Cis-acting DNA sequence at a replication origin promotes repeat expansion to fragile X full mutation

Jeannine Gerhardt,[1] Nikica Zaninovic,[2] Qiansheng Zhan,[2] Advaitha Madireddy,[1] Sarah L. Nolin,[3] Nicole Ersalesi,[3] Zi Yan,[1] Zev Rosenwaks,[2] and Carl L. Schildkraut[1]

[1]Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461
[2]Center for Reproductive Medicine and Infertility, Weill Cornell Medical College, New York, NY 10021
[3]New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY 10314

Fragile X syndrome (FXS) is caused by CGG repeat expansion that leads to *FMR1* silencing. Women with a premutation allele are at risk of having a full mutation child with FXS. To investigate the mechanism of repeat expansion, we examined the relationship between a single-nucleotide polymorphism (SNP) variant that is linked to repeat expansion in haplogroup D and a replication origin located ∼53 kb upstream of the repeats. This origin is absent in FXS human embryonic stem cells (hESCs), which have the SNP variant C, but present in the nonaffected hESCs, which have a T variant. The SNP maps directly within the replication origin. Interestingly, premutation hESCs have a replication origin and the T variant similar to nonaffected hESCs. These results suggest that a T/C SNP located at a replication origin could contribute to the inactivation of this replication origin in FXS hESCs, leading to altered replication fork progression through the repeats, which could result in repeat expansion to the FXS full mutation.

Introduction

Fragile X syndrome (FXS) is the most common inherited form of intellectual disability in males. 1 in 4,000 boys is born with this disease. FXS is caused by the expansion of a CGG trinucleotide repeat (TNR) tract in the 5′ UTR of the *FMR1* (*fragile X mental retardation 1*) gene. Normal individuals have between 6 and 54 repeats, whereas affected individuals with the full mutation have an *FMR1* allele with >200 repeats. The *FMR1* gene is methylated and transcriptionally silent in FXS patients. Loss of the *FMR1*-encoded protein FMRP leads to the disease symptoms (Fu et al., 1991; Heitz et al., 1991; Pieretti et al., 1991; Verkerk et al., 1991; Sutcliffe et al., 1992; De Rubeis and Bagni, 2011; Willemsen et al., 2011; Santoro et al., 2012). A carrier with the premutation has 55–200 repeats. 1 in ∼230 women and 1 in 360 men carry the premutation, which is associated with fragile X–related primary ovarian insufficiency and fragile X–associated tremor/ataxia syndrome in some carriers. Males with very large premutation alleles may have mild cognitive and behavioral deficits (Sherman, 2000; Hagerman and Hagerman, 2004; Sullivan et al., 2005).

Expanded repeats are inherited as a result of repeat instability that is thought to occur during germ cell generation and early stages of embryogenesis (Pearson et al., 2005; Mirkin, 2007; López Castel et al., 2010; Rousseau et al., 2011). Somatic repeat expansion does not occur postnatally in FXS patients (Reyniers et al., 1993, 1999; Wöhrle et al., 1993), indicating that this phenomenon only occurs in specific phases of very early development. Repeat instability during embryological development leads to repeat length mosaicism within and between different tissues in patients (Nolin et al., 1994; Dobkin et al., 1996). Expansion from premutation to full mutation occurs exclusively during maternal transmission in women with a premutation repeat size (Pearson et al., 2005; Mirkin, 2007; López Castel et al., 2010; Rousseau et al., 2011).

The mechanism for TNR instability is not clearly understood. Aberrant DNA replication and DNA repair are suspected to play a role in the mechanisms leading to repeat expansion (Gray et al., 2007; Kim and Mirkin, 2013). To examine CGG repeat instability in early embryonic development, we determined the replication profile at the *FMR1* locus in full mutation human embryonic stem cells (ESCs; hESCs;

The Rockefeller University Press $30.00
J. Cell Biol. Vol. 206 No. 5   599–607
www.jcb.org/cgi/doi/10.1083/jcb.201404157

JCB   599

Figure 1. **A replication initiation site upstream of the repeats overlaps with an SNP in nonaffected cells and is absent in FXS hESCs.** (A and C) The previously identified replication initiation sites ∼50 kb upstream of the repeats (Gerhardt et al., 2014) were mapped more in detail by nascent strand abundance. (top) Map of the *FMR1* locus upstream of the repeats. The SfiI restriction enzyme site and position of primers used in the nascent strand abundance assay are indicated. SNP primer pair binds at the SNP T/C (ss71651738 or WEX70) on the X chromosome (Chr. X; Fig. S1, A and B), which cosegregates with a chromosome haplotype at the highest risk for repeat expansion (Table S1; Ennis et al., 2007). Primer pairs A, B1, B2, B3, and C amplify surrounding genomic regions. (A, bottom) Nascent strand DNA was isolated from nonaffected hESCs (H14) and fetal fibroblast (GM00011). Nascent strand enrichment at the SNP was determined by real-time PCR using specific primer pairs. (B) As a control, the nascent strand enrichment was determined at the well-characterized Or6 replication origin (Gerhardt et al., 2006) on chromosome 12 (with Or6 primer pair) and at the adjacent control region (N primer pair). (C and D) Nascent strand DNA was isolated from FXS hESCs SI-214 and WCMC37 upstream of the CGG repeats (C) and at a control region (D). Enrichment of the nascent strand at the SNP and neighboring genomic regions was determined by real-time PCR with primer pairs shown above the map. No nascent strand enrichment was obtained for primer B. There was no significant enrichment of nascent strand DNA in FXS hESCs at the SNP in comparisons to control region C in contrast to nonaffected cells in three separate experiments (SI-214: **, P = 0.075; WCMC37: **, P = 0.81; H14: *, P < 0.0005; GM00011: *, P = 0.0047; also see Fig. S1). Error bars indicate the standard error of the mean of three experiments.
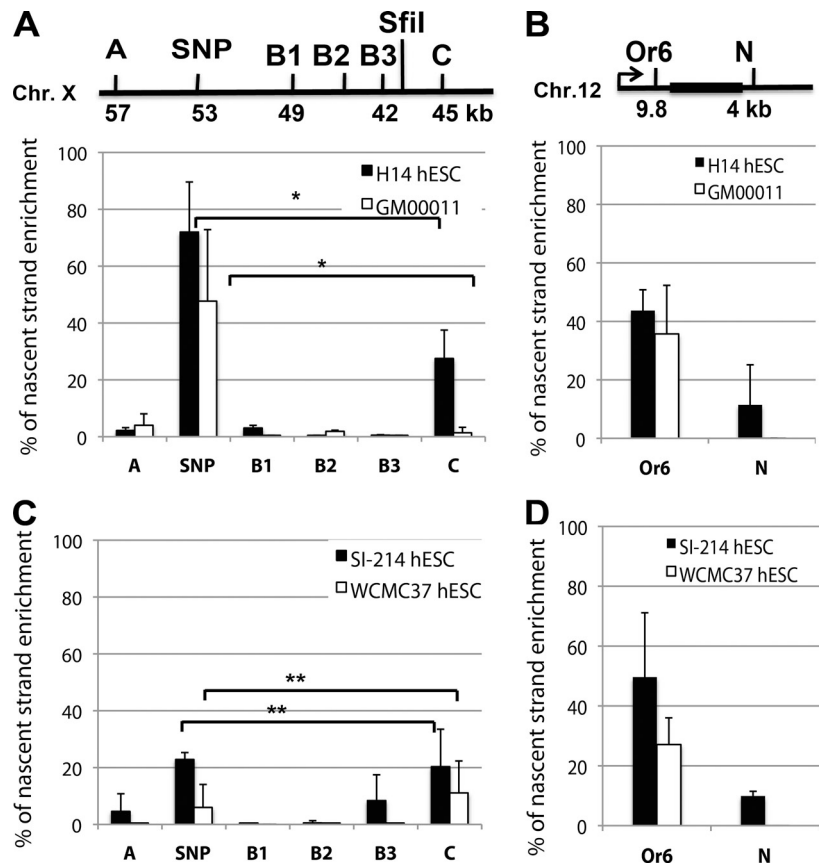
FXS hESCs; Gerhardt et al., 2014). Our results indicate that the absence of replication initiation sites ∼50 kb upstream of the CGG repeats at the *FMR1* locus leads to an alteration in replication fork progression through the CGG repeats. However, no common DNA sequences or epigenetic elements that define replication origins or lead to replication origin firing in mammalian cells have yet been identified (Méchali, 2010). Inactivation of the replication initiation sites in FXS hESCs may result from changes in the chromatin structure or in the DNA sequence.

The risk of having an FXS child is much higher in premutation carriers with family history of FXS (Nolin et al., 2011), suggesting that linked genetic factors (cis-elements) in combination with repeat size influence repeat instability. Recent studies have shown that AGG interruptions in the CGG repeats dramatically lower the risk for expansion in premutation mothers with expanded repeats (Yrigollen et al., 2012; Nolin et al., 2013). The occurrence of AGG interruptions in the CGG repeats results in a more stable CGG repeat size (Eichler et al., 1994; Pearson et al., 1998). This is probably caused by reduced formation of secondary repeat structures by the repeated triplets. However, even after accounting for the influence of repeat length and AGG interruptions, a significant portion of the variance in stability remains to be explained (Nolin et al., 2013). Furthermore, there may be at least two or more different mutational pathways causing repeat expansion associated with the FXS (Eichler et al., 1996).

Nearby cis-elements seem to play an important role in TNR expansion, as repeat instability takes place solely at the disease locus (Mangel et al., 1998; López Castel et al., 2010;

Rousseau et al., 2011). Premutation alleles without AGG interruptions are at a high risk for CGG repeat expansion. However, the risk of expansion to full mutation for a premutation mother with 55–69 repeats ranges only from 4 to 18%, implying that additional cis-elements that promote larger repeat expansions (Nolin et al., 2013) may be present. Ennis et al. (2007) identified a single-nucleotide polymorphism (SNP) variant T/C (ss71651738 or WEX70) 53 kb upstream of the CGG repeats. The SNP variant C cosegregates with a chromosome haplotype at the highest risk for repeat expansion and is located in a repetitive DNA sequence that is classified as an MRE1b (medium reiterative element 1B).

We determined whether the SNP overlaps with the replication origin upstream of the *FMR1* repeat (Gerhardt et al., 2014). First, we mapped the replication initiation sites upstream of the CGG repeats in detail and found that this SNP is located at the replication initiation site in nonaffected cells. In the FXS hESCs we examined, this replication initiation site is missing, and the SNP variant C replaces the T. We also examined hESC lines derived from embryos that contained a premutation allele. We found that the premutation hESC lines contained an active replication origin and the SNP variant T in contrast to the full mutation hESCs, which contain a C at the missing replication initiation site. This study proposes that the SNP variant C at the replication initiation site 53 kb upstream of the *FMR1* gene contributes to the silencing of this replication origin and variation in the replication program, which may promote repeat expansion to the full mutation in a subset of fragile X patients.

A

Cells	Disease status	Repeat size		SNP	AGG interruptions
H9	non-affected hESC	23,29	stable	T	1, 2
H14	non-affected hESC	30	stable	T	2
SI-214	Fragile X hESC	~450	expanded	C	1
WCMC37	Fragile X hESC	~450	expanded	C	none
GM00011	non-affected fibroblast	30	stable	T	2
GM07072	Fragile X fibroblast	~500	expanded	C	none

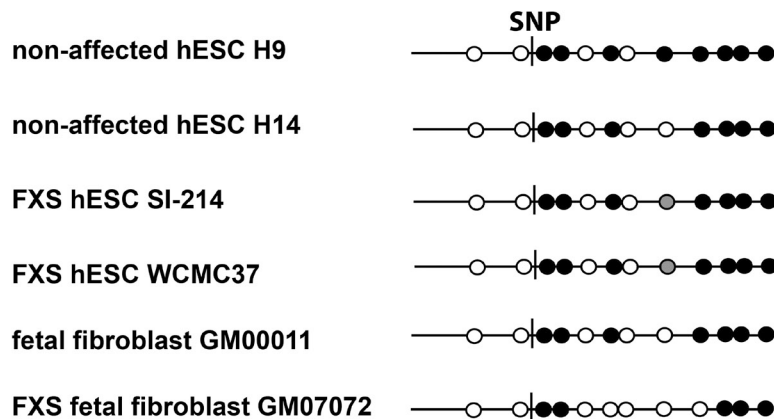**B**   **DNA methylation** ○ 0-10 %   ◉ 10-50 %   ● 50-100 %

Figure 2. **FXS hESCs contain the SNP variant C instead of T and a methylation pattern surrounding the SNP that is similar to the DNA methylation pattern in nonaffected hESCs.** (A) Table shows that DNA sequence analysis of 1-kb genomic DNA containing the SNP revealed that nonaffected hESCs H14 and H9 and fetal fibroblast GM00011 contain the SNP variant T and FXS hESCs SI-214 and WCMC37 and FXS fetal fibroblast GM07072 contain SNP variant C. The repeat sizes and AGG interruptions were determined by Asuragen PCR analysis. (B) DNA methylation pattern of the 367-bp genomic DNA segment containing the SNP was analyzed in nonaffected hESCs H14 and H9 and fetal fibroblast GM00011 and FXS hESCs SI-214 and WCMC37 and FXS fetal fibroblast GM07072. For each cell line, 10 DNA methylation analyses were performed at the genomic DNA segment including the SNP (containing 10 CpG sites and two CpA sites). A white circle indicates <10% CpG methylation, and gray circles indicate if <50% CpG methylation was detected. Filled black circles mark the positions of ≥50% methylated CpG sites.

Results and discussion

The replication initiation site ~50 kb upstream of the repeats overlaps with a previously reported SNP associated with CGG repeat expansion
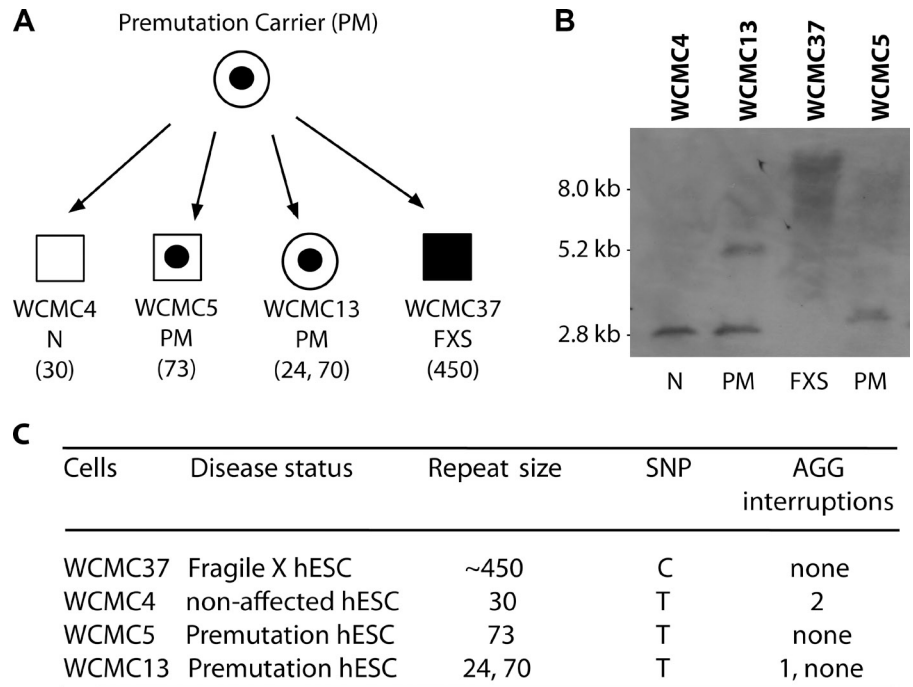
A T/C SNP (ss71651738) previously identified 53 kb upstream of the *FMR1* CGG repeats was linked to FXS patients in chromosome haplogroup D, a haplogroup at high risk of expansion (Table S1; Ennis et al., 2007). To determine whether the replication initiation site overlaps with this SNP in the MRE1b element, we mapped the replication origin upstream of the repeats in more detail using the nascent strand abundance assay (using the protocol of Liu et al., 2010). As a quality control of the isolated nascent strand DNA, we measured the enrichment of a known replication origin, Or6 (Fig. 1, B and D; Gerhardt et al., 2006). We designed six primer pairs at and surrounding the SNP ~50 kb upstream of the repeats (Fig. S1, A–C). In the nonaffected control hESC H14 and a fetal fibroblast line GM00011 (Fig. 1 A), we detected an enrichment of nascent strands at the SNP site in comparison to the surrounding DNA segments. In addition, we detected lower levels of nascent strands at the site where the C primer pair binds, probably as a result of a rarely used replication origin at this site. In FXS hESCs SI-214 and WCMC37, we detected no significant enrichment of nascent strand DNA at the SNP site in comparison to control primer C, indicating that an active replication initiation site was

missing in FXS hESCs (Fig. 1 C). Thus, the nascent strand abundance analysis revealed that the replication initiation sites ~50 kb upstream of the CGG repeats overlaps with the SNP in nonaffected cells.

FXS hESCs and FXS fetal fibroblasts contain the SNP variant C instead of T

We next examined whether the DNA sequence was altered at the replication initiation site. Therefore, we sequenced a 1-kb DNA segment containing the SNP and MRE1b element in FXS hESCs and control hESCs (Fig. 2 A). The analysis revealed that in FXS hESCs, SI-214 and WCMC37 and in FXS fetal fibroblast GM07072, the SNP variant C was found at the replication initiation site instead of the SNP variant T. We observed no other noteworthy alteration in the DNA sequences close to the SNP site 53 kb upstream of the repeats and the *FMR1* gene. In addition, we determined the number of AGG interruptions in the CGG repeats in nonaffected and FXS cells. Control cells H14, H9, and GM00011 contained two AGG interruptions, whereas FXS hESC SI-214 contained one, and FXS hESC WCMC37 and FXS fetal fibroblast GM07072 had none. These results are in agreement with previous analysis of AGG interruptions in cells from patients in whom loss of AGG interruptions increases the risk of CGG repeat expansion (Yrigollen et al., 2012; Nolin et al., 2013). Collectively, we found that the FXS cells studied here contain the SNP variant C.

Figure 3. **T/C single nucleotide change in FXS hESCs is linked to CGG repeat expansion.** (A and B) Analysis of four hESCs WCMC4, 5, 13, and 37 derived from embryos from premutation (open circle with filled black circle) mothers and nonaffected (N) fathers. Numbers in parentheses are the repeat numbers of each line. Southern blot analysis of the CGG repeat length of the four hESC lines (B). WCMC4 shows nonaffected hESCs (male, open box). WCMC5 (male, open box with filled black circle) and 13 (female, open circle with filled black circle) are premutation hESCs. WCMC37 is a full mutation FXS hESC (male, black box). (C) The table summarizes the repeat length, AGG interruptions (PCR Asuragen capillary electrophoreses analysis; Fig. S3), and SNP analysis by DNA sequencing of the genomic DNA of a 1-kb DNA segment containing the SNP.

Cells	Disease status	Repeat size	SNP	AGG interruptions
WCMC37	Fragile X hESC	~450	C	none
WCMC4	non-affected hESC	30	T	2
WCMC5	Premutation hESC	73	T	none
WCMC13	Premutation hESC	24, 70	T	1, none

The DNA methylation pattern is similar in nonaffected and FXS hESCs

The specific DNA sequences or epigenetic elements that define replication origins in mammalian cells are not known. In FXS hESCs, epigenetic changes such as DNA methylation could eliminate replication initiation sites upstream the *FMR1* gene because a more closed chromatin structure may prevent binding of replication initiation proteins and origin firing. One study has suggested that some replication origins are located in more lightly packed chromatin-like promoter regions (Hiratani et al., 2009). We compared the DNA methylation status of the CpG and CpA sites at and close to the SNP in FXS hESCs and nonaffected hESCs by bisulphite sequencing (Fig. 2 B). We observed no CpA methylation at the SNP variant C in FXS hESCs SI-214 and WCMC37 and FXS fetal fibroblast GM07072. Furthermore, we saw no major differences in the CpG methylation pattern surrounding the SNP and at the SNP in the nonaffected hESCs (H9 and H14), fetal fibroblast line GM00011, and FXS hESCs (SI-214 and WCMC37). The FXS fetal fibroblast line GM07072 had lower levels of DNA methylation than the hESC lines and nonaffected fetal fibroblasts. Although the *FMR1* gene is silenced in FXS fibroblasts because of DNA methylation at the CGG repeats, we detected less DNA methylation 53 kb upstream of the repeats than in a nonaffected fibroblast that contained an active *FMR1* gene. In summary, differences in the methylation patterns at the SNP do not appear to account for differences in origin activation.

Premutation hESCs contain the same SNP variant T as nonaffected hESCs

Because we did not detect a major difference in the DNA methylation pattern at the replication initiation site, we asked whether the T instead of a C at the SNP site in FXS cells was associated with repeat instability in FXS hESCs. To test this, we derived a few additional hESC lines from embryos of premutation carriers: WCMC4 (male normal), WCMC5 (male premutation), and WCMC13 (female premutation; Fig. 3 A and Fig. S2; Colak et al., 2014). We first examined the repeat size by Southern blotting (Fig. 3 B) and further confirmed our findings by the AmplideX PCR assay (Asuragen; Fig. S3). WCMC4 has a normal repeat size (30 repeats with two AGG interruptions; Fig. S3 A and Fig. 3 B), whereas WCMC5 and 13 contain the premutation repeat size with 73 and 70 CGG repeats, respectively, and have no AGG interruptions (Fig. S3, B and C; and Fig. 3 B). We determined the DNA sequence at and surrounding the SNP. Similar to the nonaffected cells, WCMC4, 5, and 13 contain the SNP variant T (Fig. 3 C) in contrast to the FXS hESC line WCMC37, which contains the C variant. Because of the privacy rights of the donors of human embryos, the haplotype structure of the premutation female parent cannot be established. In summary, we obtained two premutation hESCs, neither of which had AGG interruptions nor the SNP variant C in the MRE1b element.

Premutation hESCs contain a replication origin upstream of the repeats

Next, we asked whether the premutation hESCs, which contain a T, have a similar replication program to nonaffected hESCs. Nonaffected hESCs contain the SNP variant T and an active replication origin 53 kb upstream of the repeats (Gerhardt et al., 2014). This is in contrast to FXS hESCs, which contain a C and the replication origin is missing. First, we tested whether the CGG repeats in the premutation hESCs containing the SNP variant T expand further in cell culture. Therefore, we analyzed the repeat length in multiple passages of the premutation hESC WCMC5 by Southern blotting over the course of several months (Fig. 4 A). We observed no large changes in the repeat
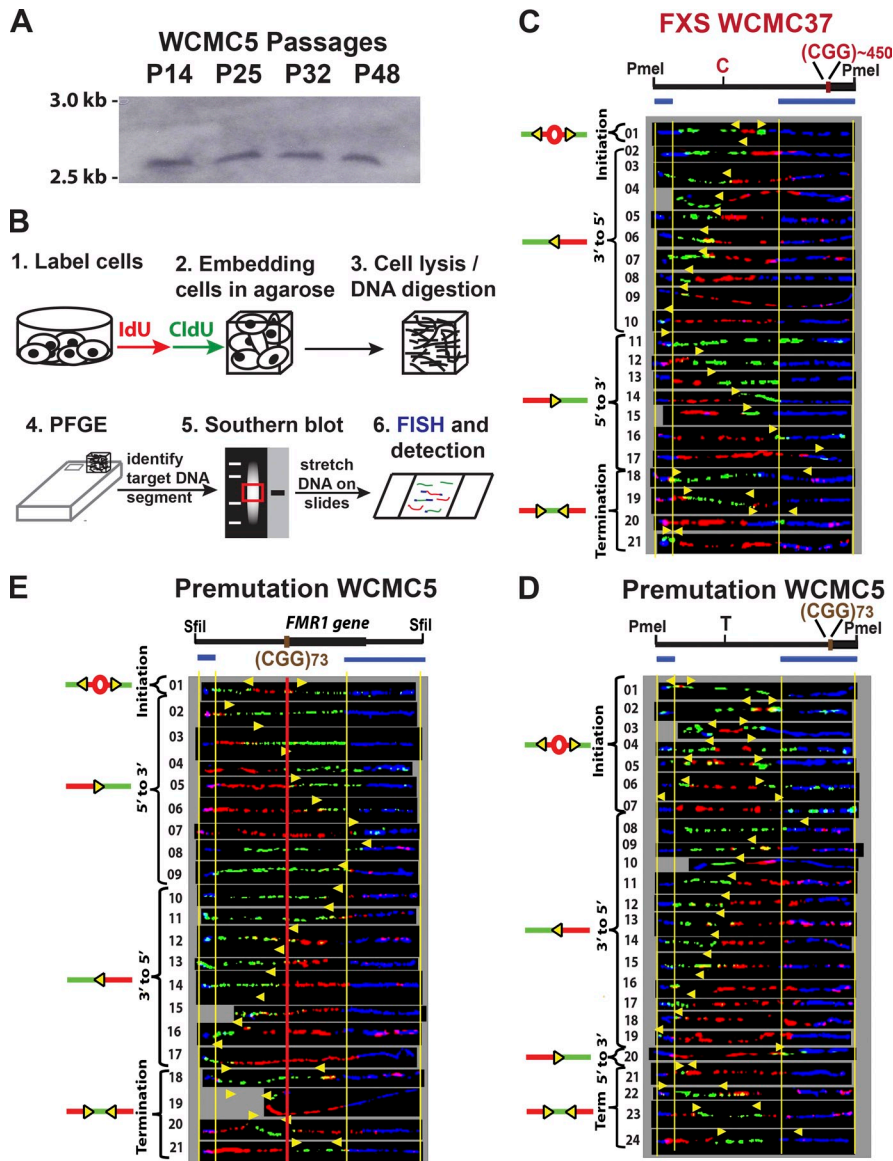
Figure 4. Premutation hESCs, which do not expand to full mutation, have a different replication program than FXS hESCs and a similar replication program to nonaffected hESCs. (A) The stability of the CGG repeats in premutation hESCs was analyzed by EcoRI and EagI Southern blotting. (B) Schematic of the steps of SMARD: First, cells are pulsed with IdU and CldU. Cells are then embedded in agarose and lysed. After digestion with a restriction enzyme, DNA molecules are separated by PFGE. Target segments were identified by Southern blotting and stretched on silanized glass slides. The replicated DNA molecules are detected by immunostaining and FISH. (C and D) The replication program was analyzed by SMARD in premutation and FXS hESCs as previously described (Gerhardt et al., 2014). (top) Map of the 102-kb PmeI segment containing the SNP variant T/C and the CGG repeats. The positions of the FISH probes are marked in blue and through the DNA molecules by yellow lines. (bottom) Photomicrographs of labeled DNA molecules from WCMC37 FXS hESCs (C) and WCMC5 premutation hESCs (D) are ordered according to replication fork (yellow arrowheads) progression; replication initiation, 5′→3′ and 3′→5′ directions, and termination are depicted. In contrast to FXS hESCs, the premutation hESCs WCMC5 contains a replication origin 53 kb upstream of the repeats. (E) The replication program was analyzed by SMARD at the *FMR1* locus containing the CGG repeats in premutation hESCs (120-kb SfiI segment). The CGG repeats are indicated by a red vertical line.

length in these cell passages. We also analyzed cell passages by Asuragen PCR analysis. All passages contained the same repeat length (73 repeats). We concluded that the repeats in the premutation hESCs WCMC5 do not expand further in cell culture.

To determine whether the replication program in the premutation hESCs is similar to nonaffected cells, we used single molecule analysis of replicated DNA (SMARD; Fig. 4 B). As a control, we determined the replication profile at this SNP site in FXS hESC line WCMC37 (Fig. 4 C). In contrast to the FXS hESC line WCMC37, we found replication initiation sites overlapping the SNP site in premutation hESC line WCMC5 similar to nonaffected hESCs (Fig. 1 A and Fig. 4 D; Gerhardt et al., 2014). We then analyzed the adjacent DNA segment containing the CGG repeats and the *FMR1* gene. In the premutation hESCs (Fig. 4 E), the replication fork progresses in both directions through the CGG repeats as a result of the active replication initiation site upstream of the repeats, as also seen in nonaffected hESCs (Gerhardt et al., 2014). These results, confirm that the SNP variant T is a mark for an active replication initiation site

53 kb upstream of the repeats and the replication profile seen in nonaffected cells. Hence, we propose that the replacement of the T with a C in the MRE1b element in FXS cells during early embryogenesis correlates with origin inactivation and an altered replication program at the *FMR1* locus, which could lead to CGG repeat expansion. We concluded that the replication initiation site 53 kb upstream of the repeats contains a cis-acting element, which could promote repeat expansion in FXS hESCs.

The cause for the altered replication program in FXS hESCs is an SNP variant 53 kb upstream of the repeats

To identify the cause of CGG repeat expansion at the *FMR1* locus and to identify nearby cis-elements, which could cause CGG repeat instability in FXS patients, we determined the DNA replication program at the endogenous *FMR1* locus in premutation and FXS hESCs. We found that in nonaffected cells, a replication initiation site is located ~50 kb upstream of the CGG repeats (Gerhardt et al., 2014). The replication initiation

site overlaps with an SNP (ss71651738) associated with CGG repeat expansion in FXS patients belonging to haplogroup D (Ennis et al., 2007). Furthermore, we found that in FXS hESCs with the SNP variant C, the replication initiation site is missing. This SNP variant C seems to contribute to CGG repeat instability in FXS patients in haplogroup D. Indeed, we found that in premutation hESCs, the presence of a T in contrast to a C in the MRE1b element was associated with replication origin activation and a replication program similar to nonaffected hESC containing stable repeats.

How can a single nucleotide substitution T to C lead to the inactivation of a replication origin? The T/C SNP could prevent the binding and proper function of proteins that bend DNA (Mysiak et al., 2004a,b) or trigger DNA looping (Snyder et al., 1986; Zahn and Blattner, 1987; Hsieh et al., 1991; Balani et al., 2010). Less flexible DNA could be more difficult to access for proteins, such as replication factors. For example, it is shown that cellular transcription factors, which bend the DNA, promote DNA replication probably by facilitating the binding of the preinitiation complex (Mysiak et al., 2004b). Another possibility might be that DNA looping could bring the replication origin into regions of high concentration of proteins, such as Cdc45, which could lead to origin firing (Knott et al., 2012). Failure in the binding of proteins that trigger DNA looping, such as forkhead transcription factors, could keep the DNA locus in a region of the nucleus with lower concentrations of replication initiation factors and prevent origin firing. A third model would be that the T/C SNP prevents the binding of chromatin remodeling complexes, such as ACF1–ISWI (Collins et al., 2002), which shift the position of nucleosomes. Nucleosomes could occupy the replication origin and affect the binding of replication initiation factors (Eaton et al., 2010; Lubelsky et al., 2011; Papior et al., 2012). These models have to be tested further experimentally.

Additional factors involved in CGG repeat expansion in FXS patients

In addition to the altered DNA replication program, errors in DNA repair could be part of the mechanism leading to repeat instability in FXS patients. Failures to resolve stalled replication forks and to repair secondary repeat structures at the *FMR1* locus could be an important aspect (Kovtun et al., 2007) of the mechanism leading to repeat expansion. Furthermore, varying levels of DNA repair capacity and expression of checkpoint proteins could contribute to repeat instability (Entezam and Usdin, 2009; Shishkin et al., 2009; Voineagu et al., 2009; Du et al., 2012, 2013; Zhang et al., 2012). An alternative possibility could be that other unknown genetic characteristics of the haplotype D together with the SNP variant C at the replication origin result in origin inactivation and repeat instability.

Several cis-elements and trans-acting factors contribute to repeat expansion in TNR diseases. We describe, in our observations, a new cis-element, which implicates in CGG repeat expansion and seems to alter the DNA replication program in a subset of fragile X families. Because large expansions leading to the FXS full mutation are seen after maternal transmission, the SNP variant C may already be present during germ

cell development in the premutation mother, and this further expansion may have taken place during embryogenesis. To summarize, we have shown that the SNP variant C in FXS hESCs appears to result in the inactivation of a DNA replication origin and could be responsible for CGG repeat expansion to the fragile X full mutation in a subset of FXS patients.

Materials and methods

Cell culture

Embryos from premutation mothers were cultured to the blastocyst stage. hESC lines WCMC4, 5, 13, and 37 (Colak et al., 2014; Gerhardt et al., 2014) were derived by laser inner cell mass dissection of the blastocyst on day 6. As described previously, isolated clumps of inner cell mass cells were plated on primary mouse embryonic fibroblasts. Outgrowth-containing cells were manually cut and propagated, resulting in a stable culture of undifferentiated hESCs. The WCMC5 line was fully characterized by stem cell markers in vitro and teratoma formation in vivo. The use of spare in vitro fertilization–derived embryos that have been diagnosed as genetically affected for the generation of hESCs was approved by Weill Cornell Medical College Institutional Review Board (protocol no. 0502007737).

H9 (WA09), H14 (WA14; Thomson et al., 1998), SI-214, WCMC4, WCMC5, WCMC13, and WCMC37 hESCs were grown on primary mouse embryonic fibroblasts plated at a density of 11,500–13,500 cells/cm^2 (GlobalStem, Inc.). ESCs were fed daily with hESC media composed of DMEM/F12 (Gibco), 20% knockout serum replacement (Gibco), 3.5 mM glutamine (Life Technologies), 0.1 mM MEM nonessential amino acids (Gibco), 55 mM 2-mercaptoethanol (Life Technologies), and 6 ng/ml FGF2 (fibroblast growth factor 2; R&D Systems). In the last passage before the experiments, hESCs were grown on Matrigel (BD) in conditioned medium. For production of conditioned medium, mouse embryonic fibroblasts were plated at 50,000 cells/cm^2 in DMEM with 10% FBS overnight. The next day, this medium was removed, and the cells were washed once with PBS. Then, hESC media were placed on the cells overnight for conditioning. The next day, the medium was removed, and 10 ng/ml FGF2 was added to conditioned medium before use. ESCs were dissociated with 1 mg/ml Dispase (Worthington Biochemical Corporation) at 37°C before transfer onto Matrigel. ESCs were dissociated into a single-cell suspension with Accutase for 40 min at 37°C (Innovative Cell Technologies) and counted before each experiment.

Human fetal fibroblasts GM00011 and GM07072 (Coriell Cell Repositories) were grown in MEM with Earle's salts, nonessential amino acids, and 15% fetal bovine serum (Invitrogen/Gibco). Fibroblasts were lifted using trypsin and passaged every 2–3 d.

Teratoma formation and histology

To evaluate the pluripotent nature of the hESCs, it is necessary to perform an in vivo study of teratoma formation. hESCs were injected into nonobese diabetic, severe combined immunodeficient mice that are known to form teratomas rather than reject human cells. To enhance the teratoma formation, approximately one to five million hESCs cells were injected together with Matrigel (1:1, vol/vol; total volume of 0.5 ml). The WMC4, -5, and -13 hESCs were injected into the dorsal back area of the nonobese diabetic, severe combined immunodeficient mice. This area was chosen as less inconvenient because the skin is more mobile and doesn't interfere with the growth of the tumor as much as in testis in which tissue growth is constrained by the capsule. After approximately 2 mo, the mouse was sacrificed, and the teratoma was processed for histological analysis.

Teratoma tissue was carefully dissected from the host mouse tissue, washed, and fixed in 4% PFA overnight. The next day, the teratoma was extensively washed from the PFA and stored in 30% ethanol. After this preparation, the teratomas were shipped in ethanol to the reference histology laboratory (Histoserv). The standard paraffin embedding and sectioning were performed followed by hematoxylin and eosin staining. Tissue sections were mounted on slides and shipped back for analysis. Histological sections were analyzed using a compound microscope (Axio Observer correlative light and electron microscopy; Carl Zeiss), and images were captured with a digital imaging system (Photoshop; Adobe). The slides were sent to a pathologist at New York Presbyterian Hospital to confirm the observations.

Flow cytometry

The SSEA-3 antibody (conjugated to Alexa Fluor 647; 561145; BD), SSEA-4 antibody (conjugated to Alexa Fluor 647; 560796; BD), and Tra-1-60

antibody (Alexa Fluor 647; 560120; BD) were used to quantify the percentage of pluripotency markers. Cells were lifted from the dish for 40 min with Accutase, washed in PBS, and then counted. Half a million cells were resuspended in PBS containing the aforementioned antibodies. Cells were incubated with antibodies on ice for 30 min, washed three times with PBS, and then analyzed on a cell sorter (FACSAria; BD).

DNA methylation analysis
0.5 µg genomic DNA was subjected to bisulfite treatment according to the instructions of the EZ DNA Methylation kit (Zymo Research). Converted DNA was amplified by PCR using a PCR kit (Advantage-GC 2; Takara Bio Inc.) with bisulfite conversion–based methylation PCR primers forward, 5′-TTATATAGTAGGAGGTGAGTGGTGG-3′, and reverse, 5′-CTTTATTA-AATCAAATAACACTTCCC-3′. After purification, the PCR product was cloned with the TOPO TA Cloning kit (Invitrogen) into the TOPO vector and transformed into DH5-α bacteria. For each cell line, 10 colonies were picked, and the plasmid DNA was purified and sequenced with M13 primers by the Albert Einstein College of Medicine Genomics Core Facility.

Southern blot
DNA was isolated by phenol extraction. 50 U EagI and 100 U EcoRI were used to digest 12 µg DNA for 4 h at 37°C. Digested genomic DNA was resolved on a 1% agarose gel without ethidium bromide (Nolin et al., 2003, 2008). Enzyme-restricted DNA was blotted onto a membrane (Hybond-XL; GE Healthcare) overnight. The membrane was hybridized overnight to a PCR-generated probe, using primers forward, 5′-GCTAGCAGGGCT-GAAGAGAA-3′, and reverse, 5′-CAGTGGAGCTCTCCGAAGTC-3′ (PCR product: 595 bp), which was labeled with ^{32}[P]CTP by RadPrime DNA Labeling System (Invitrogen). The membrane was washed twice for 5 min with 2× SSC and once for 8 min at 65°C with 2× SSC/0.5% SDS. The membrane in Saran Wrap was exposed to film (BioMax XAR; Kodak) at −80°C for 4–7 d.

Sequencing
We amplified the DNA by PCR using primers SNP1F, 5′-TTTAATTGGCT-CACGGTTCC-3′, and SNP1R, 5′-GCAACTTCAGGCTTGCTACC-3′, by Taq PCR Master Mix kit (QIAGEN). After purification, the PCR product was sequenced using either the SNP1F or SNP1R primer by the Albert Einstein College of Medicine Genomics Core Facility.

SMARD
The cells were grown at 37°C for 4 h in the presence of 25 µM 5-iodo-2′-deoxyuridine (IdU; MP Biomedicals). After washing cells with PBS, ESC medium with 25 µM 5-chloro-2′-deoxyuridine (CldU; Thermo Fisher Scientific) was added to the cultures, and the cells were incubated for an additional 4 h. The cells were lifted with Accutase or trypsin. After centrifugation, the cells were resuspended at 3 × 10⁷ cells/ml in PBS. Melted 1% InCert agarose (Lonza) in PBS was added to an equal volume of cells at 42°C. The cell suspension was pipetted into a chilled plastic mold with 0.5 × 0.2–cm wells with a depth of 0.9 cm for preparing DNA gel plugs. The gel plugs were allowed to solidify on ice for 30 min. Cells in the plug were lysed in buffer containing 1% n-lauroylsarcosine (Sigma-Aldrich), 0.5 M EDTA, and 20 mg/ml proteinase K (Roche). The gel plugs remained at 50°C for 64 h and were treated with 20 mg/ml proteinase K every 24 h. Gel plugs were then rinsed several times with TE (Tris-EDTA) and once with phenylmethanesulfonyl fluoride (Sigma-Aldrich). The plugs were washed with 10 mM MgCl₂ and 10 mM Tris-HCl, pH 8.0. The genomic DNA in the gel plugs was digested with 80 U SfiI (New England BioLabs, Inc.) at 50°C overnight. The digested gel plugs were rinsed with TE and cast into a 0.7% SeaPlaque GTG agarose gel (Lonza). A gel λ ladder pulsed-field gel electrophoresis (PFGE) marker and yeast chromosome PFGE marker (both obtained from New England BioLabs, Inc.) were cast next to the gel plugs. A Southern blot was performed to determine the location of the DNA fragment on the gel. The region of the gel containing the segment of interest was excised and set aside, whereas the rest of the DNA (which includes the chromosome ladders) was transferred to a membrane (Hybond-XL) and hybridized with a probe located near the CGG repeats (see Southern blot). Autoradiography was used to determine the location of the appropriate DNA segment. Gel slices from the appropriate positions in the PFGE were cut and melted at 72°C for 20 min. GELase enzyme (Epicentre Biotechnologies; 1 U/50 µl agarose suspension) was carefully added to digest the agarose and incubated at 45°C for a minimum of 2 h. The resulting DNA solutions were stretched on 3-aminopropyltriethoxysilane (Sigma-Aldrich)–coated glass slides. The DNA was pipetted along one side of a coverslip that had been placed on top of a silane-treated glass slide and allowed to enter by capillary action. The DNA was denatured with sodium hydroxide in ethanol and then fixed with glutaraldehyde.

The slides were hybridized overnight with a biotinylated probe (Fig. 4, C–E, the blue bars diagrammed on the maps indicate the positions of the probes used). The next day, the slides were rinsed in 2× SSC (SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 1% SDS and washed in 40% formamide solution containing 2× SSC at 45°C for 5 min and rinsed in 2× SSC–0.1% IGEPAL CA-630. After several detergent rinses (four times in 4× SSC–0.1% IGEPAL CA-630), the slides were blocked with 1% BSA for ≥20 min and treated with Avidin Alexa Fluor 350 (Invitrogen/Molecular Probes) for 20 min. The slides were rinsed with PBS containing 0.03% IGEPAL CA-630, treated with biotinylated anti–Avidin D (Vector Laboratories) for 20 min, and rinsed again. The slides were then treated with Avidin Alexa Fluor 350 for 20 min and rinsed again, as in the previous step. The slides were incubated with the IdU antibody, a mouse anti-BrdU (BD), the antibody specific for CldU, a monoclonal rat anti-BrdU (Accurate Chemical and Scientific Corporation), and biotinylated anti–Avidin D for 1 h. This was followed by incubation with Avidin Alexa Fluor 350 and secondary antibodies Alexa Fluor 568 goat anti–mouse IgG (H+L [heavy + light]; Invitrogen/Molecular Probes) and Alexa Fluor 488 goat anti–rat IgG (H+L; Invitrogen/Molecular Probes) for 1 h. After a final PBS/IGEPAL CA-630 rinse, the coverslips were mounted with gold antifade reagent (ProLong; Invitrogen). A fluorescent microscope (Axioscop 2 MOT plus with Plan Apochromat 63×/1.4 NA oil differential interference contrast objective; Carl Zeiss) with a camera (CoolSNAP HQ; Photometrics) and IPlab 4.0.8 software (BD) was used to detect the nucleoside incorporation into the DNA molecules. Images were processed with Photoshop CS5 software. Fosmids (human genome GRCh37/hg19) used were G248P81601D11, G248P87940A10, G248P80577A2, and G248P87545H3.

Nascent strand abundance analysis
Newly synthesized DNA was isolated from hESCs and fibroblasts as previously described (Liu et al., 2010). In brief, ∼8 × 10⁷ cells were harvested, washed with PBS, and loaded on a 1.25% alkaline agarose gel. After a 20-min incubation (cells were lysed by the alkaline buffer in the well), the DNA was separated by gel electrophoresis for 17 h at 35 V. Single-stranded DNA of 1–1.3 kb was cut out of the gel, and the nascent DNA was isolated using the gel extraction kit (QIAquick Gel Extraction Kit; QIAGEN). The DNA was quantified by real-time PCR. All primers used are provided in Fig. S1 A. PCR products of standard curves are provided in Fig. S1 B.

SNP (WEX70) analysis
Genomic DNA was amplified by PCR using primers forward, 5′-AGACT-GCGAGATGGGAGAAG-3′, and reverse, 5′-GGTAGAGACGCAGAGC-CAAG-3′, and the HotStar Taq Master Mix kit (QIAGEN). The PCR product was digested with AleI and analyzed by gel electrophoresis on a 2% agarose gel. For the SNP variant C, DNA digestion results in two bands of 279 and 129 bp.

Online supplemental material
Fig. S1, related to Fig. 1, lists the primer sequences and shows the primer analysis from primers used for nascent strand analysis by real-time PCR as well as lists of the p-value. Fig. S2 is related to Fig. 3 and shows the characterization of the human control, premutation, and FXS hESCs examined in this study. Fig. S3 is also related to Fig. 3 and shows the analysis of the CGG repeat length using the AmplideX PCR assay and capillary electrophoresis. Table S1 shows the analysis of the SNP variant T/C (WEX70) in individuals belonging to haplogroup D. Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.201404157/DC1.

References

Balani, V.A., Q.A. de Lima Neto, K.I. Takeda, F. Gimenes, A. Fiorini, M. Debatisse, and M.A. Fernandez. 2010. Replication origins oriGNAI3 and oriB of the mammalian AMPD2 locus nested in a region of straight DNA flanked by intrinsically bent DNA sites. *BMB Rep.* 43:744–749. http://dx.doi.org/10.5483/BMBRep.2010.43.11.744

Colak, D., N. Zaninovic, M.S. Cohen, Z. Rosenwaks, W.Y. Yang, J. Gerhardt, M.D. Disney, and S.R. Jaffrey. 2014. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science.* 343:1002–1005. http://dx.doi.org/10.1126/science.1245831

Collins, N., R.A. Poot, I. Kukimoto, C. García-Jiménez, G. Dellaire, and P.D. Varga-Weisz. 2002. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat. Genet.* 32:627–632. http://dx.doi.org/10.1038/ng1046

De Rubeis, S., and C. Bagni. 2011. Regulation of molecular pathways in the Fragile X Syndrome: insights into Autism Spectrum Disorders. *J. Neurodev. Disord.* 3:257–269. http://dx.doi.org/10.1007/s11689-011-9087-2

Dobkin, C.S., S.L. Nolin, I. Cohen, V. Sudhalter, M.G. Bialer, X.H. Ding, E.C. Jenkins, N. Zhong, and W.T. Brown. 1996. Tissue differences in fragile X mosaics: mosaicism in blood cells may differ greatly from skin. *Am. J. Med. Genet.* 64:296–301. http://dx.doi.org/10.1002/(SICI)1096-8628(19960809)64:2<296::AID-AJMG13>3.0.CO;2-A

Du, J., E. Campau, E. Soragni, S. Ku, J.W. Puckett, P.B. Dervan, and J.M. Gottesfeld. 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. http://dx.doi.org/10.1074/jbc.M112.391961

Du, J., E. Campau, E. Soragni, C. Jespersen, and J.M. Gottesfeld. 2013. Length-dependent CTG·CAG triplet-repeat expansion in myotonic dystrophy patient-derived induced pluripotent stem cells. *Hum. Mol. Genet.* 22:5276–5287. http://dx.doi.org/10.1093/hmg/ddt386

Eaton, M.L., K. Galani, S. Kang, S.P. Bell, and D.M. MacAlpine. 2010. Conserved nucleosome positioning defines replication origins. *Genes Dev.* 24:748–753. http://dx.doi.org/10.1101/gad.1913210

Eichler, E.E., J.J. Holden, B.W. Popovich, A.L. Reiss, K. Snow, S.N. Thibodeau, C.S. Richards, P.A. Ward, and D.L. Nelson. 1994. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat. Genet.* 8:88–94. http://dx.doi.org/10.1038/ng0994-88

Eichler, E.E., J.N. Macpherson, A. Murray, P.A. Jacobs, A. Chakravarti, and D.L. Nelson. 1996. Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum. Mol. Genet.* 5:319–330. http://dx.doi.org/10.1093/hmg/5.3.319

Ennis, S., A. Murray, G. Brightwell, N.E. Morton, and P.A. Jacobs. 2007. Closely linked cis-acting modifier of expansion of the CGG repeat in high risk FMR1 haplotypes. *Hum. Mutat.* 28:1216–1224. http://dx.doi.org/10.1002/humu.20600

Entezam, A., and K. Usdin. 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. http://dx.doi.org/10.1093/nar/gkp666

Fu, Y.H., D.P. Kuhl, A. Pizzuti, M. Pieretti, J.S. Sutcliffe, S. Richards, A.J. Verkerk, J.J. Holden, R.G. Fenwick Jr., S.T. Warren, 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. http://dx.doi.org/10.1016/0092-8674(91)90283-5

Gerhardt, J., S. Jafar, M.P. Spindler, E. Ott, and A. Schepers. 2006. Identification of new human origins of DNA replication by an origin-trapping assay. *Mol. Cell. Biol.* 26:7731–7746. http://dx.doi.org/10.1128/MCB.01392-06

Gerhardt, J., M.J. Tomishima, N. Zaninovic, D. Colak, Z. Yan, Q. Zhan, Z. Rosenwaks, S.R. Jaffrey, and C.L. Schildkraut. 2014. The DNA replication program is altered at the FMR1 locus in fragile X embryonic stem cells. *Mol. Cell.* 53:19–31. http://dx.doi.org/10.1016/j.molcel.2013.10.029

Gray, S.J., J. Gerhardt, W. Doerfler, L.E. Small, and E. Fanning. 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. http://dx.doi.org/10.1128/MCB.01382-06

Hagerman, P.J., and R.J. Hagerman. 2004. The fragile-X premutation: a maturing perspective. *Am. J. Hum. Genet.* 74:805–816. http://dx.doi.org/10.1086/386296

Heitz, D., F. Rousseau, D. Devys, S. Saccone, H. Abderrahim, D. Le Paslier, D. Cohen, A. Vincent, D. Toniolo, G. Della Valle, et al. 1991. Isolation of sequences that span the fragile X and identification of a fragile X–related CpG island. *Science.* 251:1236–1239. http://dx.doi.org/10.1126/science.2006411

Hiratani, I., S. Takebayashi, J. Lu, and D.M. Gilbert. 2009. Replication timing and transcriptional control: beyond cause and effect—part II. *Curr. Opin. Genet. Dev.* 19:142–149. http://dx.doi.org/10.1016/j.gde.2009.02.002

Hsieh, C.H., M. Wu, and J.M. Yang. 1991. The sequence-directed bent DNA detected in the replication origin of *Chlamydomonas reinhardtii* chloroplast DNA is important for the replication function. *Mol. Gen. Genet.* 225:25–32. http://dx.doi.org/10.1007/BF00282638

Kim, J.C., and S.M. Mirkin. 2013. The balancing act of DNA repeat expansions. *Curr. Opin. Genet. Dev.* 23:280–288. http://dx.doi.org/10.1016/j.gde.2013.04.009

Knott, S.R., J.M. Peace, A.Z. Ostrow, Y. Gan, A.E. Rex, C.J. Viggiani, S. Tavaré, and O.M. Aparicio. 2012. Forkhead transcription factors establish origin timing and long-range clustering in *S. cerevisiae. Cell.* 148:99–111. http://dx.doi.org/10.1016/j.cell.2011.12.012

Kovtun, I.V., Y. Liu, M. Bjoras, A. Klungland, S.H. Wilson, and C.T. McMurray. 2007. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature.* 447:447–452. http://dx.doi.org/10.1038/nature05778

Liu, G., X. Chen, J.J. Bissler, R.R. Sinden, and M. Leffak. 2010. Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. *Nat. Chem. Biol.* 6:652–659.

López Castel, A., J.D. Cleary, and C.E. Pearson. 2010. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* 11:165–170. http://dx.doi.org/10.1038/nrm2854

Lubelsky, Y., T. Sasaki, M.A. Kuipers, I. Lucas, M.M. Le Beau, S. Carignon, M. Debatisse, J.A. Prinz, J.H. Dennis, and D.M. Gilbert. 2011. Pre-replication complex proteins assemble at regions of low nucleosome occupancy within the Chinese hamster dihydrofolate reductase initiation zone. *Nucleic Acids Res.* 39:3141–3155. http://dx.doi.org/10.1093/nar/gkq1276

Mangel, L., T. Ternes, B. Schmitz, and W. Doerfler. 1998. New 5′-(CGG)n-3′ repeats in the human genome. *J. Biol. Chem.* 273:30466–30471. http://dx.doi.org/10.1074/jbc.273.46.30466

Méchali, M. 2010. Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat. Rev. Mol. Cell Biol.* 11:728–738. http://dx.doi.org/10.1038/nrm2976

Mirkin, S.M. 2007. Expandable DNA repeats and human disease. *Nature.* 447:932–940. http://dx.doi.org/10.1038/nature05977

Mysiak, M.E., M.H. Bleijenberg, C. Wyman, P.E. Holthuizen, and P.C. van der Vliet. 2004a. Bending of adenovirus origin DNA by nuclear factor I as shown by scanning force microscopy is required for optimal DNA replication. *J. Virol.* 78:1928–1935. http://dx.doi.org/10.1128/JVI.78.4.1928-1935.2004

Mysiak, M.E., C. Wyman, P.E. Holthuizen, and P.C. van der Vliet. 2004b. NFI and Oct-1 bend the Ad5 origin in the same direction leading to optimal DNA replication. *Nucleic Acids Res.* 32:6218–6225. http://dx.doi.org/10.1093/nar/gkh944

Nolin, S.L., A. Glicksman, G.E. Houck Jr., W.T. Brown, and C.S. Dobkin. 1994. Mosaicism in fragile X affected males. *Am. J. Med. Genet.* 51:509–512. http://dx.doi.org/10.1002/ajmg.1320510444

Nolin, S.L., C. Dobkin, and W.T. Brown. 2003. Molecular analysis of fragile X syndrome. *Curr. Protoc. Hum. Genet.* Chapter 9:Unit 9.5.

Nolin, S.L., X.H. Ding, G.E. Houck, W.T. Brown, and C. Dobkin. 2008. Fragile X full mutation alleles composed of few alleles: implications for CGG repeat expansion. *Am. J. Med. Genet. A.* 146A:60–65. http://dx.doi.org/10.1002/ajmg.a.32087

Nolin, S.L., A. Glicksman, X. Ding, N. Ersalesi, W.T. Brown, S.L. Sherman, and C. Dobkin. 2011. Fragile X analysis of 1112 prenatal samples from 1991 to 2010. *Prenat. Diagn.* 31:925–931. http://dx.doi.org/10.1002/pd.2815

Nolin, S.L., S. Sah, A. Glicksman, S.L. Sherman, E. Allen, E. Berry-Kravis, F. Tassone, C. Yrigollen, A. Cronister, M. Jodah, et al. 2013. Fragile X AGG analysis provides new risk predictions for 45-69 repeat alleles. *Am. J. Med. Genet. A.* 161A:771–778. http://dx.doi.org/10.1002/ajmg.a.35833

Papior, P., J.M. Arteaga-Salas, T. Günther, A. Grundhoff, and A. Schepers. 2012. Open chromatin structures regulate the efficiencies of pre-RC formation and replication initiation in Epstein-Barr virus. *J. Cell Biol.* 198:509–528. http://dx.doi.org/10.1083/jcb.201109105

Pearson, C.E., E.E. Eichler, D. Lorenzetti, S.F. Kramer, H.Y. Zoghbi, D.L. Nelson, and R.R. Sinden. 1998. Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. *Biochemistry.* 37:2701–2708. http://dx.doi.org/10.1021/bi972546c

Pearson, C.E., K. Nichol Edamura, and J.D. Cleary. 2005. Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.* 6:729–742. http://dx.doi.org/10.1038/nrg1689

Pieretti, M., F.P. Zhang, Y.H. Fu, S.T. Warren, B.A. Oostra, C.T. Caskey, and D.L. Nelson. 1991. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell.* 66:817–822. http://dx.doi.org/10.1016/0092-8674(91)90125-I

Reyniers, E., L. Vits, K. De Boulle, B. Van Roy, D. Van Velzen, E. de Graaff, A.J. Verkerk, H.Z. Jorens, J.K. Darby, B. Oostra, et al. 1993. The full

mutation in the FMR-1 gene of male fragile X patients is absent in their sperm. *Nat. Genet.* 4:143–146. http://dx.doi.org/10.1038/ng0693-143

Reyniers, E., J.J. Martin, P. Cras, E. Van Marck, I. Handig, H.Z. Jorens, B.A. Oostra, R.F. Kooy, and P.J. Willems. 1999. Postmortem examination of two fragile X brothers with an FMR1 full mutation. *Am. J. Med. Genet.* 84:245–249. http://dx.doi.org/10.1002/(SICI)1096-8628 (19990528)84:3<245::AID-AJMG16>3.0.CO;2-U

Rousseau, F., Y. Labelle, J. Bussières, and C. Lindsay. 2011. The fragile x mental retardation syndrome 20 years after the FMR1 gene discovery: an expanding universe of knowledge. *Clin. Biochem. Rev.* 32:135–162.

Santoro, M.R., S.M. Bray, and S.T. Warren. 2012. Molecular mechanisms of fragile X syndrome: a twenty-year perspective. *Annu. Rev. Pathol.* 7:219–245. http://dx.doi.org/10.1146/annurev-pathol-011811-132457

Sherman, S.L. 2000. Premature ovarian failure in the fragile X syndrome. *Am. J. Med. Genet.* 97:189–194. http://dx.doi.org/10.1002/1096-8628(200023)97:3<189::AID-AJMG1036>3.0.CO;2-J

Shishkin, A.A., I. Voineagu, R. Matera, N. Cherng, B.T. Chernet, M.M. Krasilnikova, V. Narayanan, K.S. Lobachev, and S.M. Mirkin. 2009. Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol. Cell.* 35:82–92. http://dx.doi.org/10.1016/j.molcel.2009.06.017

Snyder, M., A.R. Buchman, and R.W. Davis. 1986. Bent DNA at a yeast autonomously replicating sequence. *Nature.* 324:87–89. http://dx.doi.org/10.1038/324087a0

Sullivan, A.K., M. Marcus, M.P. Epstein, E.G. Allen, A.E. Anido, J.J. Paquin, M. Yadav-Shah, and S.L. Sherman. 2005. Association of FMR1 repeat size with ovarian dysfunction. *Hum. Reprod.* 20:402–412. http://dx.doi.org/10.1093/humrep/deh635

Sutcliffe, J.S., D.L. Nelson, F. Zhang, M. Pieretti, C.T. Caskey, D. Saxe, and S.T. Warren. 1992. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* 1:397–400. http://dx.doi.org/10.1093/hmg/1.6.397

Thomson, J.A., J. Itskovitz-Eldor, S.S. Shapiro, M.A. Waknitz, J.J. Swiergiel, V.S. Marshall, and J.M. Jones. 1998. Embryonic stem cell lines derived from human blastocysts. *Science.* 282:1145–1147. http://dx.doi.org/10.1126/science.282.5391.1145

Verkerk, A.J., M. Pieretti, J.S. Sutcliffe, Y.H. Fu, D.P. Kuhl, A. Pizzuti, O. Reiner, S. Richards, M.F. Victoria, F.P. Zhang, 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. http://dx.doi.org/10.1016/0092-8674(91)90397-H

Voineagu, I., C.H. Freudenreich, and S.M. Mirkin. 2009. Checkpoint responses to unusual structures formed by DNA repeats. *Mol. Carcinog.* 48:309–318. http://dx.doi.org/10.1002/mc.20512

Willemsen, R., J. Levenga, and B.A. Oostra. 2011. CGG repeat in the FMR1 gene: size matters. *Clin. Genet.* 80:214–225. http://dx.doi.org/10.1111/j.1399-0004.2011.01723.x

Wöhrle, D., I. Hennig, W. Vogel, and P. Steinbach. 1993. Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion. *Nat. Genet.* 4:140–142. http://dx.doi.org/10.1038/ng0693-140

Yrigollen, C.M., B. Durbin-Johnson, L. Gane, D.L. Nelson, R. Hagerman, P.J. Hagerman, and F. Tassone. 2012. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. *Genet. Med.* 14:729–736. http://dx.doi.org/10.1038/gim.2012.34

Zahn, K., and F.R. Blattner. 1987. Direct evidence for DNA bending at the lambda replication origin. *Science.* 236:416–422. http://dx.doi.org/10.1126/science.2951850

Zhang, Y., A.A. Shishkin, Y. Nishida, D. Marcinkowski-Desmond, N. Saini, K.V. Volkov, S.M. Mirkin, and K.S. Lobachev. 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. http://dx.doi.org/10.1016/j.molcel.2012.08.002