Human Molecular Genetics, 2014, Vol. 23, No. 11 **2940–2952**
doi:10.1093/hmg/ddu006
Advance Access published on January 12, 2014

Chromosome fragility and the abnormal replication of the *FMR1* locus in fragile X syndrome

Dmitry Yudkin[1,3], Bruce E. Hayward[1], Mirit I. Aladjem[2], Daman Kumari[1] and Karen Usdin[1,*]

[1]Section on Gene Structure and Disease, Laboratory of Cell and Molecular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, and [2]DNA Replication Group, Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA and [3]Department of Genomic Diversity and Evolution, Institute of Molecular and Cellular Biology SB RAS, Novosibirsk 630090, Russia

Received November 15, 2013; Revised December 24, 2013; Accepted January 8, 2014

Fragile X Syndrome (FXS) is a learning disability seen in individuals who have >200 CGG•CCG repeats in the 5′ untranslated region of the X-linked *FMR1* gene. Such alleles are associated with a fragile site, FRAXA, a gap or constriction in the chromosome that is coincident with the repeat and is induced by folate stress or thymidylate synthase inhibitors like fluorodeoxyuridine (FdU). The molecular basis of the chromosome fragility is unknown. Previous work has suggested that the stable intrastrand structures formed by the repeat may be responsible, perhaps via their ability to block DNA synthesis. We have examined the replication dynamics of normal and FXS cells with and without FdU. We show here that an intrinsic problem with DNA replication exists in the *FMR1* gene of individuals with FXS even in the absence of FdU. Our data suggest a model for chromosome fragility in FXS in which the repeat impairs replication from an origin of replication (ORI) immediately adjacent to the repeat. The fact that the replication problem occurs even in the absence of FdU suggests that this phenomenon may have *in vivo* consequences, including perhaps accounting for the loss of the X chromosome containing the fragile site that causes Turner syndrome (45, X0) in female carriers of such alleles. Our data on FRAXA may also be germane for the other FdU-inducible fragile sites in humans, that we show here share many common features with FRAXA.

INTRODUCTION

Fragile sites are a common feature of mammalian genomes. These sites are apparent as gaps, constrictions or breaks in the chromosome that are visible in metaphase spreads prepared from cells grown in the presence of agents like fluorodeoxyuridine (FdU), distamycin, bromodeoxyuridine or aphidicolin (APH) (see (1,2) for comprehensive reviews). Fragile sites are also frequent sites of chromosome breakage and translocation *in vivo*, suggesting that chromosome fragility is not an *in vitro* artifact, but a feature of normally growing cells. While different agents induce fragility at different chromosomal loci, all of these agents interfere with DNA replication in some way, suggesting that the different fragile sites represent sequences that are either difficult to replicate or replicate late in S phase. Cells treated with such agents could then enter mitosis before replication of the fragile site region has been completed. This would result in premature chromatin condensation and the microscopic appearance of the fragile site.

Many of the common fragile sites such as those induced by APH, a DNA polymerase α inhibitor, are spread over megabases of DNA with a heterogeneous composition (1). While these regions are often enriched for A + T-rich sequences and sequences with high flexibility and low stability (3), as yet no specific sequence responsible for slowing DNA replication has been identified in these regions. Camptothecin (CPT), a topoisomerase I (Topo I) poison, reduces APH-induced chromosome fragility (4). This has led to the suggestion that APH acts by promoting the uncoupling of the DNA polymerase from the helicase/topoisomerase complex thereby increasing the opportunity for secondary structures to form between the polymerase and the helicase on the exposed template. These structures can then cause replication fork stalling, which in turn results in fragile site production (4). It has also been suggested that common fragile sites arise from replication stalling at an ORI in an ORI-poor region under conditions of replication stress (5). Fragile sites are generated because no additional ORIs are

*To whom correspondence should be addressed at: Building 8, Room 2A19, National Institutes of Health, 8 Center Drive MSC 0830, Bethesda, MD 20892-0830, USA. Tel: +1 301 496 2189; Fax: +1 301 402 0053; Email: ku@helix.nih.gov

available that can be activated to complete replication of the region. On the other hand, there are also data to suggest that structural impediments to DNA synthesis may not be necessary for fragility, and that regions with few ORIs may be fragile simply due to their relative lack of origins (6) or because ORIs are utilized that are less efficient than ORIs elsewhere in the genome (7).

In addition to the common fragile sites, there are a number of rare fragile sites that are only seen in a subset of the human population. A subset of these fragile sites are induced by folate stress or FdU. Unlike the common APH-inducible fragile sites, FdU-inducible fragile sites all map to much smaller genomic regions that contain long CGG/CCG-repeat tracts that are often methylated. Shorter repeats at the same location are not fragile, suggesting that fragility in these cases cannot simply be the result of their chromosomal context. The best known of this class of fragile site is FRAXA, the fragile site associated with Fragile X syndrome. FXS is an intellectual disability resulting from the inheritance of >200 CGG/CCG-repeats in the 5′ UTR (8,9) of the X-linked gene *FMR1*. Such alleles are referred to as full mutations (FMs) and are usually associated with aberrant DNA hypermethylation of the promoter and transcriptional silencing of the gene. Female carriers of such alleles are at high risk of Turner syndrome in which the X chromosome carrying the FM has been lost (10). Loss of the FM chromosome may be an *in vivo* consequence of the expression of FRAXA, resulting perhaps from fusion of two broken sister chromatids at the fragile site to generate a dicentric chromatid, followed by its loss from one cell during cytokinesis. Chromosome fragility is seen at much lower frequencies in carriers of *FMR1* alleles that are either normal (<55 repeats) or that have 55–200 repeats (Premutation, PM alleles). In fact, prior to the development of a molecular assay for FXS, chromosome fragility at this locus was a key diagnostic feature of FXS and it is for this site, that the disorder is named.

FdU acts by inhibiting thymidylate synthase (TS). This leads to nucleotide pool imbalances and the slowing of replication (11). FdU can also become incorporated into DNA leading to the production of FdU:A and, occasionally, FdU:G mispairs (12). However, the molecular basis of the FdU-inducible fragile sites is unknown. Long CGG/CCG-repeats are known to form a variety of non-canonical structures including hairpins, *i*-tetraplexes, G-tetraplexes/quadruplexes and Z-DNA (13–19) some of which have been shown to be very effective at blocking DNA synthesis *in vitro* (16). These sequences are also associated with blocks to DNA synthesis *in vivo* (20) and methylation stabilizes some of these structures (18,21). Chromosome fragility may result if conditions arise that favor the formation of secondary structures by the repeat that disrupt replication through the repeat-containing region. FdU, because it primarily affects pyrimidine pools, may cause replication of the purine-rich strand to proceed more slowly than the pyrimidine-rich strand, possibly resulting in uncoupling of leading and lagging strand synthesis as has been proposed for the formation of some common fragile sites (4). This uncoupling could create conditions that favor the formation of secondary structures by the repeat that then blocks the replication fork. The effect of FdU may be exacerbated by the known tendency of transcriptionally inactive alleles to show delayed replication (22) and specifically for FM alleles to replicate very late in the cell cycle (23,24).

To better understand the basis of chromosome fragility in FXS, we have examined the expression of the FRAXA fragile site in the presence of both FdU and CPT. We have also ascertained the replication profiles of normal and FXS syndrome alleles in the presence and absence of FdU. Our data show that there is an intrinsic problem with the replication of FM alleles that is seen even under normal growth conditions and that this problem is not specifically exacerbated by FdU treatment. While CPT does reduce chromosome fragility, the replication problem is likely to be unrelated to uncoupling of the polymerase from the helicase/topoisomerase complexes as some have suggested for APH-inducible sites. Rather, we suggest that FRAXA fragility arises because FdU slows replication globally thus making this region that contains an impediment to DNA synthesis and already replicates late, finish replication even later.

RESULTS

CPT reduces FRAXA chromosome fragility

It has been suggested that CPT, a Topo I poison, reduces the fragility of some of the common APH-inducible fragile sites by reducing the opportunity for a secondary structure to form that could cause stalling of the replication fork (4). To test the effect of CPT on the FdU-inducible FRAXA fragile site, we examined the expression of the fragile site in normal and patient lymphoblastoid cells (see Table 1) using combinations of FdU and CPT. In the presence of 0.1 μM FdU but no CPT, cells from two males with normal *FMR1* alleles, GM06865 and GM06895, showed very low levels of chromosome fragility (≤1%; data not shown). A cell line carrying an allele that we have shown to be an unmethylated FM (UFM), GM06897, expressed the fragile site in 6% of cells. However, this incidence was not statistically different from that seen in cells with normal *FMR1* alleles (data not shown). In contrast, two cell lines containing a methylated FM allele, GM03200 and GM04025, expressed the FRAXA fragile site in 16–20% of their cells (Fig. 1). Addition of 3 nM CPT to cultures containing FdU decreased expression of the FX fragile site in these cell lines by 65 and 46%, respectively ($P = 0.03$ by Fisher's exact test). Addition of 30 nM CPT decreased expression even further ($P < 0.05$). A decline in fragile site expression was also seen in the case of the UFM carrier, although this did not reach statistical significance (data not shown).

CPT reduces the mitotic index of FdU-treated cells

In principle, CPT could reduce FRAXA expression either by reducing the uncoupling of the polymerase/helicase complex or by

Table 1. Cell lines used for analysis of the *FMR1* locus in Fragile X pedigrees

Cell line	Allele	Methylation	Repeat number
GM06865	Normal	No	<30
GM06895	Normal	No	<30
GM03200	FM	Yes	~530
GM04025	FM	Yes	~500
GM06897	UFM	No	~477
GM06891	Premutation	No	112
GM06891E	Premutation	No	183

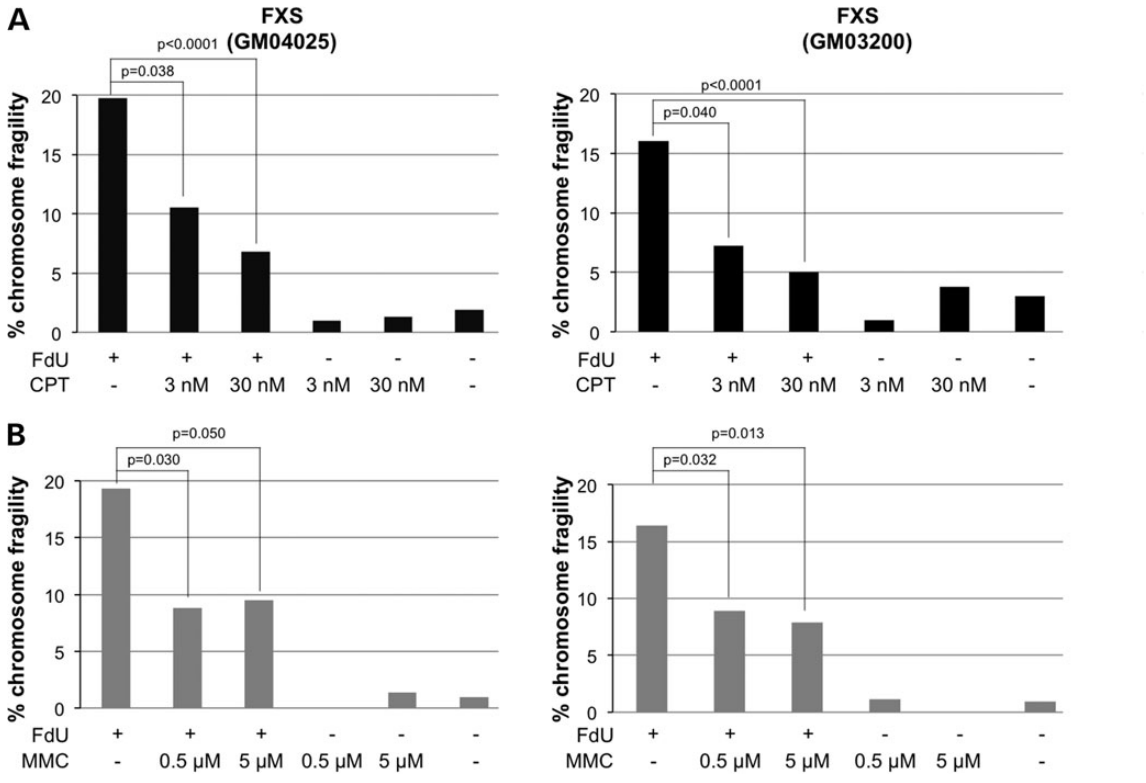

Figure 1. Effect of CPT and MMC on expression of the FRAXA fragile site. FM cell lines, GM03200 and GM04025, were treated for 18 h with FdU with and without CPT or MMC at the concentrations indicated. Cells were then treated with colcemid, fixed on slides and subjected to *in situ* hybridization with a BAC probe containing *FMR1* and the chromosomes stained with DAPI as described in the Materials and Methods. The frequency of breaks at FRAXA was then scored. Similar results were seen for samples treated with CPT or MMC for 4 h (data not shown). CPT/MMC alone had no significant effect on chromosome fragility at FRAXA and were significantly different from the FdU-treated samples at $P < 0.05$ (Fisher's exact test).

Table 2. Effect of CPT and MMC on the mitotic index (MI) of GM04025 (FM) cells

	MI	*P*-value*
FdU	0.54	
FdU + CPT (3 nM)	0.25	0.12
FdU + CPT (30 nM)	0.13	0.01
CPT (3 nM)	0.85	0.21
CPT (30 nM)	0.20	0.04
DMSO	0.81	0.30
FdU	0.60	
FdU + MMC (0.5 μM)	0.27	0.07
FdU + MMC (5 μM)	0.19	0.02
MMC (0.5 μM)	0.21	0.03
MMC (5 μM)	0.12	0.01

*P-value of the difference between FdU alone and FdU with CPT/MMC by Fisher's test.

virtue of its ability to trigger cell cycle checkpoints that would delay entry into mitosis giving cells more time to complete replication of FXS alleles. In previous experiments with common fragile sites in lymphoblastoid cells, CPT did not appear to affect the cell cycle (4). To identify any potential effect of CPT on checkpoint induction, we examined the effect of CPT on the mitotic index, a measure of the number of cells that have successfully entered mitosis. Treatment of a FM patient cell line, GM04025, with 0.1 μM FdU had no significant effect on the number of mitotic figures or cells undergoing cytokinesis ($P = 0.29$). The addition of CPT reduced the number of mitotic figures although this effect was only significant for the higher CPT dose (30 nM) (Table 2).

We also examined the effect of mitomycin C (MMC), a compound that causes DNA damage by generating DNA crosslinks (25) but that does not target Topo I. MMC reduced both FdU-inducible fragility at FRAXA (Fig. 1B) and the number of cells entering mitosis (Table 2). Thus, it may be that 30 nM CPT and both concentrations of MMC tested reduce fragility by causing DNA damage that induces cell cycle checkpoints. This would delay the exit from S phase thus giving cells more time to complete replication before mitosis and chromosome condensation begins. The fact that 3 nM CPT shows a non-significant trend toward producing fewer cells that enter mitosis, might indicate that the replication delay may not have to be very long-lived in order to provide sufficient time for cells to complete replication of the FX allele.

ORI activity spanning the FX repeat is seen in a variety of cell types

The ability of DNA damaging agents to reduce fragility would be consistent with the idea that fragility arises because the long CGG-repeat tract blocks replication fork progression. Activation of the DNA damage-induced cell cycle checkpoints increases the time that cells have to complete replication of

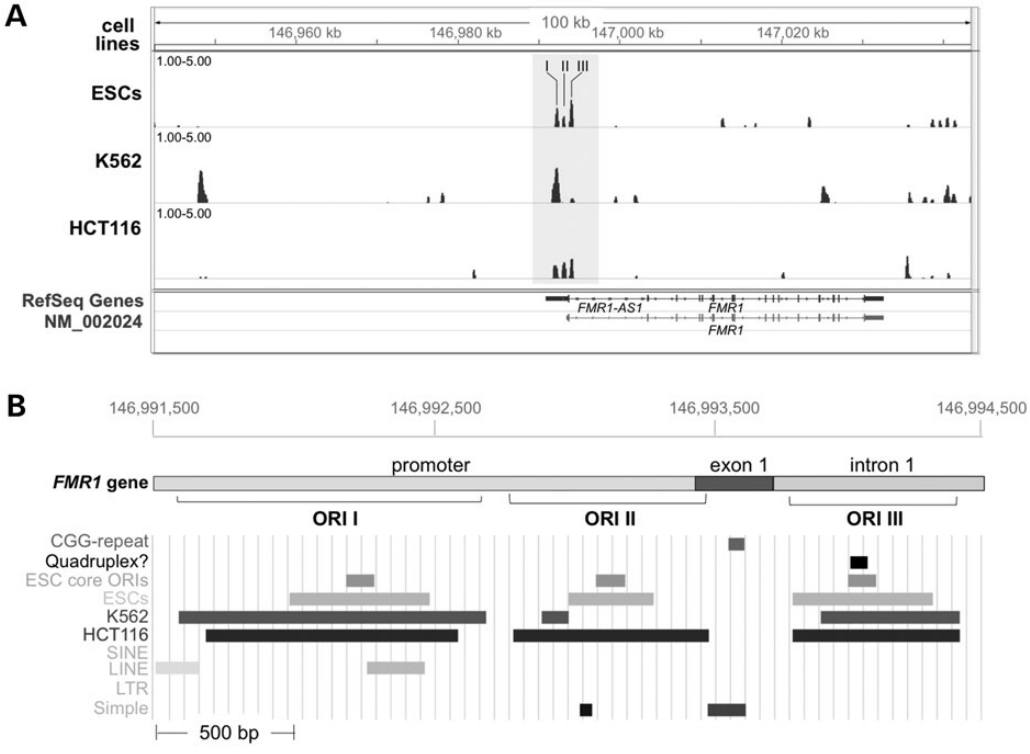

Figure 2. Sites of replication initiation in the vicinity of the *FMR1* gene in a panel of human cell lines. Panel (**A**) shows the nascent strand abundance in the 100 kb region containing *FMR1* in ESCs, K562 and HCT116 cells as determined by deep sequencing of nascent strands. The nascent strand peaks at the 5′ end of the *FMR1* gene are shown in the shaded block. This block, which spans ∼3000 bp, contains three regions that show evidence of ORI activity in some or all of the cell lines tested. These regions are designated as ORI I, II and III. Panel (**B**) shows an expanded view of the shaded region shown in (A) illustrating the position of the three ORIs along with various sequence elements present in this region including the CGG repeat, a potential quadruplex and the remnants of two long interspersed repeated DNA (LINE) elements. The ORI cores, defined as those regions showing the highest number of nascent DNA strands in deep sequencing, are also shown. In both (A and B), the numbering refers to the nucleotide numbering on the X chromosome from the GRCh37/hg19 assembly of the human genome.

this locus before mitosis begins. To examine the basis of the replication fork block, we analyzed the 5′ end of the *FMR1* gene, a region that had previously been shown in two separate studies to contain an ORI in normal (26,27) and FX (26) fibroblasts using an nascent strand abundance assay.

Analysis of genome-wide datasets for ORI activity generated by deep sequencing of purified nascent strands confirmed the presence of ORI activity in a variety of cell types (the colon cancer cell line, HCT116, the leukemia cell line, K562 and embryonic stem cells, Fig. 2). In fact, it revealed the presence of three distinct peaks of increased nascent strand abundance indicative of ORI activity in the 3 kb region containing the 5′ end of the *FMR1* gene. The presence of ORI activity at the 5′ of the *FMR1* gene in these diverse cell types increased the probability that a similar ORI activity may be found in the same region of FX patient lymphoblastoid cells.

The three nascent strand peaks were designated as ORIs I, II and III, with ORI I being the most centromeric (Fig. 2A). The three peaks span a region of ∼3 kb. ORIs I and II are located in the promoter region just upstream of the start of transcription and the CGG/CCG-repeat, while ORI III is located just downstream of the repeat in intron 1 (Fig. 2B). In HCT116 cells, ORI I, II and III are separated by very short regions containing simple repeats. Such sequences would be underrepresented in a genome-wide alignment. Thus, it may be that, in cells like HCT116, a single broad zone of origin initiation activity exists in this region. Peaks with similar nascent strand abundance are not seen for many kilobases on either side of the repeat in all three cell types (Fig. 2A). Thus, this replication zone may be used to replicate the *FMR1* locus in a significant fraction of cells. In practice, ORI I, II and III would likely operate in different cells, with the activity of the other origins within the replicon being suppressed (28,29). The likelihood that replication occurs from a given ORI would be reflected in its nascent strand abundance and this does seem to vary in different cell types.

The core ORI sequence for a given ORI region, as defined by the region containing the largest number of sequence tags obtained by deep sequencing, is likely to be closest to the point of replication initiation in that ORI. The core *FMR1* ORI sequences in ESCs, shown in Figure 2B, have a G + C-content of 52, 60 and 71%, respectively. The core ORI I region contains the remnant of an ancient non-LTR retrotransposable element, L2b, a member of the LINE-2 (L2) family (Fig. 2B). The centromeric end of the full ORI I region contains part of another non-retrotransposon, this time an L1ME5 element which belongs to the family of LINE-1 (L1) retrotransposable elements. The incorporation of parts of these elements into functional regions of the human genome represents two examples of how autonomously replicating repetitive elements present in mammalian genomes can become exapted for use by the host. It has been suggested that G-quadruplexes are an integral feature of many mammalian ORIs with a G4 motif and loops of 1–15 nucleotides

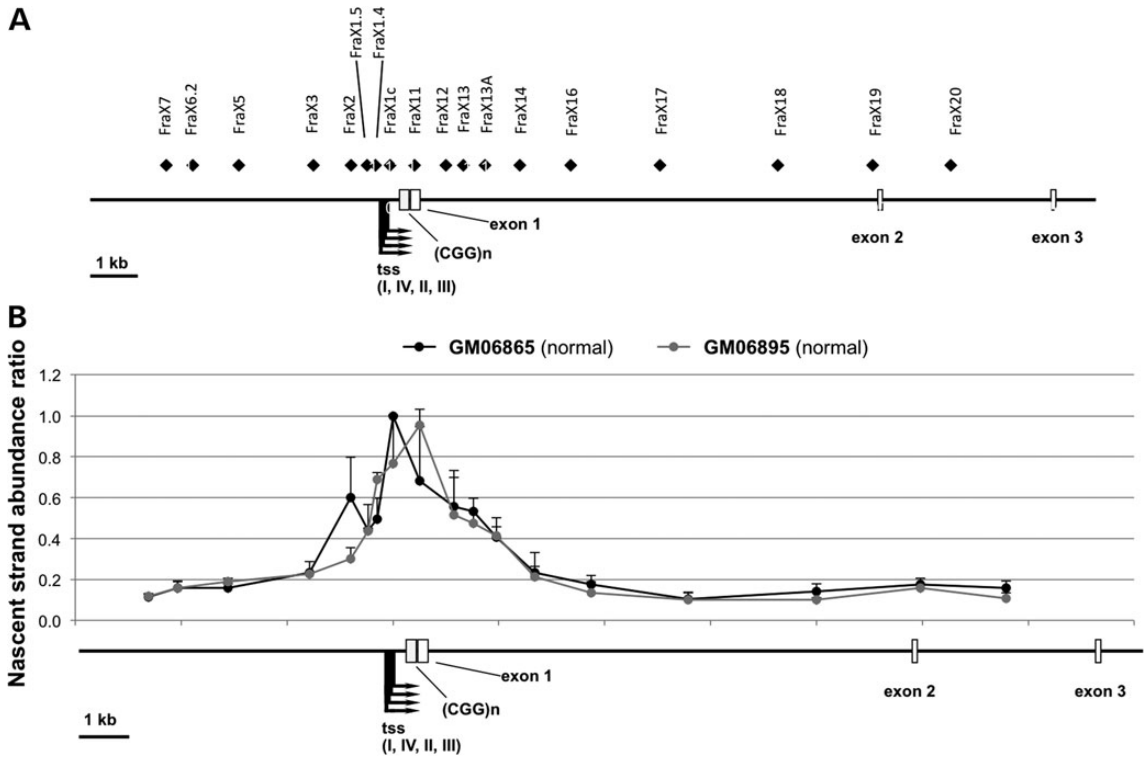

Figure 3. Nascent strand abundance profiles for the *FMR1* region of lymphoblastoid cells from two normal males (GM06865 and GM06895). Panel (**A**) shows the location of the primer pairs used in this study relative to the 5′ end of the *FMR1* gene. Panel (**B**) shows the nascent strand abundance profiles for the GM06865 and GM06895 cell lines. These profiles were determined as described in the Materials and Methods and plotted relative to the four *FMR1* transcriptional start sites (tss). To facilitate comparison between cell lines, the individual data points were expressed as a fraction of the highest peak in each dataset and then averaged.

identified in 91% of ORIs in IMR-90 cells (30). We thus assessed the 5′ end of the *FMR1* gene for the presence of the G4 motif and loop but no such sequences were found. However, quadruplexes can be conformationally complex (31–33) and one sequence with quadruplex-forming potential (5′ GGGGTGAGCTGGGG ATGGGCGAGGGCCGGCGGCAGGTACT AGAGCCGGG CGGGAAGGG 3′) was identified in the core of ORI III (Fig. 2B).

ORI activity that spans the repeat is seen in normal lymphoblastoid cells

We carried out a focused nascent strand abundance study in normal and patient lymphoblastoid cells cultured with and without FdU, using the same primers as used in one of the original fibroblast studies and using fragments 0.5–1 kb in size (26). We also extended the region interrogated by adding a number of additional primer pairs to expand coverage of the *FMR1* gene to ∼17 kb. The locations of the primer sets are shown in Figure 3. As the yield of PCR product produced by amplification through the repeat is strongly dependent on repeat size and methylation status, both of which differ between normal and patient cells, compensating for these differences would be difficult. We therefore used the yield of PCR products from the regions immediately 5′ and 3′ of the repeat as indicators of the abundance of nascent DNA from the repeat region. Quantitative analysis of PCR products from these regions demonstrated that the efficiency of amplification of these regions was the same in total genomic DNA from normal and FM cells. Thus, differences in amplification efficiency should not be a confounding factor in our experiments.

In the absence of FdU, a cell line carrying a normal *FMR1* allele (GM06895) shows a distribution of nascent strands that has a broad, roughly symmetrical peak spanning the CGG•CCG-repeat tract at the 5′ end of the *FMR1* gene (Fig. 3). Thus, there is a zone of active replication at the 5′ end of the *FMR1* gene in lymphoblastoid cells that is slightly telomeric of the ORI previously identified in normal fibroblasts using the same technique (26,27). The height of the peak seen in lymphoblastoid cells is ∼20 times that seen at the well-characterized LaminB2 ORI (data not shown). This peak is just upstream of the CGG/CCG-repeat tract and roughly coincides with ORI II identified by deep sequencing in other cell lines.

In a second normal cell line GM06865, an increase in the amount of nascent strands was also seen in this region. However, the overall nascent strand abundance relative to LaminB2 was lower than it was for GM06895, illustrating the variability in the *FMR1* ORI activity seen with cell lines in the normal range (data not shown). In addition to a peak just 5′ of the repeat, another peak further 5′ of the repeat and a small shoulder in the distribution profile could be seen downstream of the repeat (Fig. 3). A cell line GM06891 that had 112 repeats, and a cell line containing an UFM allele with 477 repeats, GM06897, both showed nascent strand abundance profiles with multiple distinct peaks in the same region (Fig. 4). A

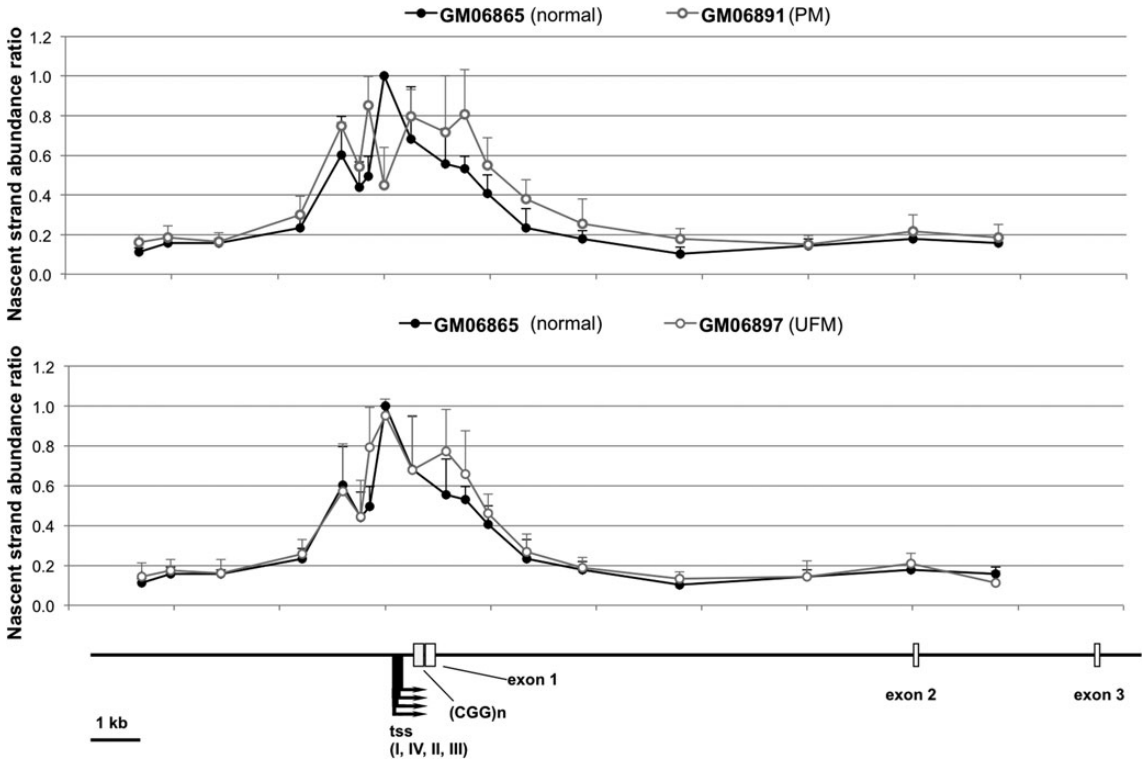

Figure 4. Comparison of the nascent strand abundance profiles of a normal male (GM06865), a PM male (GM06891) and a male UFM carrier (GM06897). The nascent strand abundance profiles were determined as described in the Materials and Methods. To facilitate comparison, the number of CGG repeats, which varied among all three cell lines was not shown to scale, and the data points were expressed as a fraction of the highest peak in each dataset.

spontaneously expanded derivative of GM06891 with 183 repeats showed a very similar nascent strand abundance profile as its 112 repeat-containing precursor (data not shown).

The peaks/shoulders seen in the nascent strand abundance profile of the GM06865 (normal) cell line roughly correspond to ORI I, II and III identified in the whole-genome dataset. The failure to see multiple distinct peaks in some of the other cell lines may reflect the lack of resolution inherent in the assay or variations in the relative usage of the three sites in cell lines from different individuals. However, in all of these cells, the region corresponding to ORI II has a higher abundance of nascent strands than positions corresponding to either ORI I or ORI III. This would be consistent with the idea that the ORI II is more commonly used than the other two ORIs in lymphoblastoid cells.

FM cell lines have nascent strand abundance profiles that are quite different from normal, PM or UFM cell lines

In contrast to what is seen with normal, PM and UFM cell lines, two FXS cell lines GM03200 and GM04025, both having ~500 methylated CGG/CCG-repeats, show profiles with a reduced nascent strand abundance in these cells at the position corresponding to ORI II. Quantification of the PCR product produced from the primer pairs immediately 5′ (Frax1C) and 3′ (Frax11) of the repeat showed the amount of Frax1C PCR product being 13–80 times lower than those found in cells with a normal *FMR1* allele and levels of the Frax11 product being 5–15 times lower. This results in a biphasic pattern with the highest abundance of nascent fragments being located at the ORI I and ORI III positions (Fig. 5). This decline in the amount of amplification from the Frax1C and Frax11 primer pairs is not due to a reduced efficiency of PCR owing to the repeats because a similar PCR efficiency is seen with these primers on total genomic DNA from normal and FX cells (see Materials and Methods). This change in nascent strand abundance profile suggests that replication of the *FMR1* region in patient cells under normal growth conditions differs from that in normal cells and that initiation of DNA replication from the preferred origin in normal cells (ORI II) is either considerably reduced or absent in FXS cells.

FdU does not alter the nascent strand abundance profile in either normal or FXS cells

In order to test the effect of FdU on replication from the *FMR1* ORI cluster, we treated normal and FXS cells with 0.1 μM FdU for 2 and 18 h before examining the nascent strand abundance at the *FMR1* locus. Treatment of cells with FdU did not change the nascent strand abundance profile or the extent of replication activity (normalized to the extent of replication at the LaminB2 ORI) at either time point (Fig. 6 and data not shown). This shows that FdU treatment does not create a specific problem with replication of the FXS allele.

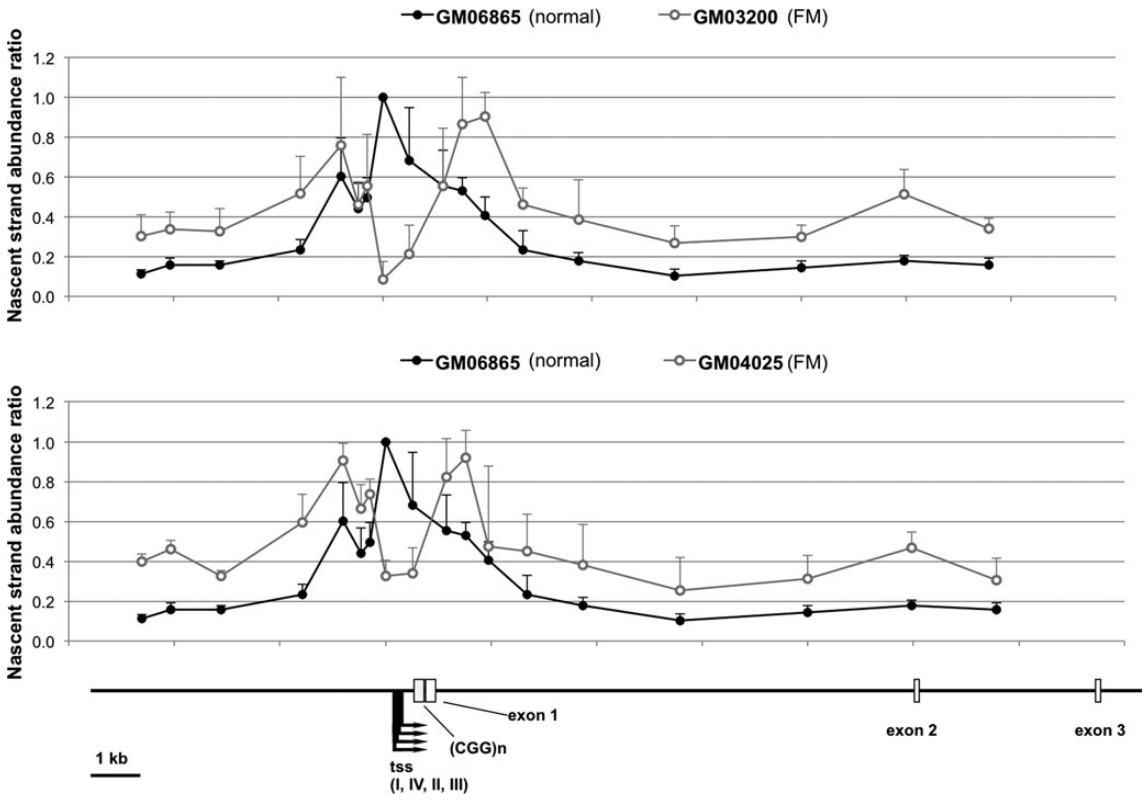

Figure 5. Comparison of the nascent strand abundance profiles of a normal male (GM06865) and two carriers of methylated FMs (GM04025 and GM03200). The nascent strand abundance profiles were determined as described in the Materials and Methods. To facilitate comparison, the number of CGG-repeats, which varied among all three cell lines was not shown to scale, and the data points were expressed as a fraction of the highest peak in each dataset.

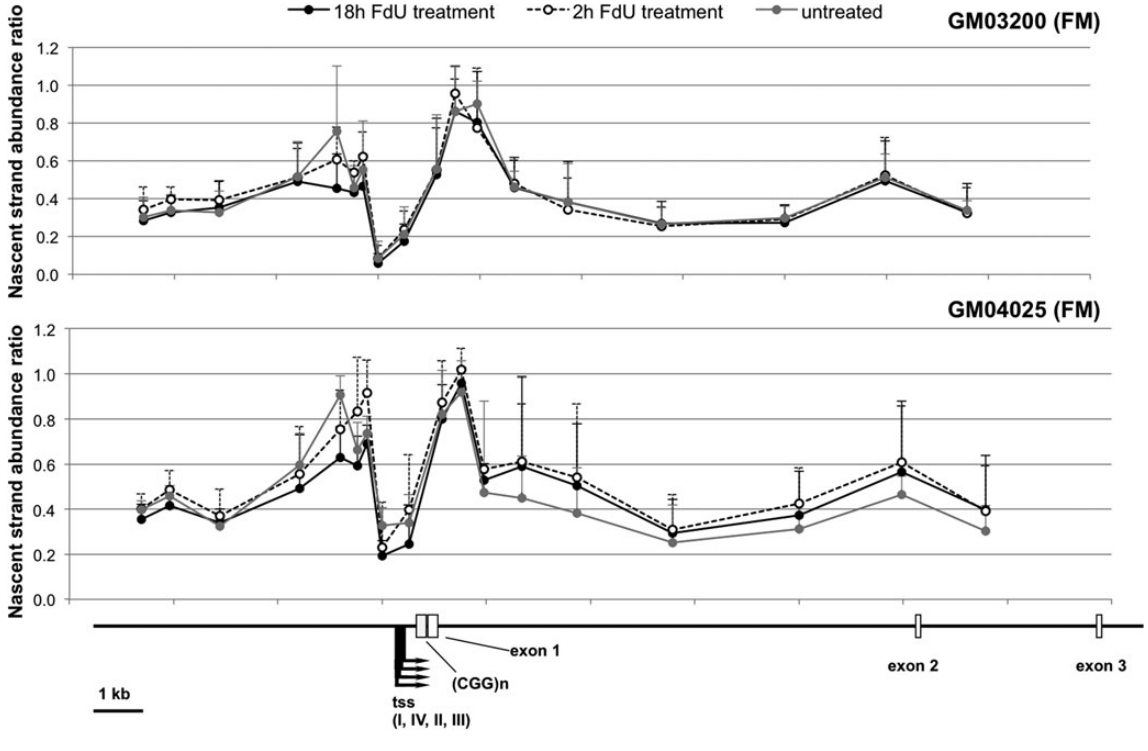

Figure 6. The effect of FdU on the nascent strand abundance profiles of lymphoblastoid cells from methylated FM carriers GM03200 (**A**) and GM04025 (**B**). The nascent strand abundance profiles from untreated cells and cells treated for 2 and 18 h with FdU were determined as described in the Materials and Methods.

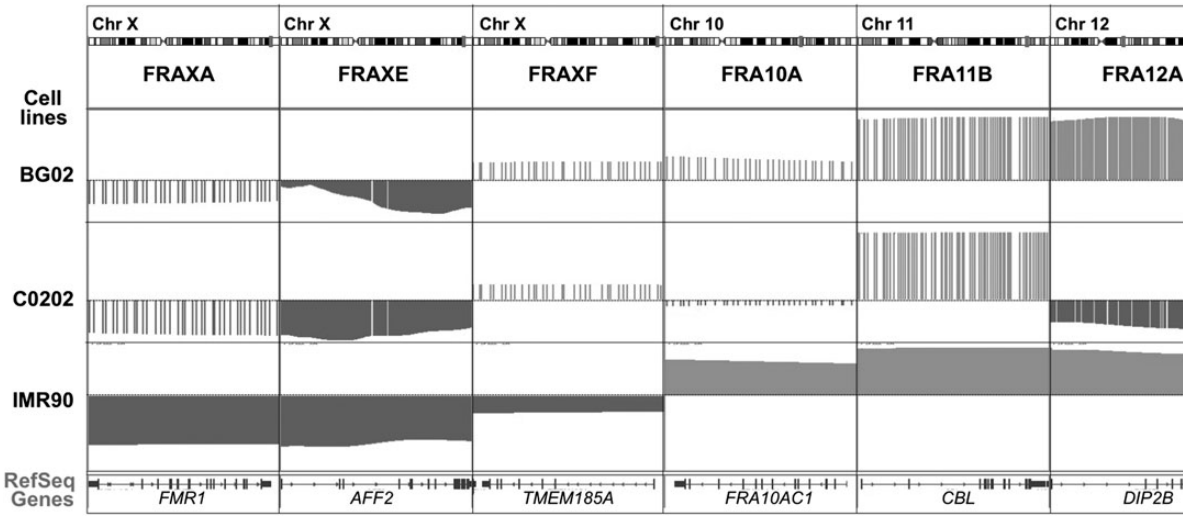

Figure 7. The replication timing of FRAXA, FRAXE, FRAXF, FRA10A, FRA11A, FRA11B and FRA12A in ESCs (BG02), lymphoblastoid (C0202-1) and fibroblast (IMR90) cells. Late replication (defined as the second half of S phase) is shown below the dotted line and early replication is shown above the line. In this representation, each gene is scaled to fit the same sized window. The line density in the BG02 and C0202 samples is a function of the actual size of each gene. The data for each cell lines is shown at the same scale. Shorter lines reflect replication closer to the middle of S phase. The data are taken from GEO datasets GSM500933 (BG02; (42)); GSM500943 (C0202-1; (42)) and GSM923447 (IMR90; ENCODE/University of Washington, (43)).

Table 3. Unmapped Folate/FdU-sensitive fragile site loci and their replication timing

Fragile site	Location	Replication
FRA1M	1p21.3	Late
FRA2A	2q11.2	Mixed
FRA2B	2q13	Mostly early
FRA2K	2q22.3	Mostly late
FRA2L	2p11.2	Mixed
FRA5G	5q35	Mostly early
FRA6A	6p23	Mixed
FRA7A	7p11.2	Mostly early
FRA8A	8q22.3	Early
FRA9A	9p21	Mostly late
FRA9B	9q32	Early
FRA12D	12q24.13	Late
FRA16A	16p13.11	Early
FRA18C	18q22.1	Late
FRA19B	19p13	Early
FRA20A	20q11.23	Early
FRA22A	22q13	Mixed

All folate/FdU-sensitive fragile sites that have been mapped share features in addition to the presence of the CGG/CCG repeat

The normal *FMR1* gene is known to replicate late and FX alleles replicate even later. To assess whether late replication is a conserved feature of folate-sensitive fragile sites, we investigated the replication timing of the other 23 regions known to contain folate-sensitive fragile sites that are expressed in lymphoblastoid cells (1). The cell lines we examined do not contain these fragile sites that are only seen in a fraction of the population. However, the corresponding normal alleles may, like the normal *FMR1* allele, also be late replicating which could influence the replication timing of the fragile allele. However, in both the case of the seven FdU/folate-sensitive fragile sites whose precise location is known (FRAXA (9), FRAXE (34,35), FRAXF (36,37), FRA10A (38), FRA11A (39), FRA11B (40), FRA12A (41)), some are late replicating while others replicate early in S phase (Fig. 7). The same is true for the other 17 FdU-inducible fragile site regions whose precise location has not yet been mapped, with some being located in a chromosomal interval that replicates early, others are located in a late replicating interval, and some are located in an interval containing a mixture of early and late replicating regions (Table 3). Thus, late replication of the locus in which the CGG repeat is situated is not a prerequisite for FdU-inducible chromosome fragility.

However, while fragile sites differ with respect to their locations in early and late replicating regions, all folate/FdU-sensitive fragile sites whose locations are precisely known, are located in the 5′ UTR of genes and are in close proximity to one or more clustered ORIs (Fig. 8).

DISCUSSION

Previous work using nascent strand abundance assays demonstrated the presence of an ORI in the promoter region of the *FMR1* gene in fibroblasts (26,27). Using the same assay, we show here that both normal and FXS patient lymphoblastoid cells show ORI activity that spans the CGG/CCG-repeat region (Figs 4–6). The nascent strand abundance profile corresponds generally to the three distinct peaks of nascent strand abundance that are seen in a variety of other cell types using deep sequencing of nascent DNA strands (Fig. 2). The ORI zone is G + C rich (60.1% overall) and the 3′ half of the ORI zone, particularly ORI III, meets the criteria for a CpG island having a G + C content of ∼69% and an observed-to-expected CpG ratio of 80%. The ORI zone spans the transcription start site(s) as is frequently seen with other ORIs in whole-genome datasets (44,45).

FXS cells, which lack *FMR1* transcriptional activity and express the fragile site strongly, show nascent strand abundance

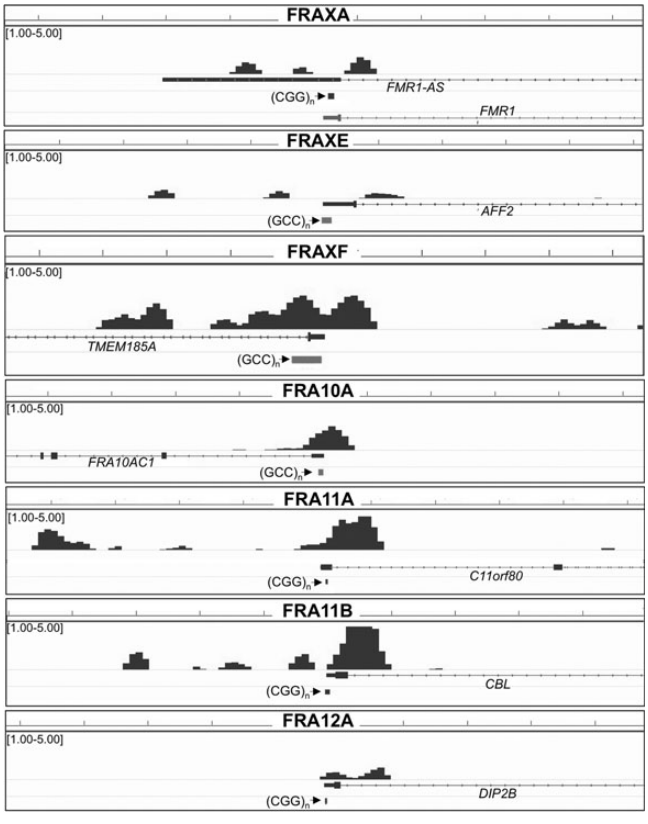

Figure 8. Origins of replication in the vicinity of the FRAXA, FRAXE, FRAXF, FRA10A, FRA11A, FRA11B and FRA12A fragile sites in ESCs. The location of the ORIs in the vicinity of each of the fragile sites was determined by analysis of a human ESC dataset produced from deep sequencing of nascent DNA strands as described in the Materials and Methods. A 10 kb window centered on the major transcription start site of the gene containing the fragile site is shown with each tick mark on the *X*-axis representing 1000 bp. The location of the repeat is shown below the depiction of the gene. The sequence of the repeat unit on the top strand is given.

profiles quite different from the profiles seen in normal, PM or UFM cells even under normal growth conditions. Our data show that the nascent strand abundance from middle of the zone of ORI activity, the region corresponding to ORI II, is reduced in FX cells. This suggests that the expanded methylated repeat affects the amount of replication initiating or extending from this region. It does not seem to reduce the extent of ORI activity at either ORI I or ORI III. However, because the nascent strand abundance assay only examines nascent fragments smaller than 1 kb in length, and ORI I is ∼1 kb away from the repeat, any problem with replication through the repeat may not be visible in the pool of nascent strands from this ORI. Previous studies of FX fibroblasts showed no patient-specific changes in the nascent strand abundance profile but, in fibroblasts, the major peak of ORI activity overlaps with ORI I (26). Thus, any effects of the repeat on replication would also not have been visible in the pool of sub 1 kb nascent strands analyzed in that cell type. As ORI III is roughly equidistant from the repeat as ORI II, an effect of the repeat on replication from ORI III should have been apparent in our assay. The fact that it is not suggests that the repeat has a differential effect on ORI II and ORI III with the most commonly used ORI in the 5′ end of the *FMR1* gene in normal lymphoblastoid cells being dysfunctional in lymphoblastoid cells from individuals that express the FRAXA fragile site. As a 'dominant' ORI in a replicon is thought to repress firing of other closely situated ORIs in the same cell (28,29), fragility may arise because in cells in which replication from ORI II is initiated, replication may need to be completed using more distant ORIs.

We had previously hypothesized that there is an intrinsic problem in this region that accounts for the chromosome fragility seen under normal growth conditions when DNA damage repair proteins like ATM and ATR are inhibited (46). The data we have presented here would be consistent with this idea because the same aberrant replication profile is seen with and without FdU. What the basis of the replication problem is remains to be determined. As FM alleles are silenced, it may be that ORI II is sensitive to the transcriptional status of the *FMR1* gene in ways that ORI I and ORI III are not. However, the fact that the entire 5′ end of the *FMR1* gene that includes the ORI zone is heavily DNA methylated is difficult to reconcile with the idea that ORI II is not used in cells in which the *FMR1* gene is not active, but ORI I and ORI III are. Furthermore, it has been demonstrated that many of the same ORIs are used on active and inactive X chromosomes (47,48) and that transcription is neither necessary nor sufficient for ORI activity at many genes (27,48,49). In fact, it has been suggested that ORI specification is set by transcription in the early embryo (45). As the *FMR1* gene is active in the early embryo even on FM alleles that are subsequently silenced (50–52), it may be that heterochromatinization and silencing of these alleles does not explain the altered pattern of ORI usage we have observed.

An alternative explanation for the effects of the repeat on replication may be found in a previous observation, we have made namely that CGG repeats form secondary structures that are effective blocks to DNA synthesis *in vitro* (16). Stalled replication forks have also been seen on cloned FX repeats transfected into primate cells (20). When ORI II is replicated, the CGG strand of the repeat would be the template for lagging strand DNA synthesis. The presence of CGG repeats on the lagging strand template causes more severe replication fork stalling in bacteria and yeast than is seen when CCG repeats are on the lagging strand template (53,54). This is thought to be due to the fact that the secondary structures formed by the CGG repeats are more stable than those formed by CCG repeats, and are thus more likely to impede the replication fork. When replication initiates at ORI III downstream of the repeat, the CCG-strand would be on the lagging strand template. Thus, ORI III may not be as sensitive as ORI II to the presence of the repeats. Any problem with replication may be exacerbated by the fact that, given the proximity of ORI II, the repeat is likely to be replicated by Pol α, the DNA polymerase involved in the initiation of replication. This polymerase is less processive than Pol δ and ϵ, the polymerases that carry out later stages of DNA replication (55,56), and is more likely to be stalled by regions that form secondary structures (56). Thus, it may be that the repeat has a more deleterious effect on replication from ORI II than it would have on replication from a more distal ORI.

Treatment of FXS cells with FdU induces the FRAXA fragile site but it does not change the nascent strand abundance profiles

of those cells suggesting that FdU does not act by exacerbating the problem with replication of this locus. Rather, it may be acting non-specifically via its effect on nucleotide pools and thus on replication in general. Slower replication as a result of the nucleotide pool imbalance would increase the proportion of FXS cells that enter mitosis before replication of the *FMR1* region is complete. If this is so, then treatment with CPT may reduce the fragility of FXS alleles not by decreasing the opportunity for replication-blocking lesions to form, as proposed for some APH-inducible fragile sites (4), but via its ability to efficiently induce a cell cycle checkpoint, thus increasing the length of time that the cells have to complete replication of the long CGG/CCG-repeat tract before mitosis begins. This interpretation makes sense in light of the fact that MMC, which does not affect Topo I, has the same effect on FX chromosome fragility. It is also consistent with our previous demonstration that UCN-01, an inhibitor of CHK1, an important S-phase checkpoint kinase, increases chromosome fragility at this locus (46), because loss of checkpoint control would mean that cells enter mitosis more rapidly.

The 23 other folate/FdU-sensitive fragile sites that have been described likely have a molecular basis similar to that of FRAXA. Those sites that are located in regions that normally replicate early may represent examples of CGG-repeat tracts that are much longer than the repeat tracts typical of FXS. Such repeat tracts may form an even more serious impediment to DNA replication than the FX repeat and thus take even longer to complete replication. They may thus still not have time to complete replication before mitosis begins despite their replication having started earlier in S phase.

Interestingly, as with FRAXA, all of the folate/FdU-inducible fragile sites whose locations have been precisely mapped also are in the 5′ UTR of genes, have one or more ORIs in close proximity (Fig. 8 and (49)) and are all associated with repeat-induced DNA methylation. It may be that their location in the 5′ UTR of genes predisposes them to fragility because the repeat is frequently replicated by Pol α initiating from a nearby ORI or because the associated gene silencing affects the activity of that ORI.

The fact that the impediment to DNA replication is seen at FRAXA even in the absence of FdU, suggests that chromosome fragility may not simply be an *in vitro* artifact. Female carriers of methylated FM alleles have a high incidence of Turner syndrome, in which the X chromosome carrying the FM allele is lost (10). This may be a downstream consequence of chromosome fragility and breakage at FRAXA; Fusion of two broken sister chromatids could result in a dicentric chromosome that would be at high risk of being lost on subsequent cell division. Whether other FdU-inducible fragile sites also make cells prone to chromosome loss is unclear. However, this may only be apparent for fragile sites on the X chromosome because cells that have lost autosomes would not be viable and thus would not be detected.

MATERIALS AND METHODS

Cell lines, primers and probes

The cell lines used in this study are lymphoblastoid cells derived from males with normal *FMR1* alleles, and from premutation (PM), who have 55–200 repeats, and FM carriers. They are listed in Table 1. They were all obtained from the Coriell collection (Camden, NJ, USA). Cells were grown in RPMI 1640 GlutaMAX medium (Life Technologies, Grand Island, NY, USA) with 10% of heat-inactivated fetal bovine serum and 1× antibiotic-antimycotic (both from Life Technologies). The primers used in this study are listed in Supplementary Material, Table S1. They were all purchased from Integrated DNA Technologies (Coralville, IA). The BAC clone RP11-489K19 containing the human *FMR1* gene was purchased from Empire Genomics (Buffalo, NY).

Chromosome fragility and the mitotic index

Ten million cells were treated for 18 h in medium supplemented with 0.1 μM FdU (Sigma-Aldrich Corp, St. Louis, MO) and/or 3 nM or 30 mM CPT (Sigma). Gibco® KaryoMAX® Colcemid™ solution (Life Technologies) was then added to a final concentration of 75 ng/ml and the cells incubated for an additional 2 h at 37°C. The cells were then incubated in 0.075 M KCl (37°C) at RT for 20 min, treated with 1 ml of Carnoy fixative (methanol: acetic acid 3:1) at 4°C for 10 min and then 2 ml of Carnoy fixative at −20°C for 30 min. The fixative was replaced and the cell suspension was dropped onto dry slides on a wet hot table (56°C). Slides were then treated with 25 ng/ml Proteinase K (Life Technologies) for 10 min at RT, washed twice in 2× SSC for 5 min, dehydrated in an alcohol series and denatured in 70% formamide in 2× SSC at 70°C for 5 min. The BAC clone containing the human *FMR1* (RP11-489K19) gene was labeled by nick-translation with biotin or digoxygenin (Roche Applied Science, Indianapolis, IN) using the supplier's instructions. For each slide, a 20 μl solution containing 0.3 μg of labeled BAC and 10 μg of human Cot1 DNA (Invitrogen), 50% formamide, 25% dextran sulfate in 2× SSC was prepared, denatured at 96°C for 8 min and incubated for 1 h at 37°C. This solution was placed on the slide, covered with a cover slip and incubated at 37°C overnight. The slide was then washed in 50% formamide in 2× SSC three times for 5 min at 45°C, in 0.1× SSC three times for 5 min at 60°C and then blocked with 3% BSA in 4× SSC/0.1%Tween20. Detection was carried with FITC-anti-digoxigenin Fab-fragments conjugated with 5 (6)-carboxy-fluorescein (Roche) or Alexa-555-streptavidin (Invitrogen) diluted 1:200 in blocking solution (3% BSA in 4× SSC/0.1% Tween20). Slides were stained in the dark with 10 μg/ml DAPI (Invitrogen) diluted in Fluoro-Gel mounting medium (Electron Microscopy Sciences, Hatfield, PA). Metaphase spreads in which fragility at FRAXA was detected were counted using a Leica DB5500 fluorescent microscope. At least 100 metaphases were analyzed for each experiment. The metaphase/mitotic index was calculated as the number of metaphase cells divided by the total number of cells in the population. For these experiments ∼2000–3000 cells were counted. Statistical significance was calculated by Fisher's two-tailed test using GraphPad Software (http://www.graphpad.com/quickcalcs/).

Nascent strand abundance assay

A nascent strand abundance assay that was previously used to study the ORI activity of the *FMR1* gene was used here with some minor modifications (26). Specifically, 70×10^6 cells

with or without 0.1 μM FdU (Sigma) were grown in a volume of 70 ml for 2 h or for 18 h. Cells were then washed with 1× PBS and collected in 240 μl of 10% glycerol in 1× PBS. Agarose (1%) was melted in water and cooled to 60°C. Sodium hydroxide and EDTA were added to a final concentration of 50 and 1 mM, respectively, and used to prepare a gel for electrophoresis. The electrophoresis tank was filled to the level of the surface of the gel with a buffer containing 50 mM NaOH and 1 mM EDTA. The cells were then loaded into the wells of the gel and cell lysis allowed to occur for 15 min. The gel was then subjected to electrophoresis for 5–6 h at 30 V, neutralized in 1× TAE and stained with ethidium bromide. DNA fragments 0.5–1 kb in length were purified from the gel using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's recommendations. Real-time PCR was carried out with 0.1 μM of the indicated primers using 10 ng of DNA and a *Power* SYBR® Green PCR Master Mix (Life Technologies) in a StepOnePlus™ real-time PCR machine (Life Technologies). The primer binding sites span ∼17 kb region of the *FMR1* gene flanking the CGG-repeat tract. All primers pairs used are listed in Table 2 and their positions on the *FMR1* gene shown in Figure 3. The primer pair FraX1c corresponds to the primer with the same name used in a previous study that identified an origin of replication (ORI), at the 5′ end of the *FMR1* gene (26). Frax1.2, Frax1.4, Frax1.5, Frax2, FraX3, FraX5, Frax11 and Frax12 also correspond to primers used in the previous report. The same general naming principle was used for additional primers that were used to interrogate regions both 5′ and 3′ of these primers. The suffixes SX and DX correspond the forward and reverse primers, respectively, as previously described (26). It had been previously reported that CGG repeat decreased the PCR efficiency for primers immediately adjacent to this repeat (26). However, comparison of the PCR yield of different primer pairs on purified genomic DNA from normal and FM carriers showed no significant differences (Supplementary Material, Table S2). It allowed us to use all the primer pairs listed without predigestion of the genomic DNA by a restriction enzyme immediately upstream or downstream of the repeat. It also allowed us to use genomic DNA from a single cell line (GM06865) to generate a standard curve for these experiments. The absolute amount of each PCR product was ascertained after normalizing by a standard curve generated from total genomic DNA. As the amount of nascent DNA obtained differed in different cell lines even after normalization to the amount of nascent DNA amplified from the lamin B2 origin (26), each data point was expressed as a fraction of the highest peak in the dataset. Three independent biological replicates were done for each cell line, the data averaged and the standard deviation calculated for each primer pair.

In silico ORI analysis

The datasets for analysis of the *FMR1* ORIs were generated by deep sequencing of nascent strands from HCT116, K562 and a human ESC lines. Replication initiation data used in this manuscript have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE28911. The K562 cell dataset has been previously described (44).

The ORIs in the *FMR1* region were visualized in two ways. IGV 2.3.18, the Broad Institute's Integrative Genomics Viewer (http://www.broadinstitute.org/igv/; (57)) was used to examine TDF files containing the data from each cell line so as to be able to visualize the ORI activity in the 100 kb region centered on the *FMR1* gene and to define the extents of the ORIs in the 5′ end of the *FMR1* gene. The ORI regions were defined and exported as BED files. The UCSC genome browser was then used to examine and annotate a much smaller region of the *FMR1* gene using the BED files as custom tracks. The Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/index.html) was used to analyze the sequence of the ORI zone. QGRS Mapper (http://bioinformatics.ramapo.edu/QGRS/analyze.php) was used to identify quadruplex motifs.

Analysis of replication timing

The GEO datasets GSM500933 (42)), GSM500943 (42)) and GSM923447 (43) for BG02, C0202-1 and IMR90 cells, respectively, were downloaded from http://www.ncbi.nlm.nih.gov/geo. For the GSM500943 dataset, a custom script was used to convert oligoIDs into hg18 genomic positions and UCSC tool liftOver (http://genome.ucsc.edu/cgi-bin/hgLiftOver) used to generate hg19 positions. In order to produce data visualization comparable with GSM500943, the WaveSignal file for the GSM923447 dataset in bigWig format was converted to bedGraph format, using the UCSC program bigWigtoBedGraph (http://genome.ucsc.edu/goldenPath/help/bigWig.html), the median of the replication values was determined and subtracted from all values.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank Eric Bouhassira and Zipora Etzion at the Albert Einstein College of Medicine Pluripotent Stem Cell Facility for culturing ESC cells, Melvenia Martin and Haiqing Fu, at the CCR sequencing facility headed by Bao Tran, and Natalie Abrams and Li Jia from the CCR bioinformatics facility for expert help with sequence analysis.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by a grant from the intramural program of the National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health to K.U.; from the Center for Cancer Research, National Cancer Institute, National Institutes of Health to M.A. and from the Russian Foundation for Basic Research to D.Y.

REFERENCES

1. Lukusa, T. and Fryns, J.P. (2008) Human chromosome fragility. *Biochim. Biophys. Acta*, **1779**, 3–16.

2. Freudenreich, C.H. (2007) Chromosome fragility: molecular mechanisms and cellular consequences. *Front. Biosci.*, **12**, 4911–4924.
3. Mishmar, D., Rahat, A., Scherer, S.W., Nyakatura, G., Hinzmann, B., Kohwi, Y., Mandel-Gutfroind, Y., Lee, J.R., Drescher, B., Sas, D.E. *et al.* (1998) Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc. Natl. Acad. Sci. USA*, **95**, 8141–8146.
4. Arlt, M.F. and Glover, T.W. (2010) Inhibition of topoisomerase I prevents chromosome breakage at common fragile sites. *DNA Repair (Amst.)*, **9**, 678–689.
5. Ozeri-Galai, E., Lebofsky, R., Rahat, A., Bester, A.C., Bensimon, A. and Kerem, B. (2011) Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. *Mol. Cell*, **43**, 122–131.
6. Letessier, A., Millot, G.A., Koundrioukoff, S., Lachages, A.M., Vogt, N., Hansen, R.S., Malfoy, B., Brison, O. and Debatisse, M. (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*, **470**, 120–123.
7. Palakodeti, A., Lucas, I., Jiang, Y., Young, D.J., Fernald, A.A., Karrison, T. and Le Beau, M.M. (2010) Impaired replication dynamics at the FRA3B common fragile site. *Hum. Mol. Genet.*, **19**, 99–110.
8. Fu, Y.H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G. Jr, Warren, S.T. *et al.* (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, **67**, 1047–1058.
9. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. *et al.* (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905–914.
10. Dobkin, C., Radu, G., Ding, X.H., Brown, W.T. and Nolin, S.L. (2009) Fragile X prenatal analyses show full mutation females at high risk for mosaic Turner syndrome: fragile X leads to chromosome loss. *Am. J. Med. Genet. A*, **149A**, 2152–2157.
11. D'Anna, J.A., Crissman, H.A., Jackson, P.J. and Tobey, R. (1985) Time-dependent changes in H1 content, H1 turnover, DNA elongation, and the survival of cells blocked in early S phase by hydroxyurea, aphidicolin, or 5-fluorodeoxyuridine. *Biochemistry*, **24**, 5020–5026.
12. Meyers, M., Hwang, A., Wagner, M.W. and Boothman, D.A. (2004) Role of DNA mismatch repair in apoptotic responses to therapeutic agents. *Environ. Mol. Mutagen.*, **44**, 249–264.
13. Nadel, Y., Weisman-Shomer, P. and Fry, M. (1995) The fragile X syndrome single strand d(CGG)n nucleotide repeats readily fold back to form unimolecular hairpin structures. *J. Biol. Chem.*, **270**, 28970–28977.
14. Mitas, M., Yu, A., Dill, J. and Haworth, I.S. (1995) The trinucleotide repeat sequence d(CGG)15 forms a heat-stable hairpin containing Gsyn. Ganti base pairs. *Biochemistry*, **34**, 12803–12811.
15. Yu, A., Barron, M.D., Romero, R.M., Christy, M., Gold, B., Dai, J., Gray, D.M., Haworth, I.S. and Mitas, M. (1997) At physiological pH, d(CCG)15 forms a hairpin containing protonated cytosines and a distorted helix. *Biochemistry*, **36**, 3687–3699.
16. Usdin, K. and Woodford, K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.*, **23**, 4202–4209.
17. Fojtik, P. and Vorlickova, M. (2001) The fragile X chromosome (GCC) repeat folds into a DNA tetraplex at neutral pH. *Nucleic Acids Res.*, **29**, 4684–4690.
18. Renciuk, D., Kypr, J. and Vorlickova, M. (2011) CGG repeats associated with fragile X chromosome form left-handed Z-DNA structure. *Biopolymers*, **95**, 174–181.
19. Mariappan, S.V., Catasti, P., Chen, X., Ratliff, R., Moyzis, R.K., Bradbury, E.M. and Gupta, G. (1996) Solution structures of the individual single strands of the fragile X DNA triplets (GCC)n.(GGC)n. *Nucleic Acids Res.*, **24**, 784–792.
20. Voineagu, I., Surka, C.F., Shishkin, A.A., Krasilnikova, M.M. and Mirkin, S.M. (2009) Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility. *Nat. Struct. Mol. Biol.*, **16**, 226–228.
21. Fry, M. and Loeb, L.A. (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. USA*, **91**, 4950–4954.
22. Goldman, M.A., Holmquist, G.P., Gray, M.C., Caston, L.A. and Nag, A. (1984) Replication timing of genes and middle repetitive sequences. *Science*, **224**, 686–692.
23. Webb, T. (1992) Delayed replication of Xq27 in individuals with the fragile X syndrome. *Am. J. Med. Genet.*, **43**, 1057–1062.
24. Hansen, R.S., Canfield, T.K., Lamb, M.M., Gartler, S.M. and Laird, C.D. (1993) Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell*, **73**, 1403–1409.
25. Iyer, V.N. and Szybalski, W. (1963) A molecular mechanism of mitomycin action: linking of complementary DNA strands. *Proc. Natl. Acad. Sci. USA*, **50**, 355–362.
26. Gray, S.J., Gerhardt, J., Doerfler, W., Small, L.E. and Fanning, E. (2007) An origin of DNA replication in the promoter region of the human fragile X mental retardation (FMR1) gene. *Mol. Cell. Biol.*, **27**, 426–437.
27. Brylawski, B.P., Chastain, P.D. II, Cohen, S.M., Cordeiro-Stone, M. and Kaufman, D.G. (2007) Mapping of an origin of DNA replication in the promoter of fragile X gene FMR1. *Exp. Mol. Pathol.*, **82**, 190–196.
28. Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R. *et al.* (2011) Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Res.*, **21**, 1438–1449.
29. Borowiec, J.A. and Schildkraut, C.L. (2011) Open sesame: activating dormant replication origins in the mouse immunoglobulin heavy chain (Igh) locus. *Curr. Opin. Cell Biol.*, **23**, 284–292.
30. Besnard, E., Babled, A., Lapasset, L., Milhavet, O., Parrinello, H., Dantec, C., Marin, J.M. and Lemaitre, J.M. (2012) Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat. Struct. Mol. Biol.*, **19**, 837–844.
31. Usdin, K. (1998) NGG-triplet repeats form similar intrastrand structures: implications for the triplet expansion diseases. *Nucleic Acids Res*, **26**, 4078–4085.
32. Weitzmann, M.N., Woodford, K.J. and Usdin, K. (1998) The mouse Ms6-hm hypervariable microsatellite forms a hairpin and two unusual tetraplexes. *J. Biol. Chem.*, **273**, 30742–30749.
33. Woodford, K.J., Howell, R.M. and Usdin, K. (1994) A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *J. Biol. Chem.*, **269**, 27029–27035.
34. Gecz, J., Gedeon, A.K., Sutherland, G.R. and Mulley, J.C. (1996) Identification of the gene FMR2, associated with FRAXE mental retardation. *Nat. Genet.*, **13**, 105–108.
35. Gu, Y., Shen, Y., Gibbs, R.A. and Nelson, D.L. (1996) Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. *Nat. Genet.*, **13**, 109–113.
36. Parrish, J.E., Oostra, B.A., Verkerk, A.J., Richards, C.S., Reynolds, J., Spikes, A.S., Shaffer, L.G. and Nelson, D.L. (1994) Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. *Nat. Genet.*, **8**, 229–235.
37. Ritchie, R.J., Knight, S.J., Hirst, M.C., Grewal, P.K., Bobrow, M., Cross, G.S. and Davies, K.E. (1994) The cloning of FRAXF: trinucleotide repeat expansion and methylation at a third fragile site in distal Xqter. *Hum. Mol. Genet.*, **3**, 2115–2121.
38. Sarafidou, T., Kahl, C., Martinez-Garay, I., Mangelsdorf, M., Gesk, S., Baker, E., Kokkinaki, M., Talley, P., Maltby, E.L., French, L. *et al.* (2004) Folate-sensitive fragile site FRA10A is due to an expansion of a CGG repeat in a novel gene, FRA10AC1, encoding a nuclear protein. *Genomics*, **84**, 69–81.
39. Debacker, K., Winnepenninckx, B., Longman, C., Colgan, J., Tolmie, J., Murray, R., van Luijk, R., Scheers, S., Fitzpatrick, D. and Kooy, F. (2007) The molecular basis of the folate-sensitive fragile site FRA11A at 11q13. *Cytogenet. Genome Res.*, **119**, 9–14.
40. Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L., Langdon, W.Y., Sutherland, G.R., Richards, R.I. and Tunnacliffe, A. (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. *Nature*, **376**, 145–149.
41. Winnepenninckx, B., Debacker, K., Ramsay, J., Smeets, D., Smits, A., FitzPatrick, D.R. and Kooy, R.F. (2007) CGG-repeat expansion in the DIP2B gene is associated with the fragile site FRA12A on chromosome 12q13.1. *Am. J. Hum. Genet.*, **80**, 221–231.
42. Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S. and Gilbert, D.M. (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.*, **20**, 761–770.
43. Hansen, R.S., Thomas, S., Sandstrom, R., Canfield, T.K., Thurman, R.E., Weaver, M., Dorschner, M.O., Gartler, S.M. and Stamatoyannopoulos, J.A.

(2010) Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc. Natl. Acad. Sci. USA*, **107**, 139–144.
44. Martin, M.M., Ryan, M., Kim, R., Zakas, A.L., Fu, H., Lin, C.M., Reinhold, W.C., Davis, S.R., Bilke, S., Liu, H. *et al.* (2011) Genome-wide depletion of replication initiation events in highly transcribed regions. *Genome Res.*, **21**, 1822–1832.
45. Sequeira-Mendes, J., Diaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N. and Gomez, M. (2009) Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS Genet.*, **5**, e1000446.
46. Kumari, D., Somma, V., Nakamura, A.J., Bonner, W.M., D'Ambrosio, E. and Usdin, K. (2009) The role of DNA damage response pathways in chromosome fragility in Fragile X syndrome. *Nucleic Acids Res.*, **37**, 4385–4392.
47. Cohen, S.M., Brylawski, B.P., Cordeiro-Stone, M. and Kaufman, D.G. (2003) Same origins of DNA replication function on the active and inactive human X chromosomes. *J. Cell. Biochem.*, **88**, 923–931.
48. Gomez, M. and Brockdorff, N. (2004) Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage. *Proc. Natl. Acad. Sci. USA*, **101**, 6923–6928.
49. Chastain, P.D. II, Cohen, S.M., Brylawski, B.P., Cordeiro-Stone, M. and Kaufman, D.G. (2006) A late origin of DNA replication in the trinucleotide repeat region of the human FMR2 gene. *Cell Cycle*, **5**, 869–872.
50. Eiges, R., Urbach, A., Malcov, M., Frumkin, T., Schwartz, T., Amit, A., Yaron, Y., Eden, A., Yanuka, O., Benvenisty, N. *et al.* (2007) Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplantation genetically diagnosed embryos. *Cell Stem Cell*, **1**, 568–577.
51. Sutcliffe, J.S., Nelson, D.L., Zhang, F., Pieretti, M., Caskey, C.T., Saxe, D. and Warren, S.T. (1992) DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.*, **1**, 397–400.
52. Malter, H.E., Iber, J.C., Willemsen, R., de Graaff, E., Tarleton, J.C., Leisti, J., Warren, S.T. and Oostra, B.A. (1997) Characterization of the full fragile X syndrome mutation in fetal gametes. *Nat. Genet.*, **15**, 165–169.
53. Samadashwily, G.M., Raca, G. and Mirkin, S.M. (1997) Trinucleotide repeats affect DNA replication *in vivo*. *Nat. Genet.*, **17**, 298–304.
54. Pelletier, R., Krasilnikova, M.M., Samadashwily, G.M., Lahue, R. and Mirkin, S.M. (2003) Replication and expansion of trinucleotide repeats in yeast. *Mol. Cell. Biol.*, **23**, 1349–1357.
55. Weiser, T., Gassmann, M., Thommes, P., Ferrari, E., Hafkemeyer, P. and Hubscher, U. (1991) Biochemical and functional comparison of DNA polymerases alpha, delta, and epsilon from calf thymus. *J. Biol. Chem.*, **266**, 10420–10428.
56. Focher, F., Gassmann, M., Hafkemeyer, P., Ferrari, E., Spadari, S. and Hubscher, U. (1989) Calf thymus DNA polymerase delta independent of proliferating cell nuclear antigen (PCNA). *Nucleic Acids Res.*, **17**, 1805–1821.
57. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.