (D). (**G**) Mock- and DHX36-FANCJ–depleted NPEs supplemented with buffer or WT DHX36 and FANCJ were analyzed by Western blot with DHX36 and FANCJ antibodies. (**H**) pG4^{TOP} was incubated in the NPEs described in (G), and products were analyzed as in (D). (**I**) Model for G4 suppression through the G-loop assembly-disassembly cycle.

is thus required for site-specific R-loop incision and G-loop disassembly. G4 unwinding coincided with G-loop assembly and required both BRCA2 and invading RNA but not downstream factors XPF-ERCC1 (fig. S4, E and L to O). Therefore, G4 resolution is a highly coordinated process that is initiated by G-loop assembly, leading to G4 structure unwinding followed by nucleolytic incision and DNA synthesis past the G4 sequence (Fig. 2I).

RNA sequence and abundance regulate G-loops

The sequence opposite the G4 in the pG4 substrates is a poly(T) stretch (fig. S1A). Therefore, the invading RNA could be poly(A) stretches in extract-derived RNA molecules. Consistent with this, replacement of the poly(T) with a poly(A) stretch in pG4^{BOT} blocked G-loop assembly (figs. S1A and S5A). To determine whether other sequences also support G-loop formation, we placed a 21-nucleotide unique DNA sequence opposite the G4 and incubated the plasmid (pG4^{BOT}-I) in NPE supplemented with a transcript containing a complementary sequence (transcript I) or a noncomplementary sequence (transcript II) (Fig. 3A). Upon incubation with the noncomplementary RNA, only minor RNA-DNA hybrid formation was detected (Fig. 3C), and the bubble structure underwent rapid degradation (Fig. 3B), as also seen when G-loop assembly was prevented (Fig. 2B and fig. S3, I and J). This degradation did not involve XPF-ERCC1 but was promoted by the MRE11 and DNA2 nucleases (fig. S5, B and C). Degradation was blocked by incubation with the complementary transcript I at a nearly 1:1 G4 plasmid-to-transcript ratio, which also supports G-loop formation and subsequent disassembly (Fig. 3, B and C). Similar results were obtained with a different G4-opposing sequence, pG4^{BOT}-II (fig. S5, D to F), supporting a universal mechanism by which homologous RNA transcripts stimulate G-loop assembly and protect them from degradation. Although increasing the level of complementary transcript to a 1:5 G4 plasmid-to-transcript ratio further promoted G-loop assembly, G-loop disassembly was inhibited (Fig. 3, B and C). Titration showed that stoichiometric concentrations ranging from 0.4 to 4 transcript copies per G4 were required for efficient G-loop disassembly, whereas higher concentrations prevented disassembly (Fig. 3D). Unlike complementary RNA, complementary ssDNA did not support either G4 suppression or stabilization (fig. S5G). We conclude that G-loop assembly and disassembly are controlled by RNA availability and levels, thereby conferring plasticity to G4 regulation (fig. S5H).

G-loop disassembly shapes G4–R-loop landscapes

To explore the role of G-loops in the regulation of cellular G4 landscapes, we deleted *Dhx36*, *Fancj*, or both genes in mouse embryonic stem cells (mESCs) and mapped genomic G4 structures by quantitative CUT&Tag (*14*) (Fig. 4A and fig. S6). More than 97% of G4 CUT&Tag peaks colocalized with PQSs, validating the specificity of this method (fig. S7A). Individual loss of either DHX36 or FANCJ resulted in an increase of G4 signal at 12,109 and 693 sites, respectively (fig. S7B). Moreover, their combined loss resulted in the up-regulation at 21,045 G4s genome wide, accounting for 29% of the total 71,445 G4s detected. Up-regulated G4s overlapped among the mutant lines (fig. S7C), and their G4 signals were further elevated in the double mutant compared with the single mutants (Fig. 4, B and C, and

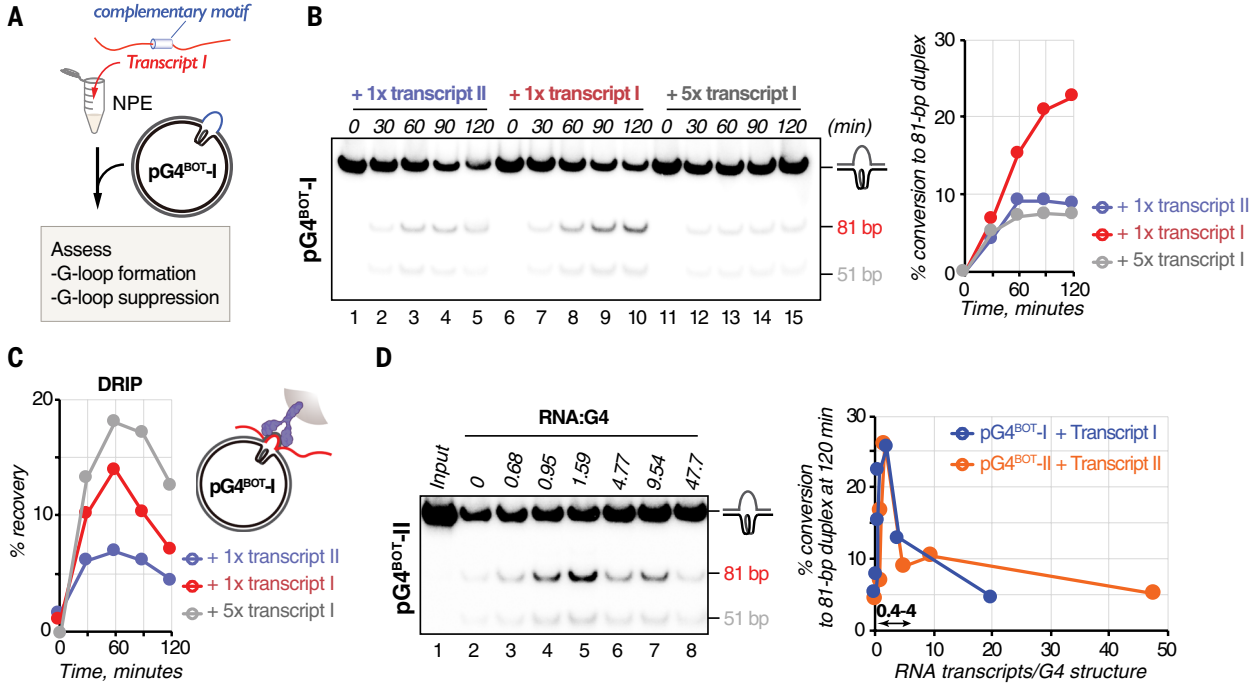

Fig. 3. RNA sequence and abundance regulate G-loops. (**A**) Experimental setup for transcript addition. A synthetic RNA transcript (transcript I) with a sequence complementary to the DNA opposite the G4 in pG4^{BOT}-I was used. DRIP and G4 conversion assays were performed in the presence of this substrate and RNA. (**B**) pG4^{BOT}-I was incubated in NPE supplemented with transcript I or a noncomplementary transcript (transcript II) at 10 nM (1×) or 50 nM (5×), and products were analyzed by G4 conversion assay (top). The 81-bp fragment was quantified, and the conversion percentage was calculated by comparing their intensity against the total intensity of all fragments at time point 0. Conversion was plotted against time (bottom). (**C**) pG4^{BOT}-I was incubated with pQuant in NPE supplemented with transcript I or transcript II at the indicated ratio, and products were analyzed by DRIP-qPCR with primers for the *G4* locus and the *pQuant* locus (for background signal measurement). After background subtraction, DRIP recovery percentages compared with input signals were plotted. (**D**) pG4^{BOT}-I and pG4^{BOT}-II were incubated in NPE supplemented with transcript I and transcript II, respectively, at the indicated ratio, and products were analyzed as in (B). The 81-bp fragment at time point 120 min was quantified, and the conversion percentage was calculated by comparing their intensity against the total intensity of all fragments at time point 0 (input). Conversion was plotted against the ratio of transcript to plasmid. G4 conversion was specifically stimulated at a 0.4:4 transcript-to-plasmid template ratio (double black arrow).

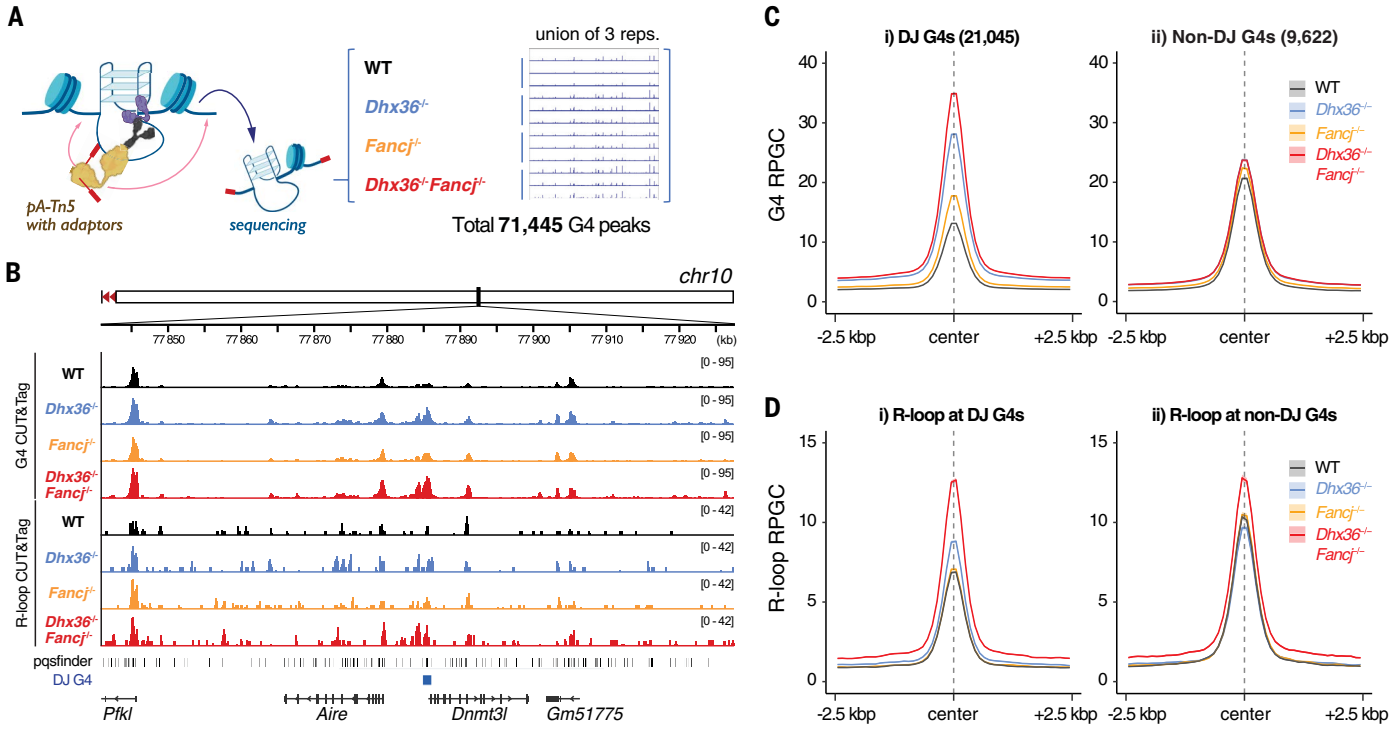

Fig. 4. G-loop disassembly shapes G-loop landscapes. (**A**) Schematic of quantitative G4 CUT&Tag. Permeabilized cells were treated with the FLAG-tagged BG4 antibody (blue) to capture G4s in chromatin. An anti-FLAG antibody (black) was subsequently added to tether the pA-Tn5 transposase bound to adaptors (red) for chromatin tagmentation. Genomic DNA was extracted after tagmentation and mixed with tagmented spike-in DNA. Tagmented fragments were amplified and sequenced. The number of total high-confidence G4 CUT&Tag peaks identified across technical triplicates in WT, *Dhx36*$^{-/-}$, *Fancj*$^{-/-}$, and *Dhx36*$^{-/-}$*Fancj*$^{-/-}$ mESCs is indicated. (**B**) Genome browser view at a representative locus showing RPGC-normalized (1× Genome Coverage) G4 and R-loop CUT&Tag peaks from triplicate (*y*-axis scale, right top) PQSs determined by pqsfinder and DJ G4s in the indicated mESCs. G4s significantly up-regulated in *Dhx36*$^{-/-}$*Fancj*$^{-/-}$ cells [$\log_2$ fold change ($\log_2$FC) >1, false discovery rate (FDR) ≤1%] were defined as DJ G4s and were significantly unchanged ($\log_2$FC <1, FDR ≤1%) as non-DJ G4s. (**C**) G4 distribution around 21,045 DJ G4 (i) and 9622 non-DJ (ii) G4 sites in the indicated mESCs. Average profiles of RPGC-normalized G4 CUT&Tag peaks within ±2.5 kilobase pairs (kbp) from their center are plotted. In WT cells, 9622 non-DJ G4 peaks shared a significantly higher signal ($P < 2.2 \times 10^{-16}$ calculated by the Wilcoxon test; Cohen's $d = 0.59$, medium) compared with the 21,045 DJ G4 peaks. Conversely, in *Dhx36*$^{-/-}$*Fancj*$^{-/-}$ cells, the 21,045 DJ G4 peaks shared a significantly higher signal ($P < 2.2 \times 10^{-16}$ calculated by the Wilcoxon test; Cohen's $d = 0.58$, medium) compared with the 9622 non-DJ G4 peaks. (**D**) R-loop CUT&Tag was performed in the indicated mESCs, and R-loop distributions around 21,045 DJ G4s (i) and 9622 non-DJ G4s (ii) were plotted with average profiles of RPGC-normalized R-loop CUT&Tag peaks within ±2.5 kbp from each G4 center.

fig. S7D), indicating partially redundant G4 suppression by DHX36 and FANCJ in mESCs. Very few G4s formed at de novo sites upon depletion of DHX36 and FANCJ; rather, the signal at existing G4 sites increased (Fig. 4B and fig. S7B). Therefore, the DHX36 and FANCJ helicases, rather than preventing de novo G4 formation, play an important role in limiting the abundance of intrinsic G4 structures throughout the genome.

Both the up-regulated 21,045 G4s (DJ G4s) and the unchanged 9622 G4s (non-DJ G4s) in double mutant cells compared with WT cells predominantly resided in promoters and enhancers and frequently overlapped with actively transcribed RNA (*3*, *34*) (fig. S7, E to G). DJ G4s were specifically found at G4 sites with low G4 signal in WT cells but were excluded from sites with high WT G4 signal, which often overlapped with non-DJ G4 sites (Fig. 4, B and C). This indicates that stably formed G4s, for example, those at active promotors, may not be substrates for DHX36-FANCJ–mediated disassembly. Consistent with the characteristic DJ G4 landscape, quantitative R-loop CUT&Tag detected a lower RNA-DNA hybrid signal at DJ G4s than at non-DJ G4s in WT cells (Fig. 4D). Loss of DHX36 and FANCJ resulted in robust accumulation of RNA-DNA hybrids at DJ G4s (Fig. 4D). Moreover, transcriptome sequencing revealed that loss of both helicases led to robust up-regulation of transcripts throughout the genome (fig. S8, A to D, and data S1 and S2). These data suggest that a defect in the dynamic G-loop assembly-disassembly cycle leads to G-loop accumulation and transcriptional dysregulation.

To further validate this mechanism, we examined another player in G-loop disassembly, XPF-ERCC1, using G4 and R-loop CUT&Tag in ERCC1-deficient mESCs (*35*) (fig. S9A). Because our biochemical data suggested that XPF-ERCC1 acts downstream of G4 unwinding (Fig. 2I), we hypothesized that its depletion would not affect G4 abundance but rather would increase R-loops due to the lack of G-loop disassembly. As expected, the loss of *Ercc1* resulted in the accumulation of R-loops at DJ G4 sites with at most a mild G4 up-regulation (fig. S9, B to D). We also addressed the recruitment of recombination factors to G4 sites by comparing BRCA2 chromatin immunoprecipitation sequencing data from WT mESCs (*36*) with our G4 sites, which revealed that 82% of BRCA2 peaks resided at the G4 sites (figs. S7A and S10). Therefore, G4-forming regions are a predominant BRCA2 substrate in unperturbed cells, supporting a genome-wide role for BRCA2 in G-loop assembly. Finally, consistent with G-loop disassembly requiring low concentrations of invading RNA in *X. laevis* egg extract, we found that the transcripts targeting to DJ G4 sites, particularly the long noncoding RNAs, were expressed at considerably lower levels compared with those to non-DJ G4s (fig. S8E). We thus conclude that the G-loop assembly-disassembly cycle through coordinated actions by the HR proteins, DHX36, FANCJ, and XPF-ERCC1 is a global mechanism used to repress G4s.

Two-tier G4 suppression ensures genome stability

Persistent G4 structures have been shown to cause DNA damage (*12*, *17*). Consistently, loss of DHX36 and FANCJ increased 53BP1-positive foci, which is indicative of DNA double-strand breaks (DSBs) (fig. S11A). Although DSBs detected in both 5-ethynyl-2′-deoxyuridine (EdU)–positive and EdU-negative cells, their accumulation was more pronounced in EdU-positive cells (Fig. 5A), indicating major damage induction in the S phase. DSBs are at least partly generated through transcription in the G_1 phase, because transcription inhibition reduced 53BP1 foci in EdU-negative cells (fig. S11B).

To further investigate the effect of *Dhx36* and *Fancj* deletion on DNA replication, we used single-cell EdU sequencing (scEdU-seq) (*37*). Two pulses of EdU labeling, in combination with single-cell sequencing, enabled us to detect fork-stalling events and genome-wide replication timing throughout the S phase (Fig. 5B). We found DJ G4s to be enriched in early replicating regions (Fig. 5C, top). FANCJ loss caused an increase in fork stalling, specifically at the early S phase. Although single DHX36 loss had little effect on DNA replication, loss of both helicases further exacerbated fork stalling at the early S phase (Fig. 5C), corresponding to reduced EdU incorporation (fig. S11, C and D). These results were not due to a change in replication timing (fig. S11, E and F) or to G4s stochastically formed at the replication fork, because replication stalling was confined to the early S phase; therefore, they instead show that G-loops transmitted from the G_1 phase led to replication fork stalling.

In addition to their role in G-loop–mediated G4 resolution, DHX36 and FANCJ also promote G4 resolution when replication forks collide with a G4 structure in *X. laevis* egg extract (*33*). Loss of this mechanism caused persistent replication stalling, which could result in enhanced γH2AX foci in mitotic chromosomes in the double-knockout cells (Fig. 5, D and E). Furthermore, these cells showed replication stress–induced genome instability, as represented by enhanced micronuclei and 53BP1 bodies and intrinsic growth defects (fig. S12). Stabilization of G4 structures by pyridostatin (*12*) phenocopied the DNA damage induction and genome instability observed in the double mutant, confirming that these adverse phenotypes were caused by persistent G4s throughout the genome (Fig. 5 and figs. S11 and S12). Loss of *Dhx36* alone, which accumulates G4 structures in mESCs (Fig. 4) but has only a mild effect on replication-coupled G4 resolution in egg extract (*33*), did not result in the accumulation of DSBs (Fig. 5, A and E, and fig. S12, A to C). Thus, the persistence of G4s upon S-phase entry is likely a major cause of replication-associated DSBs only if

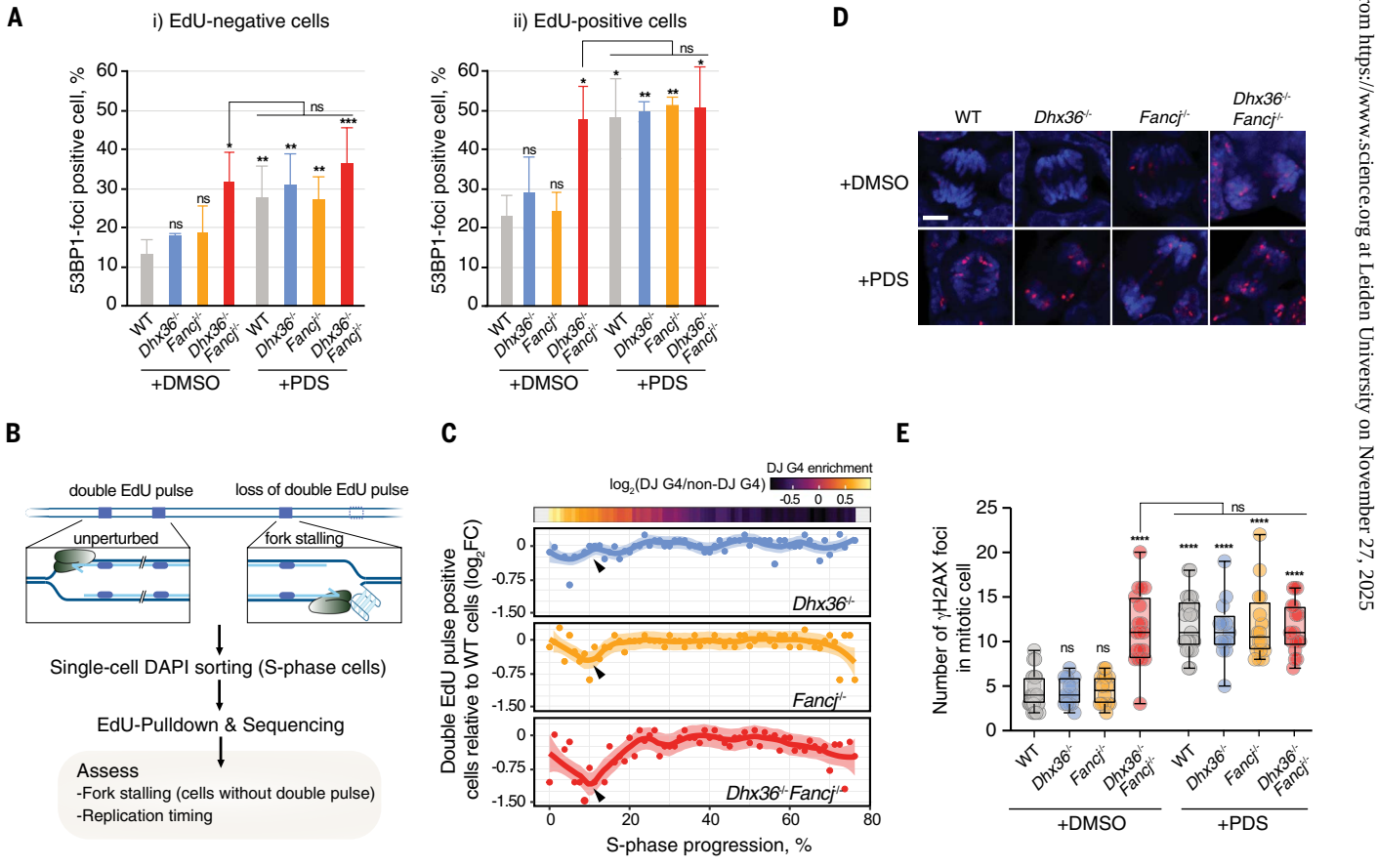

Fig. 5. Two-tier mechanism of G4 resolution. (**A**) 53BP1 foci-positive cells were counted and displayed as the mean percentage of total EdU-negative cells (i) or EdU-positive cells (ii) with SDs ($n = 3$, with >40 cells for each experiment). WT cells treated with DMSO, or where indicated, *Dhx36*$^{-/-}$*Fancj*$^{-/-}$ cells treated with DMSO, were compared for a *t* test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; ns, not significant. (**B**) Schematic of scEdU-seq. Two pulses of EdU labeling in combination with single-cell sequencing detected fork-stalling events. 4′,6-Diamidino-2-phenylindole (DAPI) intensity was used to determine genome-wide replication timing throughout the S phase. (**C**) Cells labeled with a double-EdU pulse were analyzed by scEdU-seq. The relative population of cells with the double EdU pulses compared with WT cells was plotted over S-phase progression in the indicated mESCs. Reduction of double EdU pulses indicates fork stalling. Heatmap of DJ G4 enrichment relative to non-DJ G4s is shown (top). (**D**) mESCs were treated with DMSO or pyridostatin (PDS) and visualized by immunostaining with a γH2AX antibody, along with DAPI counterstaining. Representative mitotic chromosomes are shown. Scale bar, 5 μm. (**E**) Box plots of the γH2AX focus number in mitotic cells described in (D) ($n = 2$, with >20 mitotic cells for each experiment). Center line indicates the median; box limits, 75 and 25 percentiles; and whiskers, min and max values. ****$P < 0.0001$.

replication-coupled G4 resolution is impaired. G4 resolution through G-loop assembly-disassembly throughout the cell cycle represents the first tier of protection against G4-induced DNA damage, but if this fails and G4s persist into S phase, then replication-coupled G4 unwinding provides an essential backup to ensure genome integrity.

Discussion

Our data establish a stepwise process that controls G4 dynamics through the assembly of a G-loop by the invasion of RNA transcripts across from the G4, followed by G4 unwinding coupled with G-loop disassembly (Fig. 2I). Interfering with this G-loop cycle results in severe dysregulation of the transcriptome, extensive genome instability, and proliferation defects, highlighting its vital role in cellular homeostasis and survival. This mechanism involves several unanticipated aspects. First, the G4 structure is recognized as a "DNA lesion," thereby enabling activation of the downstream DNA repair proteins in the absence of DNA damage. Second, it uses RNA molecules in trans that are not physically linked to ongoing transcription for RNA-DNA hybrid formation. Although such RNA trans action has been thought to occur at confined loci in vertebrates (*38*), our findings define it as a prevalent genome-wide mechanism that plays a twofold role in the G-loop assembly-disassembly cycle and the protection from nucleolytic attacks when G4 structures form. Third, this mechanism results in renewal of the non-G4 strand, which allows for erasing potentially mutagenic modifications on the displaced strand (*39*–*42*), thereby safeguarding the integrity of regulatory DNA motifs. Fourth, the mechanism is controlled by the concentration of invading RNAs, allowing for both G4 stabilization and suppression, and this defines RNA levels as an important regulatory parameter of G4 landscapes (fig. S5H). Finally, we observed that other DNA secondary structure motifs also promote RNA-DNA hybrid formation, but whether these are processed through a similar mechanism is a subject for future research.

Many of the proteins involved in this G-loop cycle, including RAD51, BRCA2, FANCJ, and XPF-ERCC1, also promote canonical DNA repair pathways and are directly linked to human diseases characterized by developmental abnormalities and predisposition to cancer. Although these pathways safeguard genome stability against exogenous and endogenous DNA damage, the G-loop mechanism specifically protects the genome against the accumulation of endogenous DNA structures. Given the abundance of G4 structures and DNA damage accumulation observed here, defects in G4 resolution can cause DNA damage to an extent that exceeds most other forms of endogenous DNA damage, thus posing a severe threat to cellular homeostasis. This highlights an essential, DNA lesion-independent role for DNA repair proteins in genome maintenance.

REFERENCES AND NOTES

1. D. Varshney, J. Spiegel, K. Zyner, D. Tannahill, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **21**, 459–474 (2020).
2. V. S. Chambers *et al.*, *Nat. Biotechnol.* **33**, 877–881 (2015).
3. R. Hänsel-Hertsch *et al.*, *Nat. Genet.* **48**, 1267–1272 (2016).
4. W. W. I. Hui, A. Simeone, K. G. Zyner, D. Tannahill, S. Balasubramanian, *Sci. Rep.* **11**, 23641 (2021).
5. R. Hänsel-Hertsch *et al.*, *Nat. Genet.* **52**, 878–883 (2020).
6. K. G. Zyner *et al.*, *Nat. Commun.* **13**, 142 (2022).
7. N. Maizels, *EMBO Rep.* **16**, 910–922 (2015).
8. D. Chakravarti, K. A. LaBella, R. A. DePinho, *Cell* **184**, 306–322 (2021).
9. M. L. Duquette, P. Handa, J. A. Vincent, A. F. Taylor, N. Maizels, *Genes Dev.* **18**, 1618–1629 (2004).
10. W. T. C. Lee *et al.*, *Nat. Commun.* **12**, 2525 (2021).
11. I. Cheung, M. Schertzer, A. Rose, P. M. Lansdorp, *Nat. Genet.* **31**, 405–409 (2002).
12. R. Rodriguez *et al.*, *Nat. Chem. Biol.* **8**, 301–310 (2012).
13. K. Paeschke *et al.*, *Nature* **497**, 458–462 (2013).
14. J. Lyu, R. Shao, P. Y. Kwong Yung, S. J. Elsässer, *Nucleic Acids Res.* **50**, e13 (2022).
15. P. Wulfridge *et al.*, *Mol. Cell* **83**, 3064–3079.e5 (2023).
16. M. L. Duquette, M. D. Huber, N. Maizels, *Cancer Res.* **67**, 2586–2594 (2007).
17. A. De Magis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 816–825 (2019).
18. W. S. Hoogenboom, D. Klein Douwel, P. Knipscheer, *Dev. Biol.* **428**, 300–309 (2017).
19. G. Miglietta, M. Russo, G. Capranico, *Nucleic Acids Res.* **48**, 11942–11957 (2020).
20. J. Walter, L. Sun, J. Newport, *Mol. Cell* **1**, 519–529 (1998).
21. M. P. Crossley *et al.*, *Nature* **613**, 187–194 (2023).
22. P. L. Tran, J. L. Mergny, P. Alberti, *Nucleic Acids Res.* **39**, 3282–3294 (2011).
23. S. Carpenter *et al.*, *Science* **341**, 789–792 (2013).
24. M. Feretzaki *et al.*, *Nature* **587**, 303–308 (2020).
25. S. Redon, I. Zemp, J. Lingner, *Nucleic Acids Res.* **41**, 9117–9128 (2013).
26. I. Hazan, J. Monin, B. A. M. Bouwman, N. Crosetto, R. I. Aqeilan, *Cell Rep.* **29**, 560–572.e4 (2019).
27. J. J. Gruber *et al.*, *Cell Rep.* **28**, 2182–2193.e6 (2019).
28. R. Buisson *et al.*, *Mol. Cell* **65**, 336–346 (2017).
29. C. Frattini *et al.*, *Mol. Cell* **81**, 1231–1245.e8 (2021).
30. A. Kumagai, J. Lee, H. Y. Yoo, W. G. Dunphy, *Cell* **124**, 943–955 (2006).
31. P. M. Dehé, P. H. L. Gaillard, *Nat. Rev. Mol. Cell Biol.* **18**, 315–330 (2017).
32. M. Jones *et al.*, *Nat. Commun.* **11**, 1120 (2020).
33. K. Sato, N. Martin-Pintado, H. Post, M. Altelaar, P. Knipscheer, *Sci. Adv.* **7**, eabf8653 (2021).
34. M. Tastemel *et al.*, *Stem Cell Res.* **25**, 250–255 (2017).
35. N. Bona, G. P. Crossan, *Nat. Struct. Mol. Biol.* **30**, 1434–1445 (2023).
36. Y. Wang *et al.*, *Nat. Commun.* **13**, 3493 (2022).
37. J. van den Berg *et al.*, *Nat. Methods* **21**, 1175–1184 (2024).
38. L. Statello, C. J. Guo, L. L. Chen, M. Huarte, *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
39. L. B. Alexandrov *et al.*, *Nature* **500**, 415–421 (2013).
40. P. Revy *et al.*, *Cell* **102**, 565–575 (2000).
41. M. Muramatsu *et al.*, *Cell* **102**, 553–563 (2000).
42. S. Ito *et al.*, *Science* **333**, 1300–1303 (2011).
43. S. J. Elsasser, J. Lyu, C. Navarro Luzón, K. Sato, P. Knipscheer, "G4 and Rloop qCUT&Tag - RNA transcripts regulate G-quadruplex landscapes through G-loop formation" (Zenodo, 2025); https://doi.org/10.5281/zenodo.15127429.
44. J. van den Berg, "Gloop-scEdUseq: Zenodo Gloop release" (Zenodo, 2025); https://doi.org/10.5281/zenodo.15121012.

ACKNOWLEDGMENTS

We thank M. van Vugt and M. Takata for feedback on the manuscript; the Hubrecht animal caretakers for animal support; the Hubrecht flow cytometry core facility and the Utrecht Sequencing Facility for single-cell sorting and sequencing, which are subsidized by the University Medical Center Utrecht, the Hubrecht Institute, Utrecht University, and the Netherlands X-omics Initiative (NWO project 184.034.019); T. Nyman, E. Strandback, and the members of the Knipscheer laboratory for assistance and helpful discussions; S. Noordermeer for initial help with the DRIP-qPCR protocol; G. P. Crossan, J. Garaycoechea, and J. Smink for providing the *Ercc1*$^{-/-}$ mESCs; and Protein Production Sweden (PPS) for providing BG4 antibody. **Funding:** This work was supported by the Kanae Foundation for the Promotion of Medical Science (K.S.); the Japanese Biochemical Society (K.S.); the European Research Council (ERC Consolidator Grant ERC-CoG 101003210 to P.K., ERC Advanced grant ERC-AdG 101053581-scTranslatomics to A.v.O., and ERC Starting Grant ERC-StG 851564 to F.M.); The Netherlands Organisation for Scientific Research (NWO) through the Gravitation program CancerGenomiCs.nl (P.K.); The Oncode Institute, which is partly financed by the Dutch Cancer Society (KWF) (A.v.O. and P.K.); Vetenskapsrådet Sweden (grant 2020–04313, 2023-02380 to S.J.E.); KI StratRegen Junior Project Grant (S.J.E.); The Ming Wai Lau Center for Reparative Medicine, Sweden (S.J.E.); Stiftelsen för Strategisk Forskning (FFL7 to S.J.E.); The Knut och Alice Wallenbergs Stiftelse, Sweden (S.J.E.); and the Karolinska Institutet KI Research Foundation (grant 2022-01826 to C.N.). P.P.S. is funded by the Swedish Research Council as national research infrastructure. The National Academic Infrastructure for Supercomputing in Sweden (NAISS) is funded by the Swedish Research Council through grant agreement no. 2022-06725. The Supercomputing Cluster UPPMAX was used under projects NAISS 2023/6-19, NAISS 2023/22-84, SNIC 2022/6-14, and NAISS 2024/22-108. **Author contributions:** Conceptualization: K.S., P.K.; Data curation: K.S., J.L., J.v.d.B., A.v.O., M.T., S.J.E., P.K.; Formal analysis: K.S., J.L., J.v.d.B., V.M.C., C.N.L., J.D., S.J.E.; Funding acquisition: K.S., C.N.L., F.M., A.v.O., S.J.E., P.K.; Investigation: K.S., J.L., J.v.d.B., D.B., J.S., C.E.J., M.A., J.D., A.H.; Methodology: K.S., J.L., J.v.d.B., V.M.C., C.N.L., S.J.E., J.S., C.E.J., A.H.; Resources: J.S., C.E.J., M.T.; Software: J.L., J.v.d.B., V.M.C., C.N.L., S.J.E.; Validation: K.S., J.L., J.v.d.B., D.B., J.S., C.E.J., M.A., J.D., A.H.; Visualization: K.S., J.L., J.v.d.B., V.M.C., C.N.L., S.J.E., J.S., C.E.J.; Project administration: K.S., P.K.; Supervision: F.M., A.v.O., M.T., S.J.E., P.K.; Writing – original draft: K.S., P.K., J.L., J.v.d.B., J.S., F.M., M.T., S.J.E.; Writing – review & editing: K.S., P.K., J.L., S.J.E. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. All material generated in this study (including but not limited to oligonucleotides, plasmids, purified proteins, antibodies, and cell lines) are available on request. CUT&Tag data, RNA-seq data, and scEdU-seq data generated for this study have been deposited at the Gene Expression Omnibus under GSE268153 (scEdU-seq) and GSE269084 (CUT&Tag and RNA-seq). Custom code used in this study has been deposited on Zenodo (*43*,*44*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adr0493
Materials and Methods; Figs. S1 to S17; Tables S1 to S5; References (*45*–*117*); Data S1 to S2; MDAR Reproducibility Checklist

Submitted 12 June 2024; resubmitted 28 February 2025; accepted 19 April 2025

10.1126/science.adr0493