

Large-scale expansions of Friedreich's ataxia GAA•TTC repeats in an experimental human system: role of DNA replication and prevention by LNA-DNA oligonucleotides and PNA oligomers

Anastasia Rastokina^{1,†}, Jorge Cebrián^{1,†}, Negin Mozafari^{2,†}, Nicholas H. Mandel¹, C.I. Edvard Smith², Massimo Lopes³, Rula Zain^{2,4,*} and Sergei M. Mirkin^{1,*}

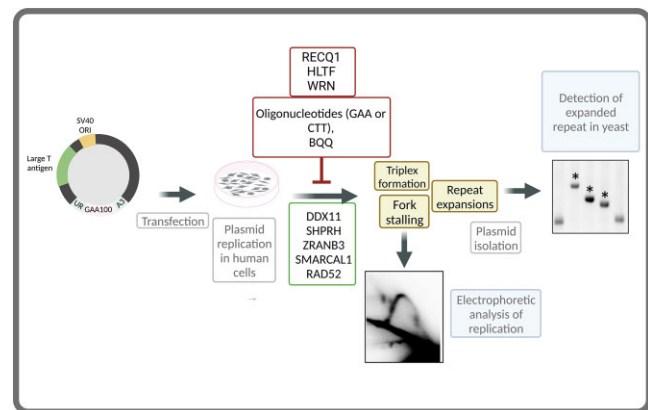
¹Department of Biology, Tufts University, Medford, MA 02155, USA, ²Department of Laboratory Medicine, Translational Research Center Karolinska (TRACK), Karolinska Institutet, Karolinska University Hospital, SE-171 77 Stockholm, Sweden, ³Institute of Molecular Cancer Research, University of Zurich, Zurich 8057, Switzerland and ⁴Center for Rare Diseases, Karolinska University Hospital, SE-17176 Stockholm, Sweden

Received October 28, 2022; Revised May 02, 2023; Editorial Decision May 03, 2023; Accepted May 20, 2023

ABSTRACT

Friedreich's ataxia (FRDA) is caused by expansions of GAA•TTC repeats in the first intron of the human *FXN* gene that occur during both intergenerational transmissions and in somatic cells. Here we describe an experimental system to analyze large-scale repeat expansions in cultured human cells. It employs a shuttle plasmid that can replicate from the SV40 origin in human cells or be stably maintained in *S. cerevisiae* utilizing ARS4-CEN6. It also contains a selectable cassette allowing us to detect repeat expansions that accumulated in human cells upon plasmid transformation into yeast. We indeed observed massive expansions of GAA•TTC repeats, making it the first genetically tractable experimental system to study large-scale repeat expansions in human cells. Further, GAA•TTC repeats stall replication fork progression, while the frequency of repeat expansions appears to depend on proteins implicated in replication fork stalling, reversal, and restart. Locked nucleic acid (LNA)-DNA mixmer oligonucleotides and peptide nucleic acid (PNA) oligomers, which interfere with triplex formation at GAA•TTC repeats *in vitro*, prevented the expansion of these repeats in human cells. We hypothesize, therefore, that triplex formation by GAA•TTC repeats stall replication fork progression, ultimately leading to repeat expansions during replication fork restart.

GRAPHICAL ABSTRACT



INTRODUCTION

Expansions of simple, tandem DNA repeats cause over 50 hereditary diseases in humans (1–3). One of these diseases is Friedreich's ataxia (FRDA): a rare, autosomal recessive degenerative disease caused by the expansion of GAA•TTC repeats in the first intron of the frataxin gene (*FXN*) (4). Normal alleles contain from 7 to 22 repeats, while FRDA patients have 66 to 1700 repeats in both alleles of the *FXN* gene (5). Lengthening of GAA•TTC tracts results in a progressive reduction of the *FXN* mRNA (6–8), frataxin deficiency, mitochondrial dysfunction and cell death (9,10). Consequently, the lengths of GAA•TTC repeats correlate directly with the severity of the disease and inversely with its age-at-onset (11).

*To whom correspondence should be addressed. Tel: +1 627 627 4794; Fax: +1 617 627 3805; Email: sergei.mirkin@tufts.edu
Correspondence may also be addressed to Rula Zain. Tel: +46 08 524 800 00; Email: rula.zain@ki.se

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

‡Rula Zain and Sergei M. Mirkin are co-senior Authors.

The nuclear *FXN* gene encodes the protein frataxin, which predominantly localizes in mitochondria, but is also weakly expressed in nuclei, endoplasmic reticulum and microsomes (12,13). It is responsible for iron-sulfur (Fe-S) cluster biosynthesis, thus, diminished levels of frataxin lead to toxic iron accumulation in mitochondria, elevation of cellular oxidative stress and subsequent cell death (14–16). Neurodegeneration in FRDA is characterized by damage in the spinal cord, dorsal root ganglia (DRG) and cerebellum (17,18). Patients often develop sensory and motor dysfunction at puberty and eventually lose their ability to walk. They also suffer from progressive cardiomyopathy, which is the leading cause of death (19).

FXN downregulation by expanded GAA•TTC repeats is triggered by the formation of an intramolecular DNA triplex, H-DNA, which impedes transcription elongation (20–22), ultimately leading to local heterochromatinization and gene silencing (23). Two types of triplex structures were shown to be formed by pathogenic GAA•TTC repeats *in vitro*: pyrimidine-purine-pyrimidine (H-y) or purine-purine-pyrimidine (H-r) where the third strand is homopyrimidine or homopurine, respectively (24–29). Another, less characterized structure called sticky-DNA involves a triplex formed by two distant GAA•TTC runs (30,31).

GAA•TTC repeats are unstable during germline transmission from parent to offspring, which can result in large-scale repeat expansions or contractions between generations (32,33). Pathogenic GAA•TTC repeats also expand further in somatic cells resulting in disease progression during the affected individual's lifespan (34–36). The most prominent expansions occur in DRG followed by the cerebellum and heart (37). It was hypothesized that formation of triplex DNA could be responsible for GAA•TTC repeat expansions in FRDA patients (38,39). One argument supporting this hypothesis is the fact that a GAAGGA•TCCTTC repeat, which cannot form an H-DNA structure, in the *FXN* gene is stable and results in a very mild and late-onset disease (40).

Mechanisms responsible for GAA•TTC repeat expansions were primarily studied in model experimental systems. First, these repeats were shown to block DNA polymerization *in vitro* (24,41) and stall replication fork progression in bacteria (35), yeast (42), mammalian cell culture (43,44) and patient-derived iPSC cells (45). Further, we have previously developed an experimental system to measure the rate of large-scale expansions of GAA•TTC repeats in yeast (46). A subsequent unbiased, genome-wide genetic screen identified several dozens of genes affecting the rate of repeat expansions, most of which encoded replication fork components (47). Replicative DNA polymerases and proteins involved in Okazaki fragment maturation strongly counteract repeat expansions (48–50). Finally, stabilization of H-DNA formed by GAA•TTC repeats by an RNA transcript (H-loop) additionally increases repeat instability (51). Altogether, these data led us to propose that GAA•TTC repeat expansions in yeast occur while the replication fork struggles to progress through the structure-prone DNA element. Replication was also implicated in the stability of GAA•TTC repeats within the SV40-based mammalian episome (52).

Other machinery implicated in GAA•TTC repeat instability is DNA mismatch repair (MMR). In yeast, the MutL α complex appears to cleave H-DNA, which results in chromosomal fragility in dividing cells (53). Notably, most affected cells in FRDA are post-mitotic, thus, repeat instability in those cells is independent of DNA replication. In an experimental yeast system to study GAA•TTC repeat instability in non-dividing, quiescent cells, MMR counteracted repeat expansions by triggering the formation of deletions or gene conversion events (54). Contrasting results were obtained in a human cell line characterized by progressive, small-scale expansions of GAA•TTC repeats: They appear to be independent of cell division and promoted by the mismatch repair complex MutL γ (55). Similarly, MMR promoted GAA•TTC repeat expansions in iPSCs derived from FRDA patient fibroblasts (56). Expansions of GAA•TTC repeats were also studied in humanized mice. In this system, intergenerational expansions of GAA•TTC repeats were inhibited by a mismatch repair system, while somatic expansions in the cerebellum and DRG were promoted by MMR (57,58).

In this paper, we aimed to study the mechanisms of large-scale expansions of GAA•TTC repeats in human cells. To this end, we designed an experimental system that allowed us to simultaneously analyze DNA replication of GAA•TTC repeats and their large-scale expansions in cultured human cells. It is based on a shuttle plasmid that expresses T-antigen (Tag), driving its extremely efficient replication from the SV40 origin in human cells (43,44). It also contains our previously described cassette for selecting repeat expansions in yeast and can be stably maintained in *Saccharomyces cerevisiae* (46). Repeat expansions accumulated during replication of this vector in mammalian cells are then detected upon its transformation into yeast. A similar strategy has been successfully used in the Lahue lab to analyze mid-scale expansions of CAG•CTG repeats in human astrocytes (59–61). Our system differs in two significant regards: (1) The presence of Tag in the plasmid magnifies its replication in human cells, making the electrophoretic analysis of the replication fork progression feasible; and (2) Our selectable cassette is tuned for studying large-scale repeat expansions.

We confirmed that GAA•TTC repeats cause stalling and reversal of the replication fork in human cells. Remarkably, large-scale expansions of GAA•TTC repeats efficiently occur in this system, making it the first experimental model for human cells. We conducted a candidate gene analysis of large-scale repeat expansions using the siRNA-mediated gene silencing approach. The depletion of 8 of the 20 proteins tested significantly impacted the frequency of GAA•TTC repeat expansions. Notably, those proteins were previously implicated in the unwinding of triplex DNA, fork reversal, and fork restoration (62).

Using our in-house triplex-specific cleavage assay, we showed that H-DNA is efficiently formed by the GAA•TTC repeat in the supercoiled plasmid used for replication/expansion studies in human cells. We have previously reported that chemically modified oligonucleotides, which bind sequence-specifically to the GAA•TTC repeats, disrupt H-DNA formation in supercoiled plasmids (26).

Here, we examined the ability of these oligonucleotides to affect the expansion of GAA•TTC repeats in human cells. We found that the LNA-DNA mixer oligonucleotides, hereafter referred to as LNA-ONs, and the corresponding PNA oligomers dramatically reduce the expansion frequency of GAA•TTC repeats. These data hold promise for the development of these compounds for the treatment of FRDA, which is currently incurable.

Altogether, our data led us to hypothesize that triplex formation by GAA•TTC repeats impairs replication fork progression, ultimately leading to their expansions during replication fork restart.

MATERIALS AND METHODS

Plasmids

The plasmid pJC_GAA100 was constructed by conventional cloning methods in several steps using the pLM113 plasmid as a backbone (44). First, pRS316 (63) was digested by Sall and SacI to excise the ARS4-CEN6 module, which was inserted between Sall and SacI sites of the pML113 plasmid, creating a plasmid called pJW12 (8835 bp-long). The pJC_GAA100 plasmid (12408 bp) was obtained by inserting the AleI-StuI fragment of pYes3-T269-GAA100 (48,64) containing our selectable UR-GAA100-A3-TRP1 cassette into the EcoRV site of pJW12. Note, in the pJC_GAA100 plasmid, GAA repeats are in the lagging strand template for replication from the SV40 origin. The pJC_GAA0 (no repeat) plasmid was obtained with the same approach except for the using AleI-StuI fragment from the no-repeat control pYES-TET644 plasmid (48,64). Plasmids were maintained in the *Escherichia coli* SURE strain (Stratagene), and the length of the repeats in isolated plasmids was confirmed by sequencing.

Oligonucleotides

LNA-ONs and DNA oligonucleotides (ONs) and PNA oligomers were purchased from Eurogentec S.A. The oligonucleotides and oligomers were purified by using reversed phase HPLC, and quality control was performed by using MALDI-TOF mass spectrometry. Control PNA was kindly provided by Prof. Peter Nielsen, Department of Cellular and Molecular Medicine, University of Copenhagen (65).

Cell culture, transfection, siRNA and ON treatment

HEK-293T (ATCC) were grown in Dulbecco's modified Eagle medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS) and MycoZap™ Plus-CL (Lonza). Cells were transfected with the pJC_GAA100 plasmid by using JetPRIME® (Polyplus-transfection) according to the manufacturer's protocol. Briefly, cells were seeded on day 0 and transfected on day 1 with siRNA. On day 2, cells were co-transfected by siRNA together with pJC_GAA100. siRNA-mediated gene silencing was confirmed by Western blot analysis. Alternatively on day 1, cells were co-transfected with the pJC_GAA100 plasmid and LNA-ON, or PNA, of interest. After additional two days, plasmid

DNA was isolated and used for the electrophoretic analysis of replication intermediates (RIs) or transfection into yeast to measure expansion frequencies.

DNA isolation

Plasmid DNA was recovered 48 h post-transfection by a modified Qiagen Miniprep protocol as described in (44). Briefly, cells were washed with PBS and then resuspended in Qiagen Buffer P1 and lysed in 0.66% sodium dodecyl sulfate, 33 mM Tris-HCl, 6 mM EDTA, 66 µg/ml RNase followed by digestion with 0.5 mg/ml proteinase K for 90 min at 37°C. Samples were subject to brief, 30 s, base extraction with 0.75 ml 0.1 M NaOH, and proteins were precipitated upon addition of Qiagen Buffer P3 (4.2 M Gu-HCl, 0.9 M potassium acetate pH 4.8). Cell debris was pelleted at 29,000g for 45 min and supernatant was loaded onto a Qiagen Miniprep spin column. Columns were washed with Qiagen Buffer PB (5 M Gu-HCl, 30% ethanol, adding 10 mM Tris-HCl, pH 6.6) and 0.75 ml Qiagen Buffer PE (10 mM Tris-HCl, pH 7.5, 80% ethanol) and plasmid DNA was eluted using two volumes of 25 µl of Qiagen EB buffer (10 mM Tris-Cl, pH 8.0).

Yeast transformation

Plasmid DNA from human cells was digested for one hour with 50 DpnI (NEB) units to eliminate unrepliated bacterial DNA, EtOH precipitated, resuspended in TE buffer and transformed into yeast SMY537 strain (*MATa*, *leu2-Δ1*, *trp1-Δ63*, *ura3-52*, *his3-200*, *bar1::HIS3*, *can1::KanMX* (64) by the lithium acetate method (66). 5% of each transformation mixture was plated onto SC-Trp plates (synthetic complete, lacking tryptophan), and the remainder 95% onto 5-FOA-Trp (synthetic complete, containing 0.09% 5-FOA and lacking tryptophan) to score for expansions. Colonies on each plate were counted after three days of growth at 30°C. All colonies from 5-FOA-Trp plates were analyzed by single-colony PCR (described below) to verify repeat expansions. Only colonies with PCR-confirmed repeat expansions were tallied in the expansion frequency calculations.

To ensure that expansion events occurred during plasmid replication in human cells and not during yeast transformation, we also established expansion-events baseline by transforming yeast directly with 100 ng of pJC_GAA100 plasmid isolated from *E. coli* without passing through mammalian cells.

PCR analysis

To verify expansions and to determine their size, individual 5-FOA-resistant yeast colonies were disrupted with Lyticase as described in (4). Subsequent PCR amplification by Phusion Polymerase (ThermoFisher) with UC1 (5'-GGTCCCAATTCTGCAGATATCCATCACAC-3') and UC6 (5'-GCAAGGAATGGTGCATGCTCGAT-3') primers flanking the repeat tract was conducted for 35 cycles: of 20 s at 98°C, 2 min at 72°C and 4 min at 72°C. The products were separated into 1.5% agarose gels. PCR product sizes were determined by comparing them with a 50 bp

DNA ladder (New England Biolabs) using ImageLab™ software (Bio-Rad).

Two-dimensional (2D) agarose gel electrophoresis

Plasmid replication intermediates, extracted from human cells as described above, were digested by DpnI, XbaI and BsrGI restriction endonucleases (New England Biolabs), EtOH precipitated and resuspended in TE buffer. The first dimension of electrophoresis was conducted in a 0.4% agarose gel in 1× TBE buffer (89 mM Tris–borate, 2 mM EDTA) at 1 V/cm at room temperature for 19 h. The second dimension was in a 1.3% agarose gel in 1× TBE buffer and was run perpendicular to the first dimension. The dissolved agarose was poured around the excised agarose lane from the first dimension, and electrophoresis was at 5 V/cm in a 4°C cold chamber for nine h in the presence of 0.3 µg/ml ethidium bromide. Gels were washed for 15 min in 0.25 N HCl before an overnight transfer to a charged nylon membrane (Hybond-XL, GE Healthcare) in 0.4 N NaOH. The membrane was hybridized with a ³²P-labeled radioactive probe corresponding to a 533 bp sequence comprising the *URA3* promoter and part of the CDS of the *URA3* gene. Membranes were washed sequentially twice with washing solution I (2× SSC, 0.1% SDS) at 65°C and twice with washing solution II (0.1 × SSC, 0.1% SDS) at 42°C, exposed on IR-sensitive screens for 1–5 days, and detection was performed on a Typhoon Imager (GE Healthcare). At least three independent transformants were tested for each siRNA knockdown.

BQQ-OP mediated DNA cleavage of H-DNA forming GAA•TTC repeats in the presence of LNA-ON oligonucleotides

Plasmids pJC-GAA100, pJC.TTC100 and pJC_GAA0 (1 µg) were incubated with 10 mM of either GAA or CTT LNA-ONs (Supplementary Table S3) in a buffer containing 10 mM sodium cacodylate, 100 mM NaCl and 2 mM MgCl₂ at pH 7.5. As a control, plasmid with no oligonucleotide addition was treated in the same way. The incubations were carried out at 37°C for 16 h. BQQ-OP (1.5 µM) and CuSO₄ (2.25 µM) were mixed and incubated at room temperature for 15 min. The BQQ-OP/CuSO₄ mixture was then added to plasmid/oligonucleotide mixture and incubated at room temperature for 45 min. The cleavage reaction was initiated by adding 2 mM 3-mercaptopropionic acid (MPA) and carried out for 3 h at 37°C. DNA samples were purified with miniprep kit (Qiagen), digested with 1U SacI (Thermoscientific) for 1 h at 37°C and separated in 0.7% agarose gel electrophoresis (90 V, 1 h) followed by SYBR-gold (Invitrogen) staining. Gels were analyzed and quantified with VersaDoc MP system (Bio-Rad) and Quantity One software (Bio-Rad), respectively.

Statistical analysis

At least three independent experiments were conducted for each data point. Statistical analysis was performed via Welch ANOVA test using GraphPad Prism version 8, GraphPad Software, San Diego, California USA.

RESULTS

Experimental system to study large-scale expansions of GAA•TTC repeats in cultured human cells

Our new experimental strategy to study large-scale expansions of GAA•TTC repeats and repeat-mediated fork stalling in cultured human cells is schematically presented in Figure 1. We created a shuttle plasmid (Figure 1A) that encodes Tag for efficient replication from the SV40 origin in human cells, an ARS4-CEN6 module for stable maintenance in yeast, and ColE1 replication origin to conduct cloning in *E. coli*. Importantly, it also contains our well-characterized cassette to detect large-scale repeat expansions, which consists of an artificially split *URA3* reporter gene with the (GAA)₁₀₀•(TTC)₁₀₀ repeat in its intron followed by the *TRP1* gene (*UR-GAA100-A3 TRP1* in Figure 1B). The addition of 10 or more repeats within the intron abrogates splicing of the *URA3* reporter, rendering yeast resistant to 5-fluoroorotic acid (5-FOA^r) (Figure 1B). Note that in our plasmid, the homopurine (GAA)₁₀₀ run is in the lagging strand template for DNA replication from the SV40 origin (Figure 1A), which is the orientation that is most prone for repeat instability (35,42,46,67–72).

This plasmid was first transfected into human HEK-293T cells, where it was allowed to replicate for 48 h. Plasmid DNA was then isolated, treated with the restriction enzyme DpnI to remove unreplicated bacterial plasmid DNA, and transformed into yeast that served as a read-out for repeat expansions occurring in human cells. Yeast transformants were plated onto either synthetic complete medium lacking tryptophan (SC-Trp), or to the SC-Trp medium with 0.09% 5-FOA. As a rule, all colonies on the 5-FOA-containing plates were further analyzed by single-colony PCR to verify repeat expansions (Figure 1C). The latter step was important since besides expansions, mutations and deletions within the *URA3* gene can also result in a 5-FOA-resistant phenotype.

Characterization and genetic controls of GAA•TTC repeat expansions in human cells

Single cell PCR analysis of colonies from 5-FOA plates confirmed that large-scale expansions of GAA•TTC repeats did occur in human cells (Figure 2A). The mean number of added repeats corresponds to 65 (Figure 2B), and we were able to detect up to 300 added repeats when the plasmid replicated in HEK-293T cells. To the best of our knowledge, this is the first system detecting large-scale expansions in mammalian cells after just 48 h in human cell culture.

The frequency of repeat expansions was estimated as the number of 5-FOA-r colonies with PCR-confirmed expansions divided by total number of Trp⁺ transformants using FluCalc calculator (70). These data were compared with the baseline frequency of expansions that either originated in bacteria or in the process of yeast transformation by transforming plasmid DNA isolated from bacteria directly into yeast. Figure 2C shows that the frequency of repeat expansions in HEK-293T significantly (~6-fold) exceeded the baseline frequency.

To identify proteins involved in large-scale GAA•TTC repeat expansions, we used candidate-gene analysis. We and

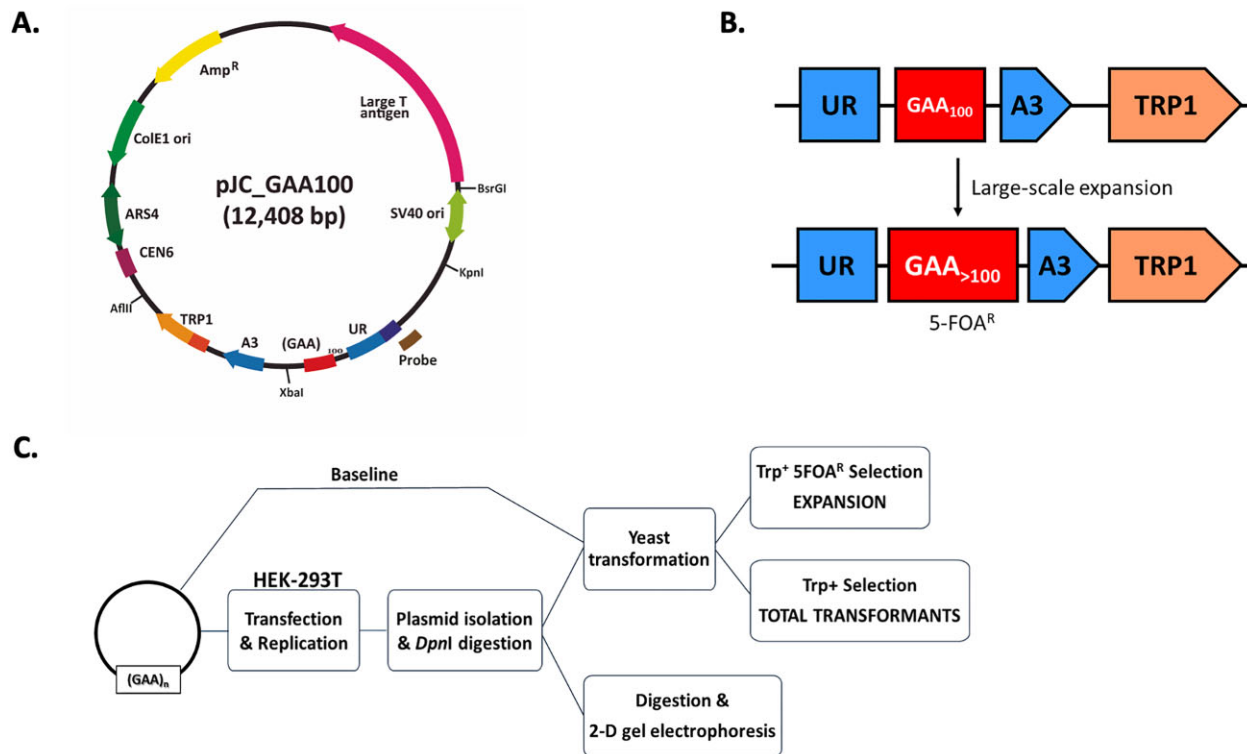


Figure 1. An experimental system to study genome instability and fork stalling caused by GAA•TTC repeats. (A) pJC_GAA100 plasmid used in this study. The relative positions of its most relevant features are indicated inside: The centromeric sequence CEN6, the Autonomous Replication Sequence (ARS4), the ColE1 unidirectional origin (ColE1 Ori), the ampicillin resistance gene (Amp^R), the Large T antigen gene, the SV40 origin of replication (SV40 ori) and the selectable cassette for repeat expansions ($UR-GAA_{100}-A3-TRP1$; depicted in B). Outside, the relative positions of sites recognized by specific restriction endonucleases are indicated. (B) Schematic of the system to select for repeat expansion in yeast. An artificially split *URA3* gene contains 100 GAA•TTC repeats such that expansion events abrogate splicing and result in resistance to a 5-fluoroorotic acid (5-FOA-r). The addition of more than 10 repeats increased the overall length of the intron beyond the splicing threshold. The selectable cassette was cloned into the pJC_GAA100 plasmid. *TRP1* - an auxotrophic marker for the selection of yeast transformants bearing the plasmids. (C) Schematic representation of the assay. Plasmids for the study were transfected into HEK-293T cells, and after culturing them for 48 h, DNA was isolated and digested with DpnI. Repeat expansions were detected upon DNA transformation into yeast. Single-colony PCR was performed for 5-FOA-r clones to confirm expansion events. Expansion frequency was calculated by dividing the number of colonies with repeat expansions by the total number of TRP^+ transformants. To calculate background expansion frequency, pJC_GAA100 plasmid isolated from *E. coli* was transformed directly into yeast. To study DNA replication through the repeats, DNA was digested with appropriate restriction enzymes and replication intermediates were analyzed by 2-dimensional (2D) agarose gel electrophoresis.

others have previously identified multiple genes and proteins involved in repeat expansions in various model systems ranging from yeast (42,53) to human cells (43–45). Different proteins were identified in different studies depending on the repeat type, expansion scale, cell type and an organism (reviewed in (1,71–73)). Notably, most of these proteins were the components of DNA replication and repair machineries. We selected 20 candidate genes that encode for proteins involved in DNA replication and post-replication repair. The selected genes were knocked down using pooled siRNAs (Supplementary Figure S1A, B), followed by measuring repeat expansion frequency.

Our candidate genes can be divided into several functional groups. The first group of genes—FEN1, TIMELESS and CLASPIN—encodes replication fork components previously implicated in repeat expansions and instability. Flap endonuclease 1 (FEN1), which is required for the flap-removal during Okazaki fragments maturation and is involved in various DNA repair pathways, was shown to prevent expansions of multiple repeats in a yeast experimental system (1), while its role

in mammalian cells remains questionable (74,75). TIMELESS (ScTof1) and CLASPIN (ScMrc1) proteins, which are components of the fork-stabilizing complex, were shown to prevent CAG•CTG repeat instability and expansions in yeast and human cells (76,77), as well as GAA•TTC repeat instability in yeast (46,47,77). In our system, however, siRNA knockdown of these proteins did not affect repeat expansion frequency (Supplementary Figure S2A).

The second group of genes—ATR (yMEC1) and ATM (yTEL1)—triggers DNA damage response (DDR) caused by the replication stress. They were previously shown to stabilize CAG•CTG repeats in yeast, mice and human cells (78). siRNA knockdown of those genes in our system did not show a statistically significant effect on GAA•TTC repeat expansions, albeit the depletion of ATM slightly (1.6-fold) increased repeat expansions (Supplementary Figure S2A).

The third group of genes—RAD51, RAD52, BRCA1 and BRCA2—includes the key components of homologous recombination (HR) and DSB-repair (79,80).

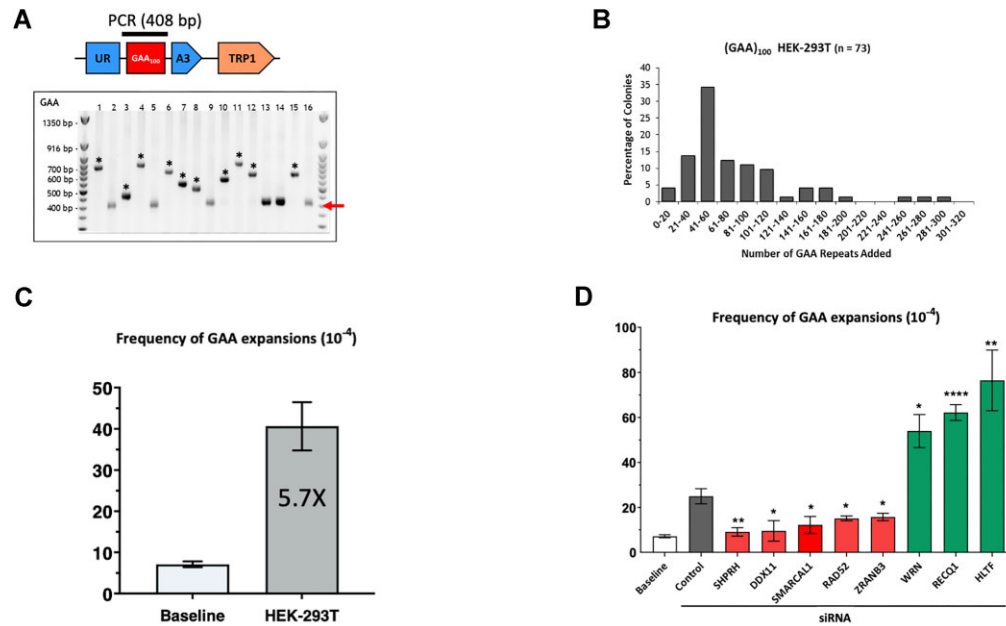


Figure 2. GAA•TTC repeat expansion upon candidate gene knockdown by siRNA. (A) Confirmation of GAA•TTC expansions in HEK-293T cells. Repeat lengths in 5-FOA-resistant colonies from independent experiments were analyzed by PCR. On the upper part, a cartoon representation of a cassette with a bar above the GAA•TTC repeats indicates the product of single-colony PCR, used to determine the repeat length. The red arrow points to the length of the PCR product carrying the starting (GAA)₁₀₀(TTC)₁₀₀ repeat. “*” marks large-scale GAA•TTC expansions. Only clones with expansions were used to calculate expansion frequencies in (D). (B) Distribution of repeats added to the (GAA)₁₀₀(TTC)₁₀₀ repeat during plasmid replication in HEK-293T cells (dark grey bars). The geometric mean of repeats added is 65 repeats (interquartile range 44.5–100.5). (C) Expansion frequencies for pJc_GAA100 plasmid derived from *E. coli* (baseline) and HEK-293T cells. Error bars indicate the standard error of the mean. “*” $P < 0.05$ and “**” indicates a $P < 0.001$ in two-way Welch ANOVA test. (D) Expansion frequencies following siRNA gene knockdown in HEK293T cells. The baseline expansion frequency upon yeast transformation by plasmid DNA isolated from *E. coli* (white bar) is shown for the comparison. Error bars indicate the standard error of the mean. Significance compared to the siControl frequency value was determined using a two-way Welch ANOVA test, “#” $P < 0.05$ versus baseline, “*” $P < 0.05$, “**” indicates a $P < 0.001$ versus siControl. See Supplementary Table S1 for details.

Homologous recombination has been implicated in both the promotion and suppression of repeat expansions and instability in various experimental systems (81). In our case, knocking down RAD51, BRCA1 or BRCA2 did not affect GAA•TTC repeat expansions (Supplementary Figure S2A). At the same time, knocking down RAD52 significantly decreased GAA•TTC repeat expansions (Figure 2D). Besides canonical HR (82–84), RAD52 is also implicated in the pathway of DSB repair called break-induced replication (BIR) (85,86). BIR, and specifically RAD52, were shown to promote CAG and CGG repeat expansions in yeast and cultured human cells (87,88). POLD3 (yPol32), a small subunit of DNA polymerase δ , is an essential protein required for DNA synthesis during BIR (89). In our system, the depletion of POLD3 did not change the frequency of GAA•TTC repeat expansions, arguing against the role of BIR. Finally, RAD52 was shown to protect reversed replication forks from degradation by exonucleases (90). We believe that this function of RAD52 might be essential for GAA•TTC repeat expansions in our system (see below and Discussion).

The fourth group of genes—HLTF and SHPRH—encodes ubiquitin ligases involved in the poly-ubiquitination of PCNA critical for DNA damage tolerance, specifically fork reversal and template switching (91,92). The yeast homolog of these genes, RAD5, was shown to promote GAA•TTC and ATTCT•AGAAT repeat expansions (46,93). In addition, the helicase activity

of HLTF has been implicated in fork reversal (94). The knockdown of HLTF, but not SHPRH, has been shown to elevate CAG repeat expansions several folds in human cells (95). Here we show that HLTF knockdown increases GAA•TTC repeat expansion frequency (Figure 2D), similarly to what was observed for CAG repeats. At the same time, the knockdown of SHPRH dramatically reduced GAA•TTC repeat expansion frequency (Figure 2D).

The fifth group of genes—SMARCAL1 and ZRANB3—encodes SWI/SNF helicases and ATPases, which catalyze replication fork reversal and restart (96–100). Consequently, depletion or inactivation of these proteins hinders the ability of a replication fork to recover from replication stress, particularly by increasing the frequency of double-strand breaks (96,101–103). We were interested in these proteins because of their role in replication fork reversal and restart, the role of which in promoting repeat expansions in various experimental systems has been widely discussed (1,44,104,105). In line with those data, depletion of either SMARCAL1 or ZRANB3 proteins prevented GAA•TTC repeat expansion in our system (Figure 2D).

The sixth group, DDX11, FANCI, and RTEL1 are members of a family of iron-sulfur-containing DNA helicases (106,107) that prevent replication stress and mediate HR repair by directly interacting with Pol δ (108,109). We were particularly interested in the DDX11 helicase, as we and others have previously shown that it unwinds triplex

structures *in vitro* (110,111). Here, we show that the depletion of DDX11 dramatically decreased repeat expansions (Figure 2D). At the same time, depletion of FANCI or RTEL1 did not affect GAA•TTC repeat expansions (Supplementary Figure S2A).

We then look at two members of the RecQ family DNA helicases: RECQ1 and WRN that were implicated in the replication fork restart. RECQ1 helicase interacts with PCNA, RPA and DNA polymerase δ (112–117), and has a 3'-5' directed DNA unwinding capacity restoring reversed forks to their original three-armed configuration *in vitro* and *in vivo* (118). WRN helicase is involved in resolving a variety of DNA substrates: three-way junctions, replication forks, flaps, D-loops, bubbles, Holliday junctions, and G-quadruplexes (G4) (119,120). Here, we show that the depletion of both RECQ1 and WRN helicases increases GAA•TTC repeat expansions (Figure 2D). Besides the RecQ family of helicases, DNA2 is also involved in the processing of reversed forks mediating DNA end resection together other proteins (121). In our system, the depletion of DNA2 did not influence repeat expansions (Supplementary Figure S2A).

As for the mismatch repair, the *MLH1* gene is known to be epigenetically silenced by promoter hypermethylation in HEK-293T cells (122). This is not the case, however, in the parental HEK-293 cells. In our system, the frequency of large-scale expansions did not drastically differ between the two cell lines (Supplementary Figure S2B). This argues against the role of MutL complexes in the expansion process in our system, but it does not rule out other components of the MMR machinery (see Discussion).

In summary, out of 20 candidates tested, five proteins, SHPRH, DDX11, SMARCAL1, RAD52 and ZRANB3 appeared to promote GAA•TTC repeat expansions in our experimental human system, while three proteins, HLTf, WRN and RECQ1 appeared to counteract them (Figure 2D and Supplementary Table S1).

Replication fork progression through the GAA•TTC repeat

We and others have previously shown that expanded GAA•TTC repeats stall replication fork progression in yeast and human cells (42–45). Here, to compare repeat expansions with their replication, we analyzed the replication fork progression through the (GAA)₁₀₀•(TTC)₁₀₀ repeat using 2D electrophoretic analysis of replication intermediates. In brief, DNA plasmids with and without the (GAA)₁₀₀•(TTC)₁₀₀ repeat were transfected into HEK-293T cells, replication intermediates were isolated 48 h post-transfection, treated with DpnI to get rid of unreplicated bacterial plasmid DNA, digested by BsrGI and XbaI restriction endonucleases (Figure 3A) and separated by 2D agarose gel electrophoresis followed by Southern hybridization. This digestion scheme deliberately positioned the repeat on the descending part of the replicative Y-arc (Figure 3B).

The simulation program for 2D gels (123) was used to predict the shape of the replication intermediates (RIs) responsible for the patterns observed. The pattern detected for the control plasmid corresponded to that expected for unconstrained replication of the circular plasmid, where the

initiation occurs at SV40 origin in a bi-directional manner (simple-Ys's shape) (Figure 3B). However, the repeat-containing plasmid produced a different pattern. In this case, replication would initiate at the SV40 origin bidirectionally, creating a bubble. The leftward (counterclockwise in Figure 1A) moving fork progresses unconstrained, while the rightward (clockwise in Figure 1A) moving fork stalls at GAA•TTC repeats, resulting in the accumulation of simple Y RIs with a mass of $\sim 1.65\times$. If this stalling is indefinite, replication will be completed by the leftward replication fork, leading to the accumulation of double-Y RIs migrating above the descending arc. The stalled fork can also reverse, forming chicken feet RIs, which should migrate somewhere in between the stall site and double-Y RIs. The fork reversed at the repeat can be restored, or plasmid replication will be completed by the opposite replication fork (104).

Experimental data shown in Figure 3B clearly demonstrate the presence of the stall site at the expected position on the descending arm and the tangent curve exiting from the stall site linking it to a spindle-shape spot. Based on its electrophoretic mobility, this spindle spot likely represents a mixture of stalled and/or reversed forks, replication of which is completed by the converging fork resulting in double-Y RIs. These results confirm that GAA•TTC repeats strongly stall replication in HEK-293T cells leading to fork reversal and completion of replication by an opposite replication fork in agreement with (44). Note that previously, we and others analyzed GAA•TTC-mediated stall sites on the ascending portion of the Y-arc. This setting makes it somewhat difficult to quantitatively analyze intermediates at different stages of fork reversal, as they can migrate closely to the ascending arc (compare Figures 3B and Supplementary Figure S3).

To quantify our results, we first normalized the signal at the stall site on the descending arm to the signal of the replication arc underneath. This quantification showed that $\sim 20\%$ of all replication forks stall at the (GAA)₁₀₀•(TTC)₁₀₀ repeat, while no stalling at this position was observed for the no-repeat control plasmid (Figure 3C). The spindle-shaped spot was quantified by comparing its intensity to the rest of the Y-arc (Supplementary Figure S4). It accounted for a $\sim 17\%$ of RIs in the repeat-containing plasmid and it was not present at all in for the no-repeat control plasmid (Figure 3D).

We then analyzed whether proteins that affected GAA•TTC repeat expansions in our system (Figure 2D) influenced replication fork progression through this repeat as well. To this end, we used 2D electrophoretic analysis of RIs isolated from cells depleted of RAD52, ZRANB3, SMARCAL1, SHPRH, HLTf, DDX11, WRN and RECQ1 proteins via siRNA (Figure 4A). Note that depletion of several of these proteins changed the shape of the stall site on the descending arc from the well-defined spot observed in the non-treated cells (Figure 4A control) to a more dispersed signal (see ZRANB3, WRN, RECQ1, SMARCAL1 in Figure 4A). This made the comparison of the stall sites between different siRNA treatments somewhat ambiguous. Thus, we decided to quantitate spindle-shaped RIs that were clearly present in every case, which we considered a proxy for termination RIs, in which

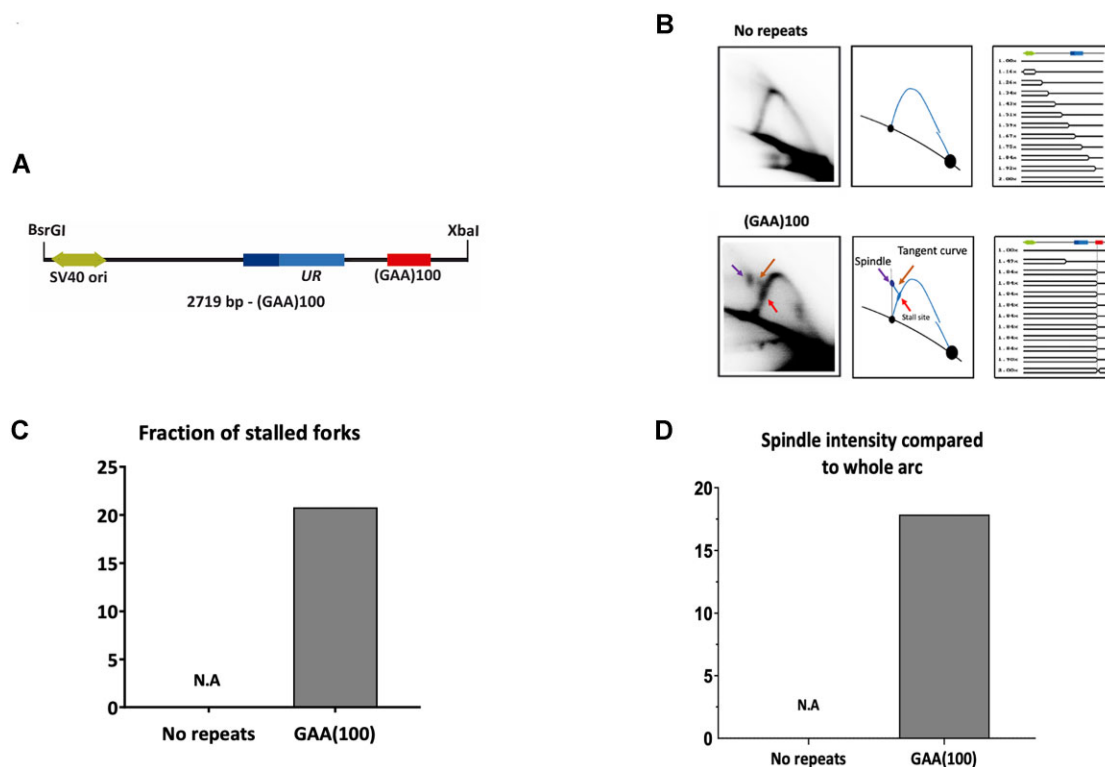


Figure 3. Analysis of pJC.GAA replication in HEK-293T cells by 2D gel electrophoresis. **(A)** Linear map of pJC.GAA100 restriction fragment. **(B)** Representative 2D gels of replication through zero and 100 GAA•TTC repeats are shown in the far-left column with their corresponding interpretative diagrams to their right. DNA was isolated, digested with BsrGI, XbaI, and DpnI, and analyzed by 2D gel electrophoresis. The simulation program 2D gel (123) was used to predict the shape of twelve consecutive replication intermediates (RIs). A linear map is shown on top of each series of RIs, showing the relative positions of SV40 ori (green) UR (blue). The red arrow points to the location of the stall at the (GAA)₁₀₀•(TTC)₁₀₀ repeat. **(C)** Quantification of the fraction of stalled forks. The ratio of radioactivity in the peak area to that corresponding area of a smooth replication arc reflects the extent of replication slowing. Quantification was done with ImageJ (NIH). N.A., non-applicable. **(D)** Quantification of spindle spot intensity compared to the whole arc. The ratio of radioactivity signal at the spindle spot was compared to the radioactivity signal of the whole arc on Image Lab®. N.A., non-applicable.

one of the two forks was reversed at the GAA•TTC repeat, but not properly restored (Figure 4B and Supplementary Table S2). This reasoning is in-line with the data that effective reversal of one fork is a pre-requisite for the opposite fork to reach the termination site (124). The fraction of those RIs was strongly decreased upon SHPRH and ZRANB3 depletion, but increased upon DDX11, SMARCA1 and RECQ1 depletion.

In sum, five out of the eight proteins that affected repeat expansions in human cells also changed the character of the replication fork progression through the GAA•TTC repeat. This partial correlation points to a link between replication and large-scale expansions of GAA•TTC repeats in human cells.

LNA-ONs and PNA oligomers bind and disrupt the intramolecular triplex formed by the (GAA)₁₀₀•(TTC)₁₀₀ repeat

We have previously shown that LNA-modified ONs, which are complementary to either the CTT or GAA strand (GAA LNA-ONs and CTT LNA-ONs, respectively) (Supplementary Table S3) interact with expanded GAA•TTC repeats derived from FRDA patient cells. The GAA LNA-ON significantly disrupted and abolished the pyrimidine-motif H-DNA (H-y DNA) present in supercoiled DNA (26) by

forming a duplex-invasion complex. On the other hand, the CTT LNA-ON was able to bind either the remaining GAA single strand of the H-DNA forming Watson-Crick hydrogen bonds or the B-DNA duplex, which resulted in the formation of an intermolecular triplex. These studies were performed using a different plasmid where longer GAA•TTC repeats and additional *FXN* derived flanking sequences were included.

To assess the ability of the (GAA)₁₀₀•(TTC)₁₀₀ repeat in the pJC.GAA100 plasmid to form a triplex structure we used the BQQ-OP triplex-specific cleavage assay. We have previously reported H-DNA formation at FRDA expanded GAA•TTC repeats using the BQQ-OP triplex-specific cleavage assay (26,28,125). BQQ-OP is a low-molecular weight compound that consists of a benzoquinoxaline derivative (BQQ) conjugated to a 1,10-*ortho*-phenanthroline (OP) (Figure 5D). BQQ-OP recognizes and intercalates into both inter- and intramolecular triplex structures, which produces double-strand DNA breaks specifically at the site of triplex formation. Furthermore, we have used the BQQ-OP assay to examine the sequence-specific binding of LNA-ONs and PNA oligomers to the triplex-forming (GAA)₁₀₀•(TTC)₁₀₀ repeat (Supplementary Table S3).

In the case of H-DNA or intermolecular triplex formation at the (GAA)₁₀₀•(TTC)₁₀₀ repeat DNA sequence in the

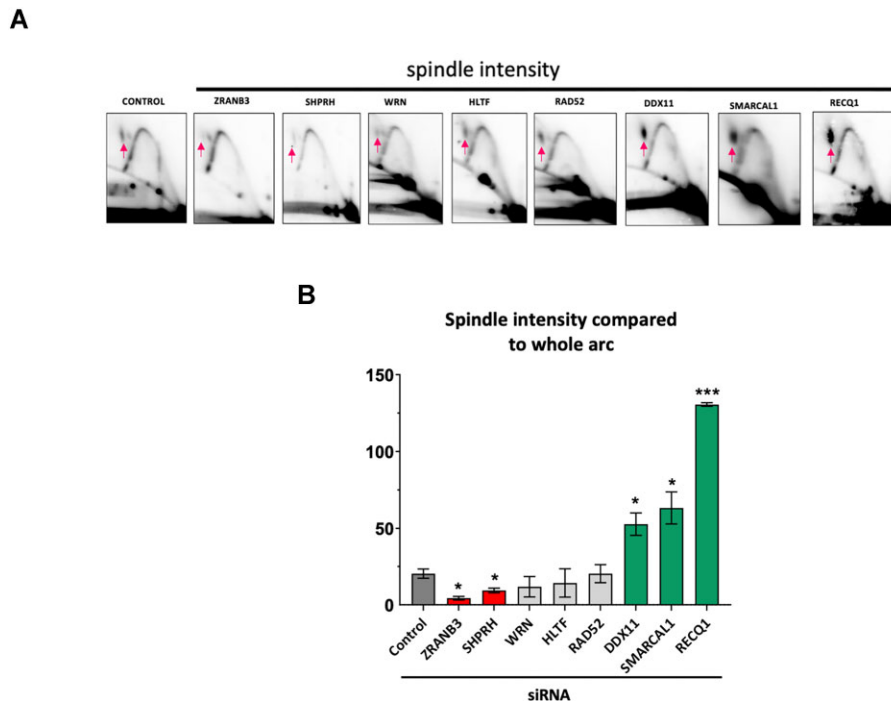


Figure 4. Analysis of replication through GAA•TTC repeats by two-dimensional (2D) agarose gel electrophoresis upon candidate gene knockdown by siRNA. (A) Representative 2D gels of replication through the (GAA)₁₀₀•(TTC)₁₀₀ repeats in HEK-293T cells. DNA was isolated, digested with BsrG1, DpnI and XbaI, and analyzed by 2D gel. The red arrow points to the location of the stalling at the GAA•TTC repeats. At least three experiments were analyzed for each siRNA treatment. (B) Quantification of the spindle spot at the replication fork. The ratio of radioactivity signal at the spindle spot was compared to the radioactivity signal of the whole arc using Image Lab®. Error bars indicate the standard error of the mean. ‘*’ $P < 0.05$, ‘***’ indicates a $P < 0.0001$ versus siControl. See Supplementary Table S2 for details.

supercoiled pJC_GAA•TTC plasmid, the BQQ-OP triplex-specific double-strand DNA cleavage followed by a unique site cleavage using the SacI restriction enzyme would produce two DNA fragments (~8431 and 3997 bp-long), which are marked as X and Y in Figure 5A. On the other hand, treatment of the control plasmid, which lacks the repeat sequence (pJC-GAA0), with BQQ-OP and SacI should only produce a single, 12 406 bp-long linear DNA fragment corresponding to the full-length plasmid cleaved by SacI (Figure 5A). Indeed, the two expected DNA fragments were detected after BQQ-OP treatment of pJC.GAA100 plasmid, confirming triplex formation at the repeat (Figure 5B, lane 3). At the same time, the control plasmid did not show any cleavage by the BQQ-OP (Figure 5B, lane 6).

Incubation of the repeat-containing plasmid with GAA LNA-ON dramatically (~10-fold) decreased BQQ-OP cleavage (Figure 5B, lane 1). These data show that the GAA LNA-ONs essentially prevented H-DNA formation, in agreement with our previous findings (26,126). On the other hand, CTT LNA-ON only marginally affected the triplex-specific DNA cleavage (Figure 5B, lane 2). It is important to note that BQQ-OP cannot distinguish between an intermolecular triplex, where the CTT LNA-ON acts as the third strand targeting the double-strand DNA repeat and the H-DNA (26,126). The difference between the current and previous CTT LNA-ONs data (26) could be due to the variation in the length of the GAA•TTC repeat and the presence of flanking DNA sequences, which might affect the nature and/or the stability of the triplex.

We obtained similar results using GAA PNA and CTT PNA oligomers (Supplementary Figure S5).

LNA-ONs inhibit GAA•TTC repeat expansions occurring during replication in human HEK-293T cells

As discussed above, the frequency of the (GAA)₁₀₀•(TTC)₁₀₀ repeat expansions increased ~6-fold over the baseline upon plasmid replication in HEK-293T cells ($P = 0.0043$). Here we examined the effect of the chemically modified oligonucleotides, which we previously reported to affect H-DNA formation, on repeat expansions in human cells. To this end, HEK-293T cells were transfected with the pJC.GAA100 plasmid preincubated with either the GAA LNA-ON or CTT LNA-ON (Supplementary Table S3). Strikingly, both oligonucleotides reduced the frequency of repeat expansions essentially down to the baseline level, whereas a control oligonucleotide, which does not bind to the target sequence, had no effect (Figure 6A). We conclude, therefore, that repeat-specific LNA-DNA mixmer oligonucleotides prevent repeat expansions in human cells.

Inhibition of GAA•TTC repeat expansions in HEK-293T cells by PNA oligomers

To examine whether binding of another class of modified oligomers to the (GAA)₁₀₀•(TTC)₁₀₀ repeat would also affect its expansion frequency in human cells, we studied peptide nucleic acid (PNA) oligomers designed to bind this

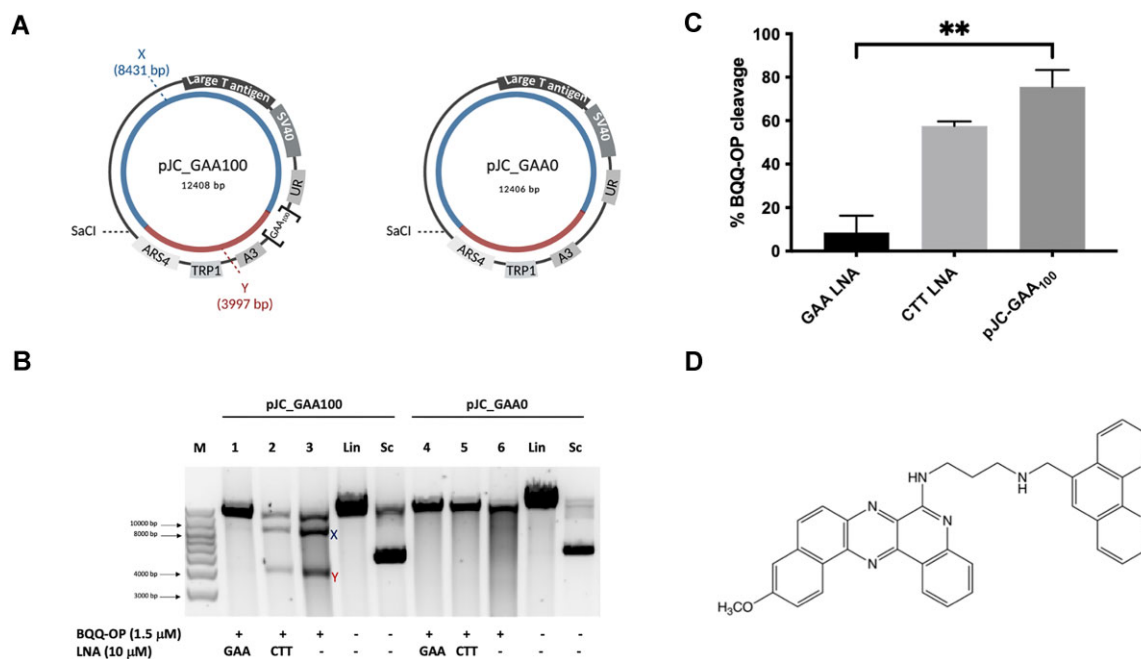


Figure 5. BQQ-OP mediated DNA cleavage of the $(GAA)_{100}\bullet(TTC)_{100}$ repeat in the presence of LNA oligomers. (A) Schematic illustration of pJC-GAA100 and pJC-GAA0. The H-DNA forming site in pJC-GAA100 is indicated as $[GAA]_{100}$; there is no H-DNA forming site in pJC-GAA0 plasmid. The two DNA fragments generated triplex cleavage by BQQ-OP (benzoquinoxaline 1,10-ortho-phenanthroline) followed by unique site restriction digestion are indicated as X (8431 bp) and Y (3997 bp) in pJC-GAA \bullet TTC and the same reaction would result in a linearized fragment only in pJC-GAA0. (B) Representative agarose gel for pJC-GAA100 and pJC-GAA0 plasmids incubated with 10 μ M GAA (LNA-DNA) mixmer (lanes 1 and 4 respectively), or CTT (LNA-DNA) mixmer (lanes 2 and 5 respectively), or in the absence of LNAs (lanes 3 and 6 respectively). BQQ-OP-mediated triplex specific cleavage of pJC-GAA100 and pJC-GAA0 was performed in the presence of Cu^{2+} and 3-mercaptopropionic acid (MPA) followed by unique site restriction digestion with *Sac*I. As controls, supercoiled (Sc) and linearized (Lin) variants of both plasmids and molecular weight DNA ladder (M) are shown. (C) Graph showing the percentage of BQQ-OP-mediated triplex specific cleavage of pJC-GAA100 in the presence of GAA and CTT (LNA-DNA) mixmers or in the absence of LNA-DNA mixmers (pJC-GAA100). The values represent the ratio between the intensity of DNA double strand cleavage (X + Y) to the total band intensity of the particular lane and are shown as mean with S.D. ($n = 2$). No cleavage was obtained in pJC-GAA0 and not included in the graph. ** indicate $P \leq 0.01$ compared to the plasmid in the absence of LNA oligomers. (D) Chemical structure of benzoquinoxaline 1,10-ortho-phenanthroline (BQQ-OP).

repeat. PNAs are DNA mimic oligomers with a pseudopeptide backbone (65) that are capable of binding and invading dsDNA in a sequence-specific manner and with high affinity owing to the lack of phosphate repulsion (127). We have previously studied the molecular interaction and binding mode of GAA- and CTT- PNAs (Supplementary Table S3, GAA-PNA and CTT-PNA, respectively) to expanded GAA \bullet TTC repeats in plasmid DNA *in vitro* (26). We have showed that GAA-PNA binds via a duplex-invasion mechanism and completely prevent H-DNA formation by the FRDA expanded repeats. CTT- PNA, on the other hand, formed either a triplex-invasion complex or a Watson-Crick duplex when binding to the complementary polypurine strand of the DNA duplex (26).

Using the same experimental setting as in the previous section, HEK-293T cells carrying the replicating pJC_GAA100 plasmid were incubated with either GAA-PNA or CTT-PNA. This was followed by plasmid DNA isolation and transformation into yeast to detect the frequency of expansion events that accumulated in human cells. Figure 6B shows that similarly to LNA-ONs, both PNA oligomers reduced repeat expansions in human cells to the baseline level, while a control PNA oligomer, which does not bind to the target sequence had only a modest inhibitory effect, which is not statistically significant. We conclude, therefore,

that repeat-specific PNAs abolish expansion of GAA \bullet TTC repeats in human cells.

Dose dependent inhibition of GAA \bullet TTC repeat expansions in human HEK-293T cells by BQQ

Our genetic analysis showed that DDX11 helicase is needed for GAA \bullet TTC repeat expansions in human cells (Figure 2D). Since this helicase is known to unwind a triplex structure formed by the GAA \bullet TTC repeat *in vitro* (111), we reasoned that extra-stabilization of triplex DNA by chemical compounds might prevent expansions as well. To test this hypothesis, we employed the benzoquinoxaline derivative (BQQ) shown in Figure 7B. We have previously reported on the efficiency of this heterocyclic compound to bind DNA polypurine \bullet polypyrimidine sequences once they form an inter- or intramolecular triplexes. BQQ intercalates in the triplex structures with its aminopropyl side chain positioned in the minor groove (128,129). BQQ has also been used to demonstrate the effect of H-DNA formation on transcription in a reporter model in bacteria (130) as well as triplex formation on a genomic level in mammalian cells (111) and its dissociation by the DDX11(ChIR1) helicase.

We treated HEK-293T cells transfected with the pJC_GAA \bullet TTC100 plasmid with BQQ at different

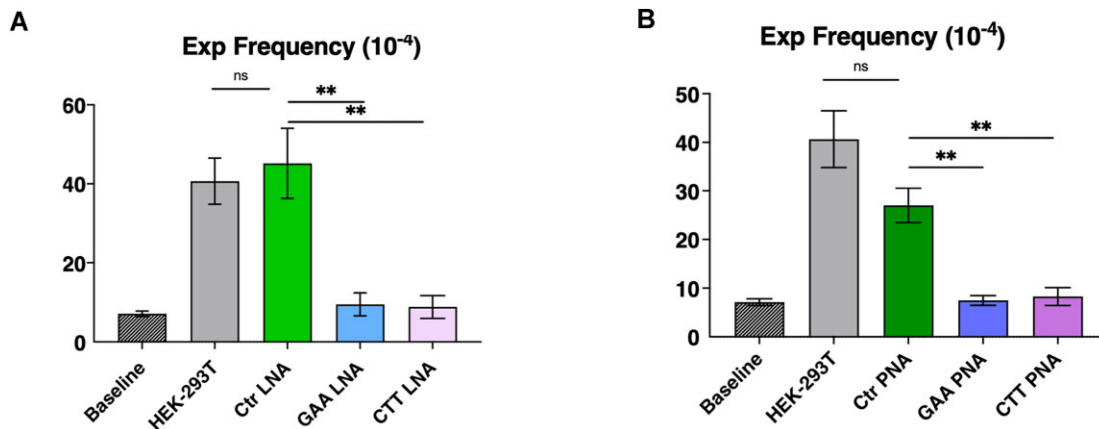


Figure 6. Inhibition of GAA•TTC repeat expansions occurring during replication in human HEK-293T cells by LNA oligonucleotides and PNA oligomers. **(A)** Baseline represents the frequency of the (GAA)₁₀₀•(TTC)₁₀₀ repeat expansion when bacterial pJC_GAA100 plasmid transformed directly into yeast. Other bars in the chart show the frequency of GAA•TTC repeat expansions occurring during plasmid replication in HEK-293T cells in the presence or absence of various LNA-DNA mixmers. Error bars indicate standard error of the mean. Significance relative to the Scramble LNA-DNA mixmer control sequence frequency value was determined using a two-way Welch ANOVA test. ***P* ≤ 0.05. **(B)** Baseline represents the frequency of the (GAA)₁₀₀•(TTC)₁₀₀ repeat expansion upon bacterial pJC_GAA100 plasmid transformed directly into yeast. Other columns show the frequency of GAA•TTC repeat expansions that occurred during plasmid replication in HEK-293T cells in the presence or absence of various PNAs. Error bars indicate standard error of the mean. Significance relative to the Scramble PNA sequence frequency value was determined using a two-way Welch ANOVA test. ***P* ≤ 0.005.

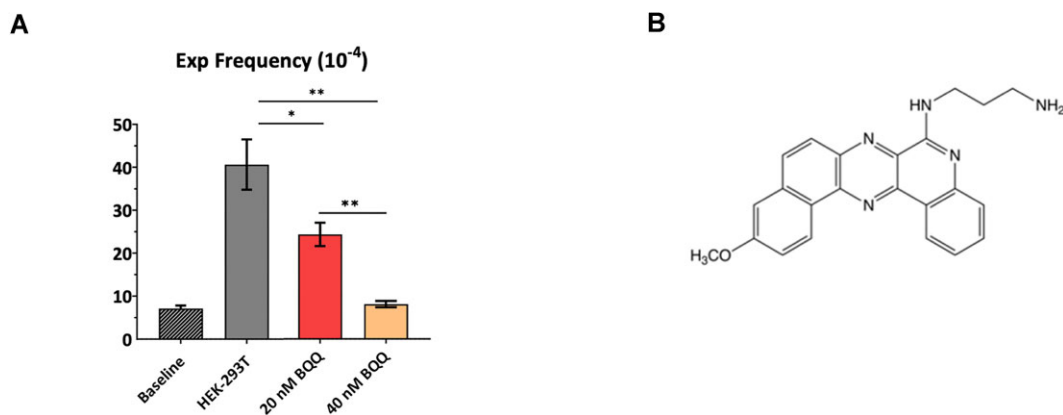


Figure 7. Dose-dependent inhibition of GAA•TTC•TTC repeat expansions in human HEK-293T cells by BQQ. **(A)** Baseline represents the frequency of the (GAA)₁₀₀•(TTC)₁₀₀ repeat expansion when bacterial pJC_GAA100 plasmid was transformed directly into yeast. Other columns show the frequency of GAA•TTC repeat expansions that occurred during plasmid replication in HEK-293T cells at two BQQ (benzoquinoxaline) concentrations (20 and 40 nM). Error bars indicate standard error of the mean. Significance relative to the HEK-293T frequency value was determined using a two-way Welch ANOVA test (for 20 nM BQQ-OP concentration *P* value is ≤ 0.05 for 40 nM, *P* ≤ 0.005). **(B)** Chemical structure of Benzoquinoxaline derivative (BQQ).

concentrations. Figure 7A shows that BQQ treatment significantly decreased the frequency of the GAA•TTC repeat expansion as compared to untreated cells. The BQQ inhibitory effect was dose-dependent: ~2-fold at 20 nM, and down to the baseline level at 40 nM. We conclude, therefore, that strong BQQ-mediated stabilization of the triplex formed by the (GAA)₁₀₀•(TTC)₁₀₀ repeat impedes repeat expansion in human cells.

DISCUSSION

In this study, we describe an experimental system to study GAA•TTC repeat instability, in which repeat expansions that occurred during replication of the SV40-based plasmid in human cells were detected upon transformation into yeast (Figure 1). Remarkably, we observed that large-scale

repeat expansions occurred very efficiently in this system: up to 300 repeats could be added, while the mean number of added repeats was 65 (Figure 2B). Our experimental setting is similar to that previously developed by Lahue *et al.* for studying CAG•CTG repeat instability (59–61) that was tuned to detect mid-scale (10-to-15 trinucleotides) expansions. The only known instance where massive repeat expansions were observed in a mammalian experimental system was a specific transgenic DM1 mouse (131), yet the reasons for big jumps in the number of CTG repeats in these mice remains unclear. Thus, our experimental system is unique in allowing the study of mechanisms and genetic controls of large-scale repeat expansions in human cells.

We and others have previously found that GAA•TTC repeats cause replication fork stalling in various experimental systems, including SV40-based plasmids (42–45). In

accord with these observations, we detected repeat-mediated replication fork stalling and reversal in our experimental plasmid (Figure 3). This data prompted us to investigate various candidate genes that affect DNA replication and post-replicative repair, and/or were previously implicated in repeat expansions in different experimental systems. Altogether, we analyzed twenty candidates and found that siRNA depletion of eight candidates significantly affected the frequency of repeat expansions (Figure 2D). Strikingly, seven of these candidate genes (SHPRH, RAD52, RECQ1, SMARCAL1, WRN, HLF and ZRANB3) were previously implicated in replication fork reversal/restoration (97–100,118–120,132–141), while the remaining DDX11 is a DNA-helicase unwinding DNA triplex structures *in vitro* (110,111).

We then looked at changes in the replication fork progression through the GAA•TTC repeat upon depletion of the proteins encoded by these eight candidate genes. Inactivation of SHPRH and ZRANB3 decreased, while DDX11, SMARCAL1 and RECQ1 depletion increased the intensity of termination RIs, in which one of the two forks is reversed at the GAA•TTC repeat, but not properly restored. While this is only a partial correlation, a combination of genetic and biochemical data point to the role of the replication fork stalling, reversal, and restart in GAA•TTC repeat expansions during plasmid replication in human cells.

The instability of FRDA GAA•TTC repeats and their role in the disease progression was linked to their ability to form triple-helix H-DNA conformation (reviewed in (142)). This structure is formed when DNA strands in half of the repeat dissociate, and one of them (either polypurine (R) or polypyrimidine (Y)) folds back to form Hoogsteen or reverse-Hoogsteen hydrogen bonds with the remaining duplex, while the complementary strand remains single-stranded (25,31,143). Formation of both H-y and H-r triplex was reported for GAA•TTC repeats at various ambient conditions in supercoiled DNA (25,28,144,145). Here using the BQQ-OP triplex-specific cleavage assay, we confirmed that the (GAA)₁₀₀•(TTC)₁₀₀ repeat forms an H-DNA in our supercoiled pJC.GAA100 plasmid *in vitro* (Figure 5). We reasoned, therefore, that if triplex formation by the GAA•TTC repeat is responsible for its instability, chemical compounds that interfere with triplex formation should affect repeat instability.

We, thus, examined the effect of two distinct classes of modified oligonucleotides LNA-ONs and PNA oligomers, on the repeat expansion frequency. PNA oligomers or LNA-modified oligonucleotides were previously shown to invade the DNA duplex at FRDA GAA•TTC repeats and completely dissolve the preformed triplex structures (26). Here we confirmed that GAA LNA-ONs and GAA-PNA abolished H-DNA formation. In our previous study, we were able to show that CTT LNA-ONs enhanced the amount of triplex formed in the plasmid with FRDA GAA•TTC repeats due to either binding to the remaining GAA single-strand of the H-DNA (H-y) or to the duplex DNA forming an intermolecular triplex. In both cases, CTT LNA-ON binding would lead to the two DNA fragments produced by the BQQ-OP double-strand cleavage. Although the detected effect in the BQQ-OP *in vitro* assay is lower in the current study, our results are still in

agreement with the previously reported findings. On the other hand, CTT-PNA, as we have reported earlier, form a triplex invasion complex with expanded GAA•TTC repeats in supercoiled plasmids, which would not be possible to detect using the BQQ-OP assay analyzed by agarose gel electrophoresis. In all cases, it is important to remember that the triplex structures formed during replication are different from those formed in supercoiled plasmids *in vitro*. Nevertheless, the *in vitro* assay demonstrates the capacity of the LNA-DNA oligonucleotides and PNA oligomers to target the GAA•TTC repeats, either through DNA duplex invasion or triplex formation.

Strikingly, we found that both GAA- and CTT- LNAs and PNAs dramatically decreased the expansion frequency in mammalian cells as compared to control oligomers. These data argue against the role of preformed H-y DNA in repeat instability in our system. Rather, both oligomers could prevent transient triplex formation by the repeat in front of the fork (42) or upon fork reversal. Also, BQQ that is known to stabilize DNA triplex structures, counteracted GAA•TTC repeat expansions in a dose-dependent manner. Altogether, chemical compounds that interact with triplex-forming repeat sequences appeared to preclude repeat expansions.

Genetic control results, electrophoretic analysis of replication intermediates, and triplex stability data altogether led us to speculate on the model of GAA•TTC repeat expansions in our system (Figure 8). During SV40-based replication, the T-antigen serves as a helicase, both the leading and lagging strands are synthesized by DNA Pol δ (146,147), and there is a relatively poor coupling between the helicase (T-antigen) and DNA-polymerase delta (Pol δ) (148), making it prone to fork reversal and restart (148,149). We propose that a triplex transiently formed by the GAA•TTC repeat during DNA replication in front of the fork causes its stalling (Figure 8A), ultimately leading to repeat expansions during replication restart and completion. Triplex formation can be counteracted by LNA-DNA mixmers, PNAs and DDX11 helicase, thus, precluding subsequent steps leading to repeat expansions.

Triplex-mediated fork stalling in the SV40-based plasmids was previously shown to cause fork reversal (44). Our data show that depletion of SHPRH (91), involved in PCNA ubiquitination, and of DNA helicase ZRANB3 reduces repeat expansions and decreases corresponding RIs (Figures 2D and 4B), implicating the two proteins at this step (Figure 8B). Also, RAD52, which is known to protect reversed forks, promotes repeat expansions in our system (90). We suggest, therefore, that fork reversal driven by these proteins is a prerequisite for repeat expansions. In contrast, the presence of triplex stabilizing BQQ makes triplex unwinding and fork reversal difficult, precluding repeat expansions. We are aware that extensive fork reversal usually requires RAD51 protein (150,151). Since we don't see any effect from the depletion of RAD51 on repeat expansions, we speculate that fork regression in our case might be limited to the repetitive DNA segment.

We show that inactivation of the RECQ1 helicase increases repeat expansion frequency and causes accumulation of RIs corresponding to reversed forks (Figures 2D and 4B). Since RECQ1 is known to efficiently restore

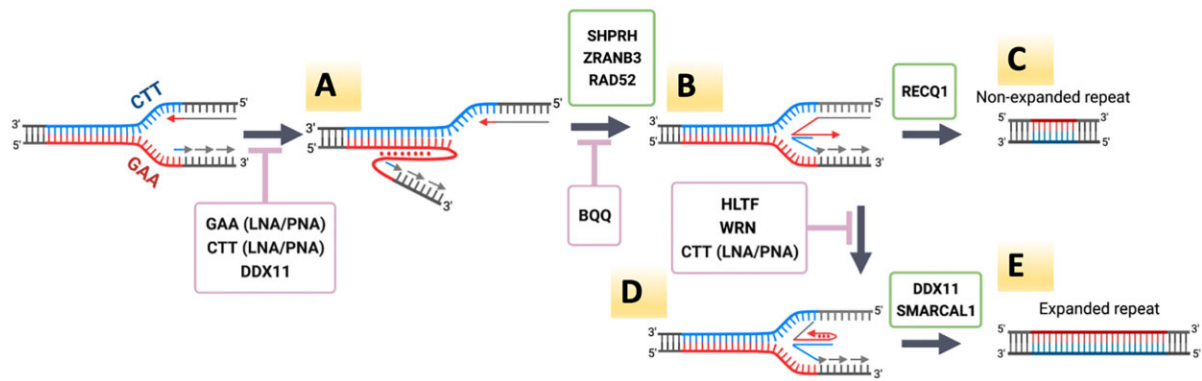


Figure 8. Proposed model for GAA•TTC large-scale expansions. (A) The (GAA)₁₀₀•(TTC)₁₀₀ repeat (red – purine strand, blue – pyrimidine strand) can form a triplex during the replication causing fork stalling. This process might be counteracted by repeat-specific LNA-ONs and PNAs, as well as by DDX11 helicase. (B) Fork reversal is promoted by SHPRH and ZRANB3, while triplex stabilization by BQQ prevents fork reversal, and the reversed fork is protected from degradation by the RAD52. (C) The reversed fork can be restored via an error-free mechanism involving RECQ1 helicase. (D) Occasionally, a nascent leading strand can fold back to form a triplex at the tip of reversed fork. This process would be counteracted by HLTF and WRN helicases or in the presence of CTT LNA or PNA. (E) Reversed fork capped by the triplex can be restored by an alternative, less precise pathway involving DDX11 and SMARCAL1 helicases, which is accompanied by out of register realignment of repetitive DNA strands, ultimately resulting in expansions.

reversed forks (100,118,152), we believe that it counteracts addition of extra repeats upon faithful fork restart (Figure 8C). Occasionally, however, a single-stranded overhang of the reversed nascent leading strand can fold back, forming a YR•R triplex at a tip of the reversed fork (Figure 8D). HLTF and WRN helicases are expected to counteract this development, as they are known to bind to single-stranded 3'-overhangs (135,139,140,153–156). Thus, depletion of both helicases would shift the equilibrium towards the reversed fork capped by a triplex. At the same time, by occluding a single-stranded overhangs CTT LNA-ON and PNA would prevent it from folding back to form DNA triplex.

We hypothesize that if DNA triplex at a tip of the reversed fork is not unraveled by the concerted action of the above helicases, an alternative pathway of fork restoration might take place. Depletion of DDX11 and SMARCAL1 decreases GAA•TTC repeat expansions while leading to a significant accumulation of RIs corresponding to reversed forks (Figures 2D and 4B). Both helicases are known to efficiently unwind various non-B DNA secondary structures, including triplex DNA (96,100,108,109). We propose, therefore, that DDX11 and/or SMARCAL1 helicases might drive an alternative, imprecise pathway of fork restoration, during which complementary strands of the repeat could realign out-of-register, ultimately leading to its expansion (Figure 8D).

Alternatively, having one fork reversed at the GAA•TTC repeat intrinsically bears the risk of re-replicating this repeat by the converging fork, which could ultimately lead to repeat expansions. A possibility for re-replication of a DNA segment within the reversed fork was previously proposed in (104).

How do our data and a model compare to other studies of GAA•TTC repeat expansions in mammalian systems? Progressive accumulation of GAA•TTC repeats in a specific HEK-293T cell line that depended on the MutL γ was observed in (55). In this case, only 1-to-2 repeats were added per cell generation. Small-scale GAA•TTC repeat expan-

sions were also observed during propagation of iPSCs cells, and they were impeded by silencing of the MutS α complex (56,157). In the humanized transgenic FRDA mouse, intergenerational expansions of GAA•TTC repeats were inhibited in the presence of a mismatch repair system (158), while somatic expansions were promoted by the MMR (159). In our case, the *MLH1* gene in the HEK-293T cells is epigenetically silenced (122). Furthermore, we see similar frequencies of repeat expansions in HEK293 cell, where *MLH1* is active (Supplementary Figure S2B). On the surface, these data argue against the involvement of MutL complexes in repeat expansions in our system. Clearly, however, more studies with various MMR-deficient cell lines are needed to elucidate the role of other components the MMR system and to decipher the role of MutL in the large-scale GAA•TTC repeat expansions observed by us.

At the same time, our interpretations of the data are in-line with the study that analyzed replication of GAA•TTC repeats in cells from FRDA patients (45). The authors showed that expanded GAA•TTC repeats in iPSCs derived from FRDA patients caused profound replication fork stalling. Notably, GAA•TTC-specific polyamides that were previously shown to alleviate expansion of the GAA•TTC repeats (56) rescued DNA replication as well (45).

Finally, there are evident similarities between our data and the data on smaller-scale CAG•CTG repeat expansions in mammalian cells implicating fork reversal and restart (76,160,161). It is foreseeable that in our case, initial smaller-size expansions of GAA•TTC repeats occurring during fork reversal and restart become larger during subsequent rounds of plasmid replication.

In summary, we developed a first of a kind, genetically tractable experimental system to study large-scale expansions of FRDA GAA•TTC repeats in cultured human cells. Our candidate gene analysis implicates fork reversal and restoration in the process. We believe that this system could be a valuable tool for elucidating the mechanisms of large-scale expansions in humans and for evaluating the efficiency of perspective FRDA drugs targeting the

instability of GAA•TTC repeats. Finally, we have showed that LNA-modified oligonucleotides and PNA oligomers targeting the GAA•TTC repeats prevent their expansion holding a promise to be developed as future therapeutics for the treatment of FRDA.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Catherine Freudenreich, Mitch McVey, Karlene Cimprich, Alessandro Vindigni and members of the Mirkin laboratory for their important input on this project. We also thank Jessica Wilks for constructing the pJW12 plasmid, Peter Nielsen for sharing with us repeat-specific PNA oligomers and Sanjukta Ghosh for her technical assistance.

FUNDING

National Institute of General Medical Sciences [R35GM130322]; White family (to S.M.M.); Hjärnfonden, The Swedish Research Council, The Swedish Governmental Agency for Innovation Systems, CIMED and Region Stockholm (to R.Z., C.I.E.S.) and NovoNordisk [NNF21OC0072778] 'Pioneer Innovator 2-2021' (to N.M.). Funding for open access charge: NIGMS [R35GM130322]. *Conflict of interest statement.* R.Z. has US and EU patent for diagnosis and treatment of Friedrich's ataxia disease.

REFERENCES

- 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.
- McMurray, C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.*, **11**, 786–799.
- Pearson, C.E. (2003) Slipping while sleeping? Trinucleotide repeat expansions in germ cells. *Trends Mol. Med.*, **9**, 490–495.
- Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalantini, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A. *et al.* (1996) Friedrich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science*, **271**, 1423–1427.
- Pandolfo, M. and Pastore, A. (2009) The pathogenesis of Friedrich ataxia and the structure and function of frataxin. *J. Neurol.*, **256**, 9–17.
- Li, Y., Lu, Y., Polak, U., Lin, K., Shen, J., Farmer, J., Seyer, L., Bhalla, A.D., Rozwadowska, N., Lynch, D.R. *et al.* (2015) Expanded GAA repeats impede transcription elongation through the FXN gene and induce transcriptional silencing that is restricted to the FXN locus. *Hum. Mol. Genet.*, **24**, 6932–6943.
- Herman, D., Jansen, K., Burnett, R., Soragni, E., Perlman, S.L. and Gottesfeld, J.M. (2006) Histone deacetylase inhibitors reverse gene silencing in Friedrich's ataxia. *Nat. Chem. Biol.*, **2**, 551–558.
- Soragni, E., Herman, D., Dent, S.Y., Gottesfeld, J.M., Wells, R.D. and Napierala, M. (2008) Long intronic GAA•TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedrich ataxia. *Nucleic Acids Res.*, **36**, 6056–6065.
- Anzovino, A., Lane, D.J., Huang, M.L. and Richardson, D.R. (2014) Fixing frataxin: 'ironing out' the metabolic defect in Friedrich's ataxia. *Br. J. Pharmacol.*, **171**, 2174–2190.
- Koutnikova, H., Campuzano, V., Foury, F., Dolle, P., Cazzalini, O. and Koenig, M. (1997) Studies of human, mouse and yeast homologues indicate a mitochondrial function for frataxin. *Nat. Genet.*, **16**, 345–351.
- Santoro, L., Perretti, A., Lanzillo, B., Coppola, G., De Joanna, G., Manganello, F., Cocozza, S., De Michele, G., Filla, A. and Caruso, G. (2000) Influence of GAA expansion size and disease duration on central nervous system impairment in Friedrich's ataxia: contribution to the understanding of the pathophysiology of the disease. *Clin. Neurophysiol.*, **111**, 1023–1030.
- Delatycki, M.B. and Bidichandani, S.I. (2019) Friedrich ataxia-pathogenesis and implications for therapies. *Neurobiol. Dis.*, **132**, 104606.
- Campuzano, V., Montermini, L., Lutz, Y., Cova, L., Hindelang, C., Jiralerspong, S., Trottier, Y., Kish, S.J., Faucheux, B., Trouillas, P. *et al.* (1997) Frataxin is reduced in Friedrich ataxia patients and is associated with mitochondrial membranes. *Hum. Mol. Genet.*, **6**, 1771–1780.
- Tai, G., Corben, L.A., Yiu, E.M., Milne, S.C. and Delatycki, M.B. (2018) Progress in the treatment of Friedrich ataxia. *Neurol. Neurochir. Pol.*, **52**, 129–139.
- Evans-Galea, M.V., Pebay, A., Dottori, M., Corben, L.A., Ong, S.H., Lockhart, P.J. and Delatycki, M.B. (2014) Cell and gene therapy for Friedrich ataxia: progress to date. *Hum. Gene Ther.*, **25**, 684–693.
- Pandey, A., Gordon, D.M., Pain, J., Stemmler, T.L., Dancis, A. and Pain, D. (2013) Frataxin directly stimulates mitochondrial cysteine desulfurase by exposing substrate-binding sites, and a mutant Fe-S cluster scaffold protein with frataxin-bypassing ability acts similarly. *J. Biol. Chem.*, **288**, 36773–36786.
- Koeppen, A.H., Becker, A.B., Qian, J. and Feustel, P.J.F.A. (2017) Hypoplasia of spinal cord and dorsal root ganglia. *J. Neuropathol. Exp. Neurol.*, **76**, 101–108.
- Koeppen, A.H. (2011) Friedrich's ataxia: pathology, pathogenesis, and molecular genetics. *J. Neurol. Sci.*, **303**, 1–12.
- Ocana-Santero, G., Diaz-Nido, J. and Herranz-Martín, S. (2021) Future prospects of gene therapy for Friedrich's. *Int. J. Mol. Sci.*, **22**, 1815.
- Punga, T. and Buhler, M. (2010) Long intronic GAA repeats causing Friedrich ataxia impede transcription elongation. *EMBO Mol. Med.*, **2**, 120–129.
- Butler, J.S. and Napierala, M. (2015) Friedrich's ataxia—a case of aberrant transcription termination? *Transcription*, **6**, 33–36.
- Burnett, R., Melander, C., Puckett, J.W., Son, L.S., Wells, R.D., Dervan, P.B. and Gottesfeld, J.M. (2006) DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA•TTC repeats in Friedrich's ataxia. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 11497–11502.
- Al-Mahdawi, S., Pinto, R.M., Ismail, O., Varshney, D., Lymperi, S., Sandi, C., Trabzuni, D. and Pook, M. (2008) The Friedrich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. *Hum. Mol. Genet.*, **17**, 735–746.
- Gacy, A.M., Goellner, G.M., Spiro, C., Chen, X., Gupta, G., Bradbury, E.M., Dyer, R.B., Mikesell, M.J., Yao, J.Z., Johnson, A.J. *et al.* (1998) GAA instability in Friedrich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol. Cell*, **1**, 583–593.
- Potaman, V.N., Oussatcheva, E.A., Lyubchenko, Y.L., Shlyakhtenko, L.S., Bidichandani, S.I., Ashizawa, T. and Sinden, R.R. (2004) Length-dependent structure formation in Friedrich ataxia (GAA)_n(TTC)_n repeats at neutral pH. *Nucleic Acids Res.*, **32**, 1224–1231.
- Bergquist, H., Rocha, C.S., Álvarez-Asencio, R., Nguyen, C.H., Rutland, M.W., Smith, C.I., Good, L., Nielsen, P.E. and Zain, R. (2016) Disruption of higher order DNA structures in Friedrich's Ataxia (GAA)_n repeats by PNA or LNA targeting. *PLoS One*, **11**, e0165788.
- Ohno, M., Fukagawa, T., Lee, J.S. and Ikemura, T. (2002) Triplex-forming dnas in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma*, **111**, 201–213.

28. Bergquist, H. (2009) Structure-specific recognition of Friedreich's ataxia (GAA)_n repeats by benzoquinolinoxaline derivatives. *ChemBioChem*, **10**, 2629–2637.
29. Zain, R. and Sun, J.S. (2003) Do natural DNA triple-helical structures occur and function in vivo? *Cell. Mol. Life Sci.*, **60**, 862–870.
30. Son, L.S., Bacolla, A. and Wells, R.D. (2006) Sticky DNA: in vivo formation in *E. coli* and in vitro association of long GAA*_nTTC tracts to generate two independent supercoiled domains. *J. Mol. Biol.*, **360**, 267–284.
31. Sakamoto, N., Chastain, P.D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J.D. and Wells, R.D. (1999) Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. *Mol. Cell*, **3**, 465–475.
32. Delatycki, M.B., Paris, D., Gardner, R.J., Forshaw, K., Nicholson, G.A., Nassif, N., Williamson, R. and Forrest, S.M. (1998) Sperm DNA analysis in a Friedreich ataxia premutation carrier suggests both meiotic and mitotic expansion in the FRDA gene. *J. Med. Genet.*, **35**, 713–716.
33. Monros, E., Molto, M.D., Martinez, F., Canizares, J., Blanca, J., Vilchez, J.J., Prieto, F., de Frutos, R. and Palau, F. (1997) Phenotype correlation and intergenerational dynamics of the Friedreich ataxia GAA trinucleotide repeat. *Am. J. Hum. Genet.*, **61**, 101–110.
34. De Biase, I., Rasmussen, A., Monticelli, A., Al-Mahdawi, S., Pook, M., Coccozza, S. and Bidichandani, S.I. (2007) Somatic instability of the expanded GAA triplet-repeat sequence in Friedreich ataxia progresses throughout life. *Genomics*, **90**, 1–5.
35. Pollard, L.M., Sharma, R., Gomez, M., Shah, S., Delatycki, M.B., Pianese, L., Monticelli, A., Keats, B.J. and Bidichandani, S.I. (2004) Replication-mediated instability of the GAA triplet repeat mutation in Friedreich ataxia. *Nucleic Acids Res.*, **32**, 5962–5971.
36. Long, A., Napierala, J.S., Polak, U., Hauser, L., Koepfen, A.H., Lynch, D.R. and Napierala, M. (2017) Somatic instability of the expanded GAA repeats in Friedreich's ataxia. *PLoS One*, **12**, e0189990.
37. De Biase, I., Rasmussen, A., Endres, D., Al-Mahdawi, S., Monticelli, A., Coccozza, S., Pook, M. and Bidichandani, S.I. (2007) Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. *Ann. Neurol.*, **61**, 55–60.
38. Zhao, J., Bacolla, A., Wang, G. and Vasquez, K.M. (2010) Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.*, **67**, 43–62.
39. Wang, G. and Vasquez, K.M. (2014) Impact of alternative DNA structures on DNA damage, DNA repair, and genetic instability. *DNA Repair (Amst.)*, **19**, 143–151.
40. Ohshima, K., Sakamoto, N., Labuda, M., Poirier, J., Moseley, M.L., Montermini, L., Ranun, L.P., Wells, R.D. and Pandolfo, M. (1999) A nonpathogenic GAAGGA repeat in the Friedreich gene: implications for pathogenesis. *Neurology*, **53**, 1854–1857.
41. Baran, N., Lapidot, A. and Manor, H. (1991) Formation of DNA triplexes accounts for arrests of DNA synthesis at d(TC)_n and d(GA)_n tracts. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 507–511.
42. 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.
43. Chandok, G.S., Patel, M.P., Mirkin, S.M. and Krasilnikova, M.M. (2012) Effects of Friedreich's ataxia GAA repeats on DNA replication in mammalian cells. *Nucleic Acids Res.*, **40**, 3964–3974.
44. Follonier, C., Oehler, J., Herrador, R. and Lopes, M. (2013) Friedreich's ataxia-associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nat. Struct. Mol. Biol.*, **20**, 486–494.
45. Gerhardt, J., Bhalla, A.D., Butler, J.S., Puckett, J.W., Dervan, P.B., Rosenwaks, Z. and Napierala, M. (2016) Stalled DNA replication forks at the endogenous GAA repeats drive repeat expansion in Friedreich's *Cell Rep.*, **16**, 1218–1227.
46. Shishkin, A.A., Voineagu, I., Matera, R., Cherng, N., Chernet, B.T., Krasilnikova, M.M., Narayanan, V., Lobachev, K.S. and Mirkin, S.M. (2009) Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol. Cell*, **35**, 82–92.
47. Zhang, Y., Shishkin, A.A., Nishida, Y., Marcinkowski-Desmond, D., Saini, N., Volkov, K.V., Mirkin, S.M. and Lobachev, K.S. (2012) Genome-wide screen identifies pathways that govern GAA/TTC repeat fragility and expansions in dividing and nondividing yeast cells. *Mol. Cell*, **48**, 254–265.
48. Shah, K.A., Shishkin, A.A., Voineagu, I., Pavlov, Y.I., Shcherbakova, P.V. and Mirkin, S.M. (2012) Role of DNA polymerases in repeat-mediated genome instability. *Cell Rep.*, **2**, 1088–1095.
49. Tsutakawa, S.E., Thompson, M.J., Arvai, A.S., Neil, A.J., Shaw, S.J., Algasier, S.I., Kim, J.C., Finger, L.D., Jardine, E., Gotham, V.J.B. et al. (2017) Phosphate steering by Flap Endonuclease 1 promotes 5'-flap specificity and incision to prevent genome instability. *Nat. Commun.*, **8**, 15855.
50. Neil, A.J., Kim, J.C. and Mirkin, S.M. (2017) Precarious maintenance of simple DNA repeats in eukaryotes. *Bioessays*, **39**, <https://doi.org/10.1002/bies.201700077>.
51. Neil, A.J., Liang, M.U., Khristich, A.N., Shah, K.A. and Mirkin, S.M. (2018) RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)_n repeats via break-induced replication. *Nucleic Acids Res.*, **46**, 3487–3497.
52. MRindler, P., Clark, R.M., Pollard, L.M., De Biase, I. and Bidichandani, S.I. (2006) Replication in mammalian cells recapitulates the locus-specific differences in somatic instability of genomic GAA triplet-repeats. *Nucleic Acids Res.*, **34**, 6352–6361.
53. Kim, H.M., Narayanan, V., Mieczkowski, P.A., Petes, T.D., Krasilnikova, M.M., Mirkin, S.M. and Lobachev, K.S. (2008) Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. *EMBO J.*, **27**, 2896–2906.
54. Neil, A.J., Hisey, J.A., Quasem, I., McGinty, R.J., Hitzzenko, M., Khristich, A.N. and Mirkin, S.M. (2021) Replication-independent instability of Friedreich's ataxia GAA repeats during chronological aging. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2013080118.
55. Halabi, A., Fuselier, K.T.B. and Grabczyk, E. (2018) GAA*_nTTC repeat expansion in human cells is mediated by mismatch repair complex MutLgamma and depends upon the endonuclease domain in MLH3 isoform one. *Nucleic Acids Res.*, **46**, 4022–4032.
56. Du, J., Campau, E., Soragni, E., Ku, S., Puckett, J.W., Dervan, P.B. and Gottesfeld, J.M. (2012) Role of mismatch repair enzymes in GAA-TTC triplet-repeat expansion in Friedreich ataxia induced pluripotent stem cells. *J. Biol. Chem.*, **287**, 29861–29872.
57. Anjomani, S., Ezzatizadeh, V., Sandi, C., Sandi, M., Al-Mahdawi, S., Chutake, Y. and Pook, M.A. (2015) A novel GAA-repeat-expansion-based mouse model of Friedreich's ataxia. *Dis. Model. Mech.*, **8**, 225–235.
58. Clark, R.M., De Biase, I., Malykhina, A.P., Al-Mahdawi, S., Pook, M. and Bidichandani, S.I. (2007) The GAA triplet-repeat is unstable in the context of the human FXN locus and displays age-dependent expansions in cerebellum and DRG in a transgenic mouse model. *Hum. Genet.*, **120**, 633–640.
59. Claassen, D.A. and Lahue, R.S. (2007) Expansions of CAG/CTG repeats in immortalized human astrocytes. *Hum. Mol. Genet.*, **16**, 3088–3096.
60. Stevens, J.R., Lahue, E.E., Li, G.M. and Lahue, R.S. (2013) Trinucleotide repeat expansions catalyzed by human cell-free extracts. *Cell Res.*, **23**, 565–572.
61. Farrell, B.T. and Lahue, R.S. (2006) CAG*CTG repeat instability in cultured human astrocytes. *Nucleic Acids Res.*, **34**, 4495–4505.
62. Quinet, A., Lemacon, D. and Vindigni, A. (2017) Replication fork reversal: players and guardians. *Mol. Cell*, **68**, 830–833.
63. Myers, A.M., Tzagoloff, A., Kinney, D.M. and Lusty, C.J. (1986) Yeast shuttle and integrative vectors with multiple cloning sites suitable for construction of lacZ fusions. *Gene*, **45**, 299–310.
64. Shah, K.A., McGinty, R.J., Egorova, V.I. and Mirkin, S.M. (2014) Coupling transcriptional state to large-scale repeat expansions in yeast. *Cell Rep.*, **9**, 1594–1602.
65. Nielsen, P.E., Egholm, M., Berg, R.H. and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **254**, 1497–1500.
66. Gietz, R.D. (1993) In: *Molecular Genetics of Yeast*. Oxford University Press, pp. 121–134.
67. Samadashwily, G.M., Dayn, A. and Mirkin, S.M. (1993) Suicidal nucleotide sequences for DNA polymerization. *EMBO J.*, **12**, 4975–4983.
68. Krasilnikov, A.S., Panyutin, I.G., Samadashwily, G.M., Cox, R., Lazurkin, Y.S. and Mirkin, S.M. (1997) Mechanisms of triplex-caused polymerization arrest. *Nucleic Acids Res.*, **25**, 1339–1346.

69. Ohshima, K., Montermini, L., Wells, R.D. and Pandolfo, M. (1998) Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication in vivo. *J. Biol. Chem.*, **273**, 14588–14595.
70. Radchenko, E.A., McGinty, R.J., Aksenova, A.Y., Neil, A.J. and Mirkin, S.M. (2018) Quantitative analysis of the rates for repeat-mediated genome instability in a yeast experimental system. *Methods Mol. Biol.*, **1672**, 421–438.
71. Brown, R.E. and Freudenreich, C.H. (2020) Structure-forming repeats and their impact on genome stability. *Curr. Opin. Genet. Dev.*, **67**, 41–51.
72. Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932–940.
73. 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.
74. Moe, S.E., Sorbo, J.G. and Holen, T. (2008) Huntingtin triplet-repeat locus is stable under long-term Fen1 knockdown in human cells. *J. Neurosci. Methods*, **171**, 233–238.
75. Spiro, C. and McMurray, C.T. (2003) Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. *Mol. Cell. Biol.*, **23**, 6063–6074.
76. Liu, G., Chen, X., Gao, Y., Lewis, T., Barthelemy, J. and Leffak, M. (2012) Altered replication in human cells promotes DMPK (CTG)_n·(CAG)_n repeat instability. *Mol. Cell. Biol.*, **32**, 1618–1632.
77. Gellon, L., Kaushal, S., Cebrian, J., Lahiri, M., Mirkin, S.M. and Freudenreich, C.H. (2019) Mre1 and Tof1 prevent fragility and instability at long CAG repeats by their fork stabilizing function. *Nucleic Acids Res.*, **47**, 794–805.
78. Entezam, A. and Usdin, K. (2009) ATM and ATR protect the genome against two different types of tandem repeat instability in Fragile X premutation mice. *Nucleic Acids Res.*, **37**, 6371–6377.
79. Krejci, L., Altmannova, V., Spirek, M. and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic Acids Res.*, **40**, 5795–5818.
80. Prakash, R., Zhang, Y., Feng, W. and Jasin, M. (2015) Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb. Perspect. Biol.*, **7**, a016600.
81. Polleys, E.J. and Freudenreich, C.H. (2021) Homologous recombination within repetitive DNA. *Curr. Opin. Genet. Dev.*, **71**, 143–153.
82. Sugiyama, T., New, J.H. and Kowalczykowski, S.C. (1998) DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6049–6054.
83. Sung, P. (1997) Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *J. Biol. Chem.*, **272**, 28194–28197.
84. Feng, Z., Scott, S.P., Bussen, W., Sharma, G.G., Guo, G., Pandita, T.K. and Powell, S.N. (2011) Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 686–691.
85. Anand, R.P., Lovett, S.T. and Haber, J.E. (2013) Break-induced DNA replication. *Cold Spring Harb. Perspect. Biol.*, **5**, a010397.
86. Malkova, A. and Ira, G. (2013) Break-induced replication: functions and molecular mechanism. *Curr. Opin. Genet. Dev.*, **23**, 271–279.
87. 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.
88. Kim, J.C., Harris, S.T., Dinter, T., Shah, K.A. and Mirkin, S.M. (2017) The role of break-induced replication in large-scale expansions of (CAG)_n/(CTG)_n repeats. *Nat. Struct. Mol. Biol.*, **24**, 55–60.
89. 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.
90. Malacaria, E., Pugliese, G.M., Honda, M., Marabitti, V., Aiello, F.A., Spies, M., Franchitto, A. and Pichierri, P. (2019) Rad52 prevents excessive replication fork reversal and protects from nascent strand degradation. *Nat. Commun.*, **10**, 1412.
91. Motegi, A., Liaw, H.J., Lee, K.Y., Roest, H.P., Maas, A., Wu, X., Moinova, H., Markowitz, S.D., Ding, H., Hoeijmakers, J.H. et al. (2008) Polyubiquitination of proliferating cell nuclear antigen by HLTf and SHPRH prevents genomic instability from stalled replication forks. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 12411–12416.
92. Lin, J.R., Zeman, M.K., Chen, J.Y., Yee, M.C. and Cimprich, K.A. (2011) SHPRH and HLTf act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. *Mol. Cell*, **42**, 237–249.
93. Cherng, N., Shishkin, A.A., Schlager, L.I., Tuck, R.H., Sloan, L., Matera, R., Sarkar, P.S., Ashizawa, T., Freudenreich, C.H. and Mirkin, S.M. (2011) Expansions, contractions, and fragility of the spinocerebellar ataxia type 10 pentanucleotide repeat in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 2843–2848.
94. Bai, G., Kermi, C., Stoy, H., Schiltz, C.J., Bacal, J., Zaino, A.M., Hadden, M.K., Eichman, B.F., Lopes, M. and Cimprich, K.A. (2020) HLTf promotes fork reversal, limiting replication stress resistance and preventing multiple mechanisms of unrestrained DNA synthesis. *Mol. Cell*, **78**, 1237–1251.
95. Frizzell, A., Nguyen, J.H., Petalcorin, M.I., Turner, K.D., Boulton, S.J., Freudenreich, C.H. and Lahue, R.S. (2014) RTEL1 inhibits trinucleotide repeat expansions and fragility. *Cell Rep.*, **6**, 827–835.
96. Ciccia, A., Bredemeyer, A.L., Sowa, M.E., Terret, M.E., Jallepalli, P.V., Harper, J.W. and Elledge, S.J. (2009) The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev.*, **23**, 2415–2425.
97. Betous, R., Mason, A.C., Rambo, R.P., Bansbach, C.E., Badu-Nkansah, A., Sirbu, B.M., Eichman, B.F. and Cortez, D. (2012) SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes Dev.*, **26**, 151–162.
98. Ciccia, A., Nimonkar, A.V., Hu, Y., Hajdu, I., Achar, Y.J., Izhar, L., Petit, S.A., Adamson, B., Yoon, J.C., Kowalczykowski, S.C. et al. (2012) Polyubiquitinated PCNA recruits the ZRANB3 translocase to maintain genomic integrity after replication stress. *Mol. Cell*, **47**, 396–409.
99. Bhat, K.P., Betous, R. and Cortez, D. (2015) High-affinity DNA-binding domains of replication protein A (RPA) direct SMARCAL1-dependent replication fork remodeling. *J. Biol. Chem.*, **290**, 4110–4117.
100. Qiu, S., Jiang, G., Cao, L. and Huang, J. (2021) Replication fork reversal and protection. *Front. Cell Dev. Biol.*, **9**, 670392.
101. Yuan, J., Ghosal, G. and Chen, J. (2009) The annealing helicase HARP protects stalled replication forks. *Genes Dev.*, **23**, 2394–2399.
102. Bansbach, C.E., Boerkoel, C.F. and Cortez, D. (2010) SMARCAL1 and replication stress: an explanation for SIOD? *Nucleus*, **1**, 245–248.
103. Dugrawala, H., Bhat, K.P., Le Meur, R., Chazin, W.J., Ding, X., Sharan, S.K., Wessel, S.R., Sathe, A.A., Zhao, R. and Cortez, D. (2017) RADX promotes genome stability and modulates chemosensitivity by regulating RAD51 at replication forks. *Mol. Cell*, **67**, 374–386.
104. Neelsen, K.J. and Lopes, M. (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nat. Rev. Mol. Cell Biol.*, **16**, 207–220.
105. Polleys, E.J., House, N.C.M. and Freudenreich, C.H. (2017) Role of recombination and replication fork restart in repeat instability. *DNA Repair (Amst.)*, **56**, 156–165.
106. Lerner, L.K. and Sale, J.E. (2019) Replication of G quadruplex DNA. *Genes*, **10**, 95.
107. Santos, D., Mahtab, M., Boavida, A. and Pisani, F.M. (2021) Role of the DDX11 DNA helicase in Warsaw breakage syndrome etiology. *Int. J. Mol. Sci.*, **22**, 2308.
108. Simon, A.K., Kummer, S., Wild, S., Lezaja, A., Teloni, F., Jozwiakowski, S.K., Altmeyer, M. and Gari, K. (2020) The iron-sulfur helicase DDX11 promotes the generation of single-stranded DNA for CHK1 activation. *Life Sci Alliance*, **3**, e201900547.
109. Jegadesan, N.K. and Branzei, D. (2021) DDX11 loss causes replication stress and pharmacologically exploitable DNA repair defects. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2024258118.
110. Hirota, Y. and Lahti, J.M. (2000) Characterization of the enzymatic activity of hChIR1, a novel human DNA helicase. *Nucleic Acids Res.*, **28**, 917–924.
111. Guo, M., Hundseth, K., Ding, H., Vidhyasagar, V., Inoue, A., Nguyen, C.H., Zain, R., Lee, J.S. and Wu, Y. (2015) A distinct triplex

- DNA unwinding activity of ChlR1 helicase. *J. Biol. Chem.*, **290**, 5174–5189.
112. Murray, J.M., Lindsay, H.D., Munday, C.A. and Carr, A.M. (1997) Role of schizosaccharomyces pombe RecQ homolog, recombination, and checkpoint genes in UV damage tolerance. *Mol. Cell. Biol.*, **17**, 6868–6875.
 113. Watt, P.M., Louis, E.J., Borts, R.H. and Hickson, I.D. (1995) Sgs1: a eukaryotic homolog of E. coli RecQ that interacts with topoisomerase II in vivo and is required for faithful chromosome segregation. *Cell*, **81**, 253–260.
 114. Lebel, M., Spillare, E.A., Harris, C.C. and Leder, P. (1999) The Werner syndrome gene product co-purifies with the DNA replication complex and interacts with PCNA and topoisomerase I. *J. Biol. Chem.*, **274**, 37795–37799.
 115. Brosh, R.M. Jr, Orren, D.K., Nehlin, J.O., Ravn, P.H., Kenny, M.K., Machwe, A. and Bohr, V.A. (1999) Functional and physical interaction between WRN helicase and human replication protein A. *J. Biol. Chem.*, **274**, 18341–18350.
 116. Gangloff, S., McDonald, J.P., Bendixen, C., Arthur, L. and Rothstein, R. (1994) The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol. Cell. Biol.*, **14**, 8391–8398.
 117. Kamath-Loeb, A.S., Johansson, E., Burgers, P.M. and Loeb, L.A. (2000) Functional interaction between the Werner Syndrome protein and DNA polymerase delta. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 4603–4608.
 118. Berti, M., Ray Chaudhuri, A., Thangavel, S., Gomathinayagam, S., Kenig, S., Vujanovic, M., Odreman, F., Glatter, T., Graziano, S., Mendoza-Maldonado, R. et al. (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat. Struct. Mol. Biol.*, **20**, 347–354.
 119. Orren, D.K., Theodore, S. and Machwe, A. (2002) The Werner syndrome helicase/exonuclease (WRN) disrupts and degrades D-loops in vitro. *Biochemistry*, **41**, 13483–13488.
 120. Mohaghegh, P., Karow, J.K., Brosh, R.M. Jr, Bohr, V.A. and Hickson, I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.*, **29**, 2843–2849.
 121. Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P. and Kowalczykowski, S.C. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.*, **25**, 350–362.
 122. Trojan, J., Zeuzem, S., Randolph, A., Hemmerle, C., Brieger, A., Raedle, J., Plotz, G., Jiricny, J. and Marra, G. (2002) Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology*, **122**, 211–219.
 123. Viguera, E., Rodríguez, A., Hernández, P., Krimer, D.B., Trellez, O. and Schwartzman, J.B. (1998) A computer model for the analysis of DNA replication intermediates by two-dimensional agarose gel electrophoresis. *Gene*, **217**, 41–49.
 124. Amunugama, R., Willcox, S., Wu, R.A., Abdullah, U.B., El-Sagheer, A.H., Brown, T., McHugh, P.J., Griffith, J.D. and Walter, J.C. (2018) Replication fork reversal during DNA interstrand crosslink repair requires CMG unloading. *Cell Rep.*, **23**, 3419–3428.
 125. Zaid, A., Sun, J.S., Nguyen, C.H., Bisagni, E., Garestier, T., Grierson, D.S. and Zain, R. (2004) Triple-helix directed cleavage of double-stranded DNA by benzoquinoxaline-1,10-phenanthroline conjugates. *ChemBioChem*, **5**, 1550–1557.
 126. Bergquist, H., Nikravesh, A., Fernandez, R.D., Larsson, V., Nguyen, C.H., Good, L. and Zain, R. (2009) Structure-specific recognition of Friedreich's ataxia (GAA)_n repeats by benzoquinoxaline derivatives. *ChemBioChem*, **10**, 2629–2637.
 127. Demidov, V.V., Yavnilovich, M.V., Belotserkovskii, B.P., Frank-Kamenetskii, M.D. and Nielsen, P.E. (1995) Kinetics and mechanism of polyamide ("peptide") nucleic acid binding to duplex DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 2637–2641.
 128. Escudé, C., Nguyen, C.H., Kukreti, S., Janin, Y., Sun, J.S., Bisagni, E., Garestier, T. and Hélène, C. (1998) Rational design of a triple helix-specific intercalating ligand. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 3591–3596.
 129. Zain, R., Marchand, C., Sun, J., Nguyen, C.H., Bisagni, E., Garestier, T. and Helene, C. (1999) Design of a triple-helix-specific cleaving reagent. *Chem. Biol.*, **6**, 771–777.
 130. Amiri, H., Nekhotiaeva, N., Sun, J.S., Nguyen, C.H., Grierson, D.S., Good, L. and Zain, R. (2005) Benzoquinoxaline derivatives stabilize and cleave H-DNA and repress transcription downstream of a triplex-forming sequence. *J. Mol. Biol.*, **351**, 776–783.
 131. Gomes-Pereira, M., Foiry, L., Nicole, A., Huguet, A., Junien, C., Munnich, A. and Gourdon, G. (2007) CTG trinucleotide repeat "big jumps": large expansions, small mice. *PLoS Genet.*, **3**, e52.
 132. Pike, A.C., Gomathinayagam, S., Swuec, P., Berti, M., Zhang, Y., Schnecke, C., Marino, F., von Delft, F., Renault, L., Costa, A. et al. (2015) Human RECQ1 helicase-driven DNA unwinding, annealing, and branch migration: insights from DNA complex structures. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 4286–4291.
 133. Van, C., Yan, S., Michael, W.M., Waga, S. and Cimprich, K.A. (2010) Continued primer synthesis at stalled replication forks contributes to checkpoint activation. *J. Cell Biol.*, **189**, 233–246.
 134. Poole, L.A. and Cortez, D. (2017) Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability. *Crit. Rev. Biochem. Mol. Biol.*, **52**, 696–714.
 135. Shen, J.C., Gray, M.D., Oshima, J. and Loeb, L.A. (1998) Characterization of Werner syndrome protein DNA helicase activity: directionality, substrate dependence and stimulation by replication protein A. *Nucleic Acids Res.*, **26**, 2879–2885.
 136. Huang, S., Beresten, S., Li, B., Oshima, J., Ellis, N.A. and Campisi, J. (2000) Characterization of the human and mouse WRN 3'→5' exonuclease. *Nucleic Acids Res.*, **28**, 2396–2405.
 137. Suzuki, N., Shimamoto, A., Imamura, O., Kuromitsu, J., Kitao, S., Goto, M. and Furuichi, Y. (1997) DNA helicase activity in Werner's syndrome gene product synthesized in a baculovirus system. *Nucleic Acids Res.*, **25**, 2973–2978.
 138. Machwe, A., Xiao, L., Groden, J., Matson, S.W. and Orren, D.K. (2005) RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. *J. Biol. Chem.*, **280**, 23397–23407.
 139. Gray, M.D., Shen, J.C., Kamath-Loeb, A.S., Blank, A., Sopher, B.L., Martin, G.M., Oshima, J. and Loeb, L.A. (1997) The Werner syndrome protein is a DNA helicase. *Nat. Genet.*, **17**, 100–103.
 140. Shen, J.C., Gray, M.D., Oshima, J., Kamath-Loeb, A.S., Fry, M. and Loeb, L.A. (1998) Werner syndrome protein. I. DNA helicase and dna exonuclease reside on the same polypeptide. *J. Biol. Chem.*, **273**, 34139–34144.
 141. Bansbach, C.E., Betous, R., Lovejoy, C.A., Glick, G.G. and Cortez, D. (2009) The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev.*, **23**, 2405–2414.
 142. Masnovi, C., Lobo, A.F. and Mirkin, S.M. (2022) Replication dependent and independent mechanisms of GAA repeat instability. *DNA Repair (Amst.)*, **118**, 103385.
 143. Voloshin, O.N., Mirkin, S.M., Lyamichev, V.I., Belotserkovskii, B.P. and Frank-Kamenetskii, M.D. (1988) Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA. *Nature*, **333**, 475–476.
 144. Grabczyk, E. and Usdin, K. (2000) The GAA*_nTTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner. *Nucleic Acids Res.*, **28**, 2815–2822.
 145. Zhang, J., Fakharzadeh, A., Pan, F., Roland, C. and Sagui, C. (2020) Atypical structures of GAA/TTC trinucleotide repeats underlying Friedreich's ataxia: DNA triplexes and RNA/DNA hybrids. *Nucleic Acids Res.*, **48**, 9899–9917.
 146. Stillman, B. (2015) Reconsidering DNA polymerases at the replication fork in eukaryotes. *Mol. Cell*, **59**, 139–141.
 147. Fortune, J.M., Pavlov, Y.I., Welch, C.M., Johansson, E., Burgers, P.M. and Kunkel, T.A. (2005) Saccharomyces cerevisiae DNA polymerase delta: high fidelity for base substitutions but lower fidelity for single- and multi-base deletions. *J. Biol. Chem.*, **280**, 29980–29987.
 148. Borowiec, J.A., Dean, F.B., Bullock, P.A. and Hurwitz, J. (1990) Binding and unwinding—how T antigen engages the SV40 origin of DNA replication. *Cell*, **60**, 181–184.
 149. Hu, L., Filippakis, H., Huang, H., Yen, T.J. and Gjoerup, O.V. (2013) Replication stress and mitotic dysfunction in cells expressing simian virus 40 large T antigen. *J. Virol.*, **87**, 13179–13192.

150. Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A. and Lopes, M. (2015) Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *J. Cell Biol.*, **208**, 563–579.
151. Hashimoto, Y., Ray Chaudhuri, A., Lopes, M. and Costanzo, V. (2010) Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. *Nat. Struct. Mol. Biol.*, **17**, 1305–1311.
152. Kitao, S., Ohsugi, I., Ichikawa, K., Goto, M., Furuichi, Y. and Shimamoto, A. (1998) Cloning of two new human helicase genes of the RecQ family: biological significance of multiple species in higher eukaryotes. *Genomics*, **54**, 443–452.
153. Chavez, D.A., Greer, B.H. and Eichman, B.F. (2018) The HIRAN domain of helicase-like transcription factor positions the DNA translocase motor to drive efficient DNA fork regression. *J. Biol. Chem.*, **293**, 8484–8494.
154. Kile, A.C., Chavez, D.A., Bacal, J., Eldirany, S., Korzhnev, D.M., Bezsonova, I., Eichman, B.F. and Cimprich, K.A. (2015) HLTf's ancient HIRAN domain binds 3' DNA ends to drive replication fork reversal. *Mol. Cell*, **58**, 1090–1100.
155. Shen, J.C., Lao, Y., Kamath-Loeb, A., Wold, M.S. and Loeb, L.A. (2003) The N-terminal domain of the large subunit of human replication protein A binds to Werner syndrome protein and stimulates helicase activity. *Mech. Ageing Dev.*, **124**, 921–930.
156. Dusek, C.O., Dash, R.C., McPherson, K.S., Calhoun, J.T., Bezsonova, I., Korzhnev, D.M. and Hadden, M.K. (2022) DNA sequence specificity reveals a role of the HLTf HIRAN domain in the recognition of trinucleotide repeats. *Biochemistry*, **61**, 992–1004.
157. Ku, S., Soragni, E., Campau, E., Thomas, E.A., Altun, G., Laurent, L.C., Loring, J.F., Napierala, M. and Gottesfeld, J.M. (2010) Friedreich's ataxia induced pluripotent stem cells model intergenerational GAA-TTC triplet repeat instability. *Cell Stem Cell*, **7**, 631–637.
158. Ezzatizadeh, V., Pinto, R.M., Sandi, C., Sandi, M., Al-Mahdawi, S., Te Riele, H. and Pook, M.A. (2012) The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model. *Neurobiol. Dis.*, **46**, 165–171.
159. Bourn, R.L., De Biase, I., Pinto, R.M., Sandi, C., Al-Mahdawi, S., Pook, M.A. and Bidichandani, S.I. (2012) Pms2 suppresses large expansions of the (GAA.TTC)_n sequence in neuronal tissues. *PLoS One*, **7**, e47085.
160. 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.
161. 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.