FUNCTIONAL COMPONENTS OF A MAMMALIAN ORIGIN OF DNA REPLICATION AND THE POSSIBLE ROLE OF AN ORIGIN IN HUMAN FRAGILE X MENTAL RETARDATION

By

Steven James Gray

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Molecular Biology

December, 2006

Nashville, Tennessee

Approved:

Ellen Fanning

Wallace Lestourgeon

James Patton

Katherine Friedman

Terrence Dermody

ACKNOWLEDGEMENTS

There are many people who helped with the work presented in this document, and with my personal and professional growth. First, I'd like to thank my mentor and Ph.D. advisor, Ellen Fanning. She has always pushed me to be the best scientist that I could become, and served as an excellent role model of an outstanding scholar. I'd also like to thank my graduate committee members, Wally Lestourgeon, Kathy Friedman, Jim Patton, and Terry Dermody for their support and guidance. I'd like to thank the experimental and collaborative contributions of Jeannine Gerhardt, Amy Altman, Guoqi Liu, Lawrence Small, and Walter Doerfler, who helped to generate some of the data presented in this document. I thank the members of the Fanning lab for stimulating discussions (scientific and otherwise), friendship, and help.

I would like to thank M. Leffak for plasmids and the Hela 406 cell line, N. Heintz for the pMCD plasmid, and L.J. Zwiebel for LightCycler access. I thank F. Grummt for the pUC-SB2+ plasmid and J. Hayes for the pXP10 plasmid. I thank C. Pearson and J. Cleary for stimulating discussion in the early phase of this project. I also thank G. Liu, D. Schaarschmidt, and R. Knippers for advice and reagents, X. Qin for help optimizing PCR conditions, and L.J. Zwiebel for LightCycler access. I acknowledge the Coriell Institute for Medical Research, T. R. Peters, P. Wright, and M. Ikizler for providing cells used in these studies.

The financial support of the NIH (GM 52948 and training grant 5 T32 CA09385-20), the Howard Hughes Medical Institute, and Vanderbilt University are greatly appreciated.

A major component of my professional growth came from my role as a continuing mentor in the Howard Hughes Medical Institute Community of Scholars program at Vanderbilt University, under the direction of Ellen Fanning and Kathy Friedman. This outstanding program paired graduate students or post-doctoral researchers with undergraduate students to participate in collaborative, hands-on research. Through this program, I had the opportunity to work with several amazing and bright undergraduate students: Michael Gleason, Lawrence Small, Geoff Todd, Prabal Tiwaree, and Xan Paxton.

I wish to thank my closest friends who have helped me immensely to achieve this point in my life. Joel and Sara Schwartz became like a brother and sister to me, graciously offering help and support but asking for nothing in return. Vitaly Klimovich worked next to me for my entire time in the lab, and he was a great friend, colleague and role model. Eric and Sabrina Warren were always there when I needed help or a nice diversion from work, and I thank them for stimulating (usually not scientific) discussions.

My parents raised me to be what I am today, and I could not be where I am without them. Marilyn and Eddie Whitmire. Bernie and Robin Gray. I love each of them, and I will always be grateful for what they have done for me. I think they have done more than they will ever realize or acknowledge. I thank my brothers, Chris, Kyle, and Matt, and my sister-in-law Nikki, for doing the things that siblings do and for being more than just family.

Finally, I want to thank my wife Christy, and my daughter Aubrey. Christy has been the best thing that ever happened to me, and I don’t know how I would survive without her. She gave me the most wonderful daughter in the world. Aubrey doesn’t

know it yet, but she did more to keep me sane this last year than anything else. No matter how bad of a day I could have, a smile from her and “Daddy” made everything better. She kept life in perspective, and she is keeping me young when everything else is pushing me to grow up. She is my greatest accomplishment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Adenosine
ACS	ARS consensus sequence
ARS	Autonomously replicating sequence
bp	Base pair
°C	Degrees celsius
C	Cytidine/ cytosine
ChIP	Chromatin immunoprecipitation
CHO	Chinese hamster ovary
CO_2	Carbon dioxide
CTP	Cytosine 5'-triphosphate
DHFR	Di-hydrofolate reductase
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
DNR	Dinucleotide repeat
E. coli	*Escherichia coli*
FMR1	Fragile X Mental Retardation protein 1
FRAXA	Fragile X Mental Retardation, in relation to the FMR1 gene
G	Guanidine/ guanosine
GAF	GAGA factor
HD	Huntington disease
Hyg	Hygromycin

IP	Immunoprecipitation
IR	Initiation region
kb	Kilobase
kDa	Kilo Dalton
MCM	Mini-chromosome maintainance
μg	Mirogram
mL	Milliliter
mM	Millimolar
MNase	Micrococcal nuclease
mRNA	Messenger RNA
Neo	Neomycin
ng	Nanogram
nt	Nucleotide
NPE	Nucleosome positioning element
OBR	Origin of bidirectional replication
OIZ	Okazaki initiation zone
ORC	Origin recognition complex
ori	Origin
PCNA	Proliferating cell nuclear T antigen
PCR	Polymerase chain reaction
rDNA	Ribosomal DNA
RIP60	60 kDa replication initiation protein
RNA	Ribonucleic acid

SB2	Murine 28S rDNA Sal box 2
S. cerevisiae	*Saccharomyces cerevisiae*
SEM	Standard Error of the Mean
SV40	Simian Virus 40
T	Thymidine
TK	Thymidine kinase
TNR	Trinucleotide repeat
TTF-1	Transcription termination factor 1
WT	Wild-type

CHAPTER I

INTRODUCTION

In order for genetic information to be inherited from one generation to the next, the genome of the parent must be faithfully replicated and passed on to the progeny. Although DNA replication is one of the most fundamental aspects of life, our understanding of how it occurs and how it is regulated in eukaryotic cells is still far from complete. The most basic regulation of DNA replication occurs at the level of the DNA itself. The replicon model originally proposed by Jacob et al. in 1963 (83) (Figure 1) hypothesized that specific *cis*-acting elements (DNA) were acted upon by *trans*-acting elements (proteins) to initiate DNA replication. This model suggested that DNA sequences ultimately determined replication start sites, and that protein factors provided the triggering mechanism for the start of new DNA synthesis. The replicon model has since been refined in breadth and detail, but the basic concept has held true. Although many DNA sequences have been identified that are necessary and sufficient to serve as origins of DNA replication, it is still unclear how those DNA sequence elements are able to dictate if, when, and where DNA replication begins.

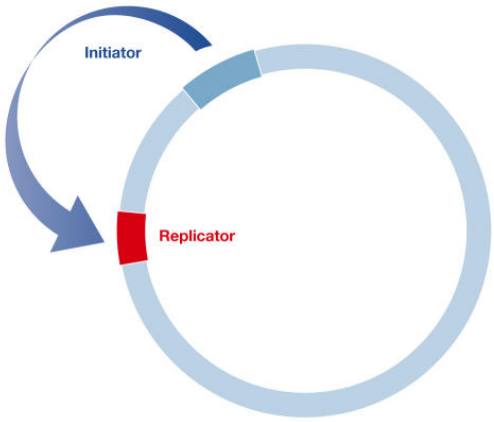


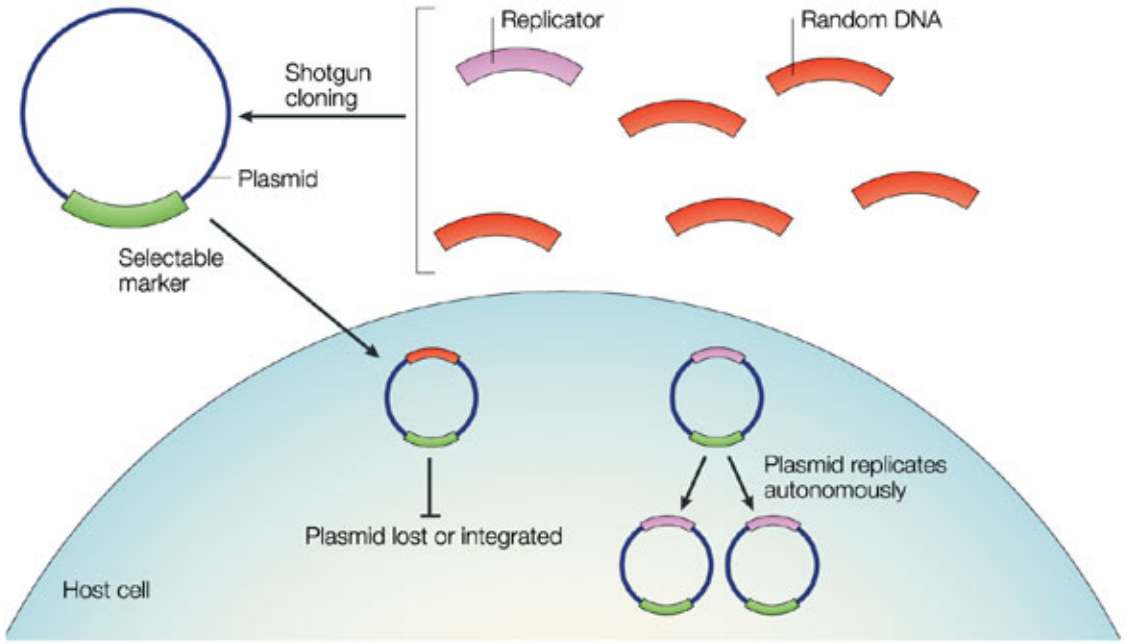
Figure 1. The replicon model proposed by Jacob et al. in 1963 (83). Reproduced from Aladjem and Fanning, 2004 (4).

What Is an Origin of DNA Replication?

An origin of DNA replication (origin) is a specific sequence that serves as a start site for DNA replication, and for the purpose of this dissertation, will also be defined to include the nearby DNA sequence elements necessary for DNA replication to start at that site. The replicon model (83) is most applicable in describing the relatively simple origin, *oriC*, found in *Escherichia coli*. In this example, the DNA sequence *OriC* is the replicator and the protein DnaA is the initiator (127). The 265 bp *E. coli OriC* is composed of five binding sites for DnaA, an AT-rich sequence, and binding sites for additional auxillary factors (87). Initiation occurs when helical distortions mediated by DnaA oligomerization melt the AT-rich tract, and the open "bubble" of single-stranded DNA can serve as a template for replication fork proteins to begin to synthesize new DNA (127).

Replication origins in budding yeast, the ARS1 model. In the budding yeast *Saccharomyces cerevisiae*, genome-wide mapping of replication origins by multiple techniques has provided a wealth of information about the characteristics of origins (115,

152, 179, 214). Origins in *S. cerevisiae* tend to be small (100-200 bp), and they are composed of multiple necessary sequence elements. It has been possible to genetically and biochemically dissect yeast origins through the use of an Autonomously Replicating Sequence (ARS) assay, in which an origin-containing DNA fragment is placed on a plasmid with a centromere (32, 180). The plasmid will replicate autonomously in yeast cells in a manner identical to the chromosomes as long as it contains a functional origin (Figure 2).



Nature Reviews | Molecular Cell Biology

Figure 2. The ARS assay. Taken from Gilbert, 2004, box1 (60). DNA fragments can be cloned onto a plasmid containing a selectable marker and transformed into yeast cells at low copies such that each cell should get no more than 1 plasmid. The transformed yeast are grown in the presence of selection, and surviving colonies contain plasmids with the ability to autonomously replicate. These plasmids are further analyzed for their ability to replicate autonomously, and in this manner ARS elements can be identified.

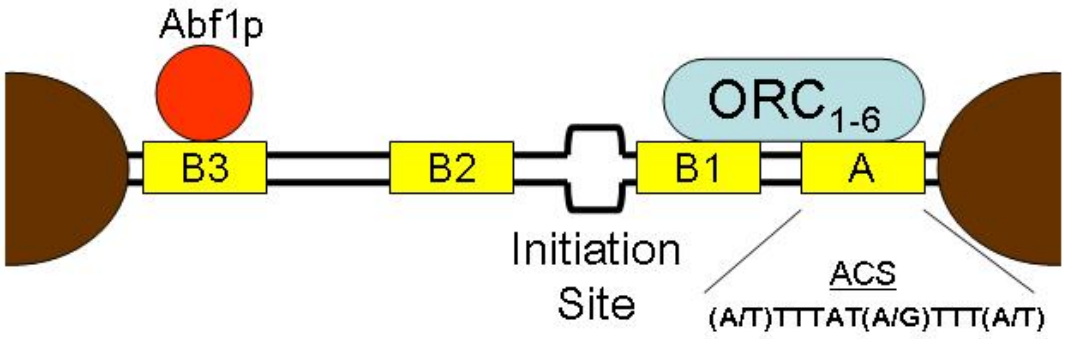


Figure 3. The *S. cerevisiae* ARS1. The ~125 bp origin sequence contains four necessary sequence elements (yellow boxes) with a single initiation site. The region is maintained free of nucleosomes as shown, with positioned nucleosomes flanking the origin (brown half circles). The ARS consensus sequence (ACS) for all budding yeast origins is shown.

The best characterized example of a yeast origin is ARS1 (180). ARS1 is ~125 bp long and is composed of 4 necessary sequence elements: the ARS consensus sequence (ACS), B1, B2, and B3 elements (61, 187) (Figure 3). The ACS sequence is an 11 bp sequence that is conserved with at least a 9/11 bp match in all *S. cerevisiae* origins. The origin recognition complex (ORC) initiator binds to the ACS and B1 elements and recruits additional replication factors. The B2 element functions as a DNA unwinding element and may also serve other roles (111). The B3 element binds the transcription factor Abf1 and maintains a specific chromatin conformation at the origin (120). New DNA synthesis begins at a specific nucleotide between the B1 and B2 elements (19).

Replication origins in multicellular eukaryotes are larger and more complex. Although the basic features of origins in multicellular eukaryotes may be similar to budding yeast, they are larger, more complex, and more difficult to study. For these reasons, much less is known about what constitutes an origin of replication in higher eukaryotes. Like origins of *S. cerevisiae*, origins in multicellular organisms bind ORC and have multiple DNA sequences that are necessary for the initiation of DNA

replication (4, 60). However, multicellular eukaryotic origins are ~1-6 kb long and contain no recognized consensus sequence. Chromatin organization has also been identified as an important component in origin activity (60).

The ARS assay that was so useful in characterizing yeast origins so far has not been successfully recapitulated in mammal cells. The inability to use this tool has slowed the progress of research in identifying functional requirements for DNA replication in mammals. Genetic dissection of mammalian origins has relied on the stable integration of WT or mutant origins at ectopic chromosomal locations to test the effect of mutations on origin function. Although this strategy works, it is cumbersome, restricting both the number of origins analyzed and the range of mutations that can be practically tested.

Specific start sites for replication initiation are used to replicate the mammalian genome. However, the determinant for origin specification on mammalian chromosomes is still controversial. In *Xenopus* oocyte extracts, any sequence can serve as a replication start site, arguing that sequence is not a major determinant in origin selection (35). Furthermore, ORC seems to bind any DNA sequence equally well (35). This suggests that such factors as transcription, chromatin, and nuclear localization may play dominant roles in determining specific start site for DNA replication (35, 60). However, specific sequence elements do contribute to origin selection. This has been clearly demonstrated in multiple investigations that move a DNA fragment containing an origin to an ectopic location, where it retains its function as a start site for DNA replication (5, 7, 8, 113, 118, 141, 203). Moreover, deletion of short sequences within these ectopic origin fragments can cause them to lose origin function (4, 5, 113, 203, 204). Although specific DNA sequences are important for origin function, it is unclear what those sequence elements

are doing. At this time, it is unclear what type of dynamic interaction exists with sequence-directed origin selection in the context of other factors such as transcription, chromatin, and nuclear organization.

Factors Affecting the Usage of Origins or the Timing of Origin Firing

Nucleosome positioning. In several model origins in yeast and eukaryotic viruses, the ordered positioning of nucleosomes at origins of replication is crucial for replication initiation. In the *S. cerevisiae* ARS1, the start site for DNA replication is maintained free of nucleosomes by the concerted effort of ORC and the transcription factor Abf1 (112, 172) (Figure 3). The Epstein-Barr virus also has a specific pattern of positioned nucleosomes at the viral origin, oriP (221). The concerted action of simian virus 40 (SV40) T-antigen and the CHRomatin Accessibility Complex (CHRAC) remodels nucleosomes at the SV40 origin, allowing DNA replication to initiate (6, 85).

Chromatin modifications. Epigenetic modifications and chromatin remodeling have been shown to play a role in the selection and activity of origins, but the exact nature of this role is not clear. The "Jesuit model" (many are called, but few are chosen) of site-specific initiation of DNA replication suggests that many more potential origins exist in the genome than are normally used (39). Consistent with this model, *Xenopus* and *Drosophila* embryos, which lack epigenetic modifications, have tightly spaced (~5 kb) start sites for DNA replication and no apparent sequence-directed specificity of origin selection (81, 103, 161, 201). Introduction of epigenetic modifications and the definition of transcription loci later in development correlate with an increase in inter-origin spacing

and a decrease in the total number of origins used (81, 161). In somatic cells, DNA replication start sites are normally restricted to specific locations.

In *Drosophila*, origins can be inactivated by targeted histone deacetylation and reactivated by targeted histone acetylation (3). Global disruption of histone deacetylation by mutation of the histone deacetylase dmRpd3 leads to an increase in overall DNA replication (3). Consistent with these findings, a similar investigation in *S. cerevisiae* determined that deletion of the histone deacetylase scRpd3 caused individual origins to fire earlier in S phase, and targeted tethering of the histone acetyltransferase Gcn5p caused a late firing origin to initiate replication early in S phase (198). Global changes in histone acetylation caused by the drug trichostatin A in human cells changes the pattern of origin usage at multiple loci (90).

Transcription, origin usage, and replication timing. Origins of DNA replication are commonly associated with gene promoters and CpG islands (12, 38, 116, 189). This association was mostly anecdotal until genome- and chromosome-wide origin identification studies were performed in *S. cerevisiae* and *Drosophila*, respectively. Genome-wide analyses in budding yeast showed that origins were more likely to be found near gene promoters, and that transcriptionally active euchromatic regions were associated with earlier S phase replication than heterochromatic regions (115, 152, 214). A similar chromome-wide investigation in *Drosophila* focused directly on the relationship between origin usage and transcription and found a strong statistical correlation between ORC and RNA Pol II localization (116). Although the link between active transcription and early-S phase replication is clearly established, this correlation is

not absolute. Early-firing origins have been identified in heterochromatic regions, and *vice versa* (164).

There are several possible reasons for the link between transcription and replication. One is that the chromatin environment that favors DNA accessibility for the loading of RNA polymerase might also favor the mechanisms for replication initiation. Chromatin remodeling complexes associated with gene promoters have been labeled as "transcription factors," but in fact their general purpose is to facilitate the binding of any DNA-binding complex (93). An illustration of this concept is seen during V(D)J recombination, where transcription factor binding, but not transcription, is required for recombination (140, 167). Another potential reason for the link between replication and transcription may involve epigenetic inheritance of modified histones. Supporting a role for DNA replication in epigenetic re-organization, changes in DNA replication at the HoxB locus precede, and are necessary for, developmental changes in the transcription of the Hox genes (49).

Multiple investigations link gene activity with replication timing and, in some cases, a change in the pattern of origin usage (27, 42, 45, 49, 51, 133, 171). A survey of 6 origins of replication associated with CpG islands at gene promoters on the X chromosome found that upon X inactivation, CpG methylation, and gene silencing, origin usage was not affected, but the timing of replication was delayed (63). Thus, a certain class of origins have a transcriptional link with the timing of their firing in S phase, but their overall usage is independent of transcription. Although it has been a common notion to think of replication timing as an effect of transcriptional activity, a recent report demonstrated the opposite. When an origin of replication was integrated with a transgene

at a repressive locus, a delay in replication timing precedes histone deacetylation and gene silencing (51). This report supports earlier investigations linking gene activity with replication timing (27, 42, 45, 133, 171), but it also adds a new dimension in that the delay of replication occurred before gene inactivation and therefore was not a consequence of gene inactivation.

CpG methylation. The consequence of CpG methylation on replication initiation is still unclear. Rein et al. (1999) reported that at the locus containing the active Chinese hamster DHFR origin, most CpG dinucleotides are normally methylated, but in cells that have reduced CpG methylation the origin is not active (155). However, the CpG island seen near the c-myc origin of replication is normally not methylated (155), and in Xenopus egg extracts DNA methylation prevents ORC assembly and inhibits DNA replication (74). Methylation of a CpG island in the human HPRT gene, and in 6 other origins associated with genes and CpG islands on the X chromosome, had no observed effect on the usage of an origin located there (34, 63). These conflicting results suggest that the effect of DNA methylation on origin activity may be specific for individual loci, and that additional mechanisms may regulate individual origins.

Drugs inhibiting replication fork progression, and their effect on origin selection. Treatment of cells with aphidicolin will uncouple DNA polymerases from the rest of the replication fork, leading to large regions of unwound and unreplicated DNA (106, 177). This uncoupling will quickly lead to a cell-cycle arrest and makes the unwound sections of single-stranded DNA susceptible to breakage, leading to a DNA damage response (95). Low doses of aphidicolin slow replication forks without completely inhibiting DNA synthesis. Treatment of cells with the nucleotide synthesis

inhibitors hydroxyurea (173), methotrexate, fluorodeoxyuridine, or excess thymidine also leads to cell-cycle arrest and activation of a DNA damage response (95, 197).

Upon depletion of nucleotide pools during S phase, a change in the profile of origin usage in hamster cells was observed in which normally dormant origins fired (11, 37). The activation of these "dormant" origins is hypothesized to occur in response to the slowing of replication forks, so that the slowly replicating portions of the genome will not remain unreplicated at the end of S phase. Supporting this hypothesis, normally dormant origins at the yeast HML locus will fire when neighboring early-firing origins are inactivated (199). An explanation for both cases is that at a given locus, multiple origins are licensed and ready to initiate replication. The dominant origin will fires first. The secondary "dormant" origins are passively replicated from forks emanating from the dominant origin in a normal setting, but if the dominant origin fails to fire or replication forks are slowed, the dormant origins will have time to fire before they are passively replicated.

The Chinese hamster DiHydroFolate Reductase (DHFR) origin beta as a model. The Chinese hamster DHFR origin-beta is one of the best characterized origins of DNA replication currently under study. It will be the focus of some of the investigations presented in this thesis. Ori-beta is the strongest of 3 preferred start sites of replication identified in a 55 kb intergenic zone of replication initiation located downstream of the dihydrofolate reductase (DHFR) locus in Chinese hamster cells (4, 189) (Figure 4). Ori-beta is located 17 kb downstream of the DHFR gene, ori-beta' is located approximately 4 kb downstream of ori-beta, and ori-gamma is located 23 bp further downstream (9, 23, 24, 44, 76, 92, 105, 144, 205). Peaks of replication initiation

were determined over ori-beta and ori-beta' by quantitation of nascent strands, and the valley between these peaks occurred next to a GA+CA dinucleotide repeat (DNR) sequence (92), suggesting that DNR may act as a boundary between these two origins.

Replication initiation at origins in the DHFR intergenic region is linked with transcription of the DHFR gene. Deletion of the DHFR promoter causes a delay in replication initiation at origins in the intergenic region that can be reversed by substitution with a *Drosophila* promoter (158). Transcription termination signals exist at the 3' end of the DHFR gene that block transcription elongation into the intergenic region, and deletion of these signals causes a loss of replication initiation in the entire intergenic region (126). Transcription of DHFR and neighboring genes prevents replication initiation within the transcribed regions, but this block is suppressed upon treatment of the cells with a transcription inhibitor (160). Thus, many potential origins exist at and around the DHFR gene, but transcription limits those chosen to fire.

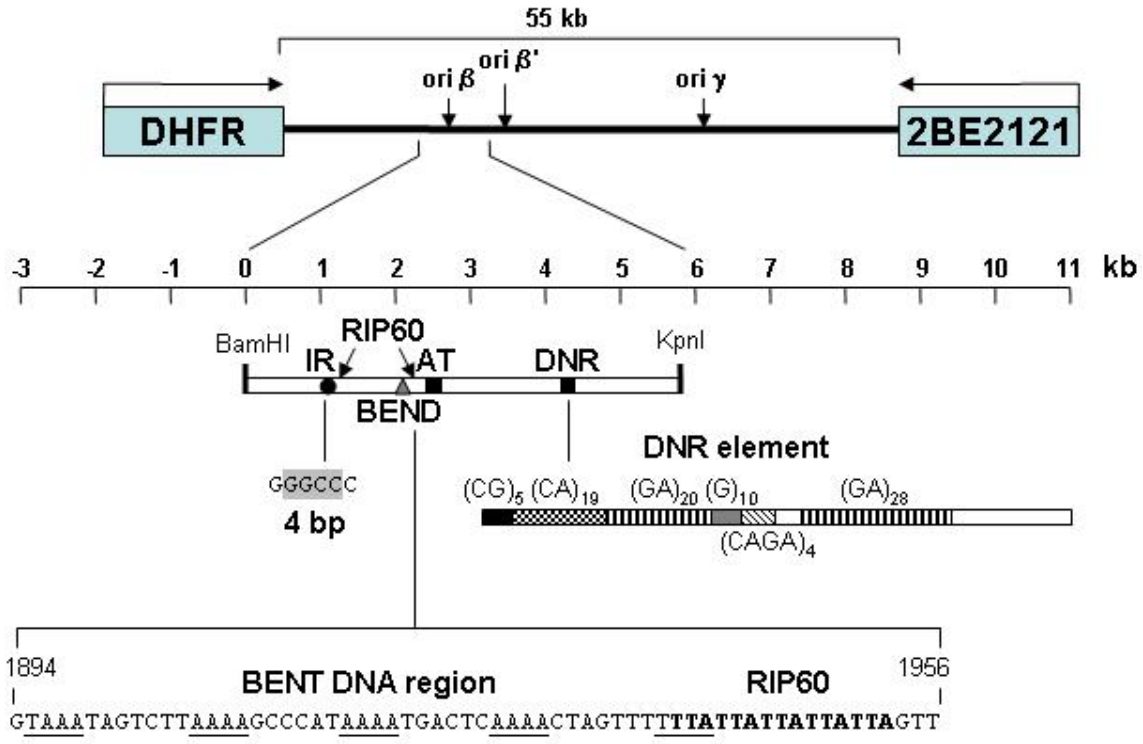


Figure 4. The DHFR intergenic region and origin beta. Necessary sequence features in the 5.8 kb BamHI/KpnI ori-beta fragment are shown. IR, initiation region; RIP60, 60 kDa replication initiation protein; BEND, sequence-induced bend; AT, 343 bp AT-rich region containing $(AT)_{47}(A)_{14}(T)_{20}$; DNR, 235 bp CA+GA dinucleotide repeat element (7, 8). The 4 bp (gray highlight), bent DNA (underlined), and RIP60 binding site (bold) sequences are given. The DNR element is enlarged to display the composition of its nucleotide repeat tracts.

Five *cis*-acting sequence elements have been identified in the Chinese hamster DHFR ori-beta that are necessary for the full activity of the origin when ori-beta is integrated at random ectopic locations (7, 8) (Figure 4). One of the functionally important elements in DHFR ori-beta is a 235 bp DNR sequence containing a $(CA)_{19}$ $(GA)_{69}$ dinucleotide repeat. Deletion of DNR reduced ectopic ori-beta activity ~9-fold (7) in hamster cells. Deletion of the 1.5 kb NK fragment at the 3' end of ori-beta, which contains DNR, reduced ectopic ori-beta activity 3.5-fold (Altman and Fanning, unpublished). These data suggest that the DNR sequence element is required for the

origin to function properly. The second element is identified as a 4 bp sequence, GGCC, within a GGGCCC hexanucleotide palindrome. Deletion of this sequence or replacement with CATG reduced ectopic ori-beta activity in hamster cells by at least half (7, 8). A 344 bp AT-rich element in ectopic ori-beta was also identified that, when deleted, reduced ori-beta activity ~6-fold (7) in both hamster and human cells. A region of stably bent DNA between the ori-beta initiation region and the AT element was shown to be important when mutation of the sequences responsible for the bend were mutated and ectopic ori-beta initiation activity in hamster cells was lost (8). Mutation of one of two binding sites for the 60 kDa replication initiation protein (RIP60) resulted in a ~5-fold decrease in ectopic ori-beta initiation activity in hamster cells (8). Based on these analyses of ori-beta, along with similar mutational analyses of the human c-myc (113, 118), laminB2 (141), and beta-globin origins (5, 203), mammalian origins require specific functional elements for replication initiation to occur.

Interestingly, replacement of the AT element with a similar sequence from the human laminB2 origin restored the activity of ectopic ori-beta in hamster or human cells, but replacement with a spacer DNA did not (8), thus providing evidence of conservation of functional elements between mammalian origins from different species. This result supports the idea that a candidate-based replacement strategy can be used to elucidate the function of necessary elements.

Replication Dynamics Affecting Tri-Nucleotide Repeat (TNR) Stability

At least 20 genetic diseases are caused by the expansion of trinucleotide repeats, including myotonic dystrophy and Huntington's disease (30, 54, 143). All TNR diseases begin with a "pre-mutation" state when the repeat number becomes unstable for unknown reasons. Through successive generations, as the repeats expand, the disease phenotypes appear and increase in severity (30, 31, 54, 129, 143). It is possible that these mutations all share a similar mechanism for repeat instability. However, a "global" change in the nucleus or in *trans*-acting protein factors cannot easily account for the disease-relevant repeat expansions, since expansion of repeats occurs at a specific locus, not at multiple loci at once (30, 104, 129, 130). When considering potential mechanisms for repeat expansions, it is productive to focus on events or defects that are specific for the disease locus, whether that event/defect is a random mistake in normal DNA metabolism, a very specific *trans*-acting protein factor, an epigenetic modification, or a mutation in a *cis*-acting DNA sequence.

Candidate DNA replication origins have been discovered near disease-relevant TNRs. These include the spinal cerebellar ataxia type 7 (SCA-7), Huntington disease (HD), and spinal and bulbar muscular atrophy (SBMA) genes (54), as well as the FMR2 (FRAXE) gene (28). The FMR2 and FMR1 genes include a similar arrangement of CGG repeats, transcriptional start sites, promoters, and CpG islands (54, 55, 66). Like the FMR1 locus, the FMR2 locus is also reported to replicate later in S phase in Fragile X patient cells (181), and the origin seems to be localized very close to its CGG repeats. For the TNR-proximal HD origin, there was no obvious difference in initiation between

affected and unaffected individuals (51). Although the effect of origin placement on tri-nucleotide repeat stability has been investigated in model systems, origins near TNR tracts have not been extensively characterized, and their implications for TNR stability have not been addressed.

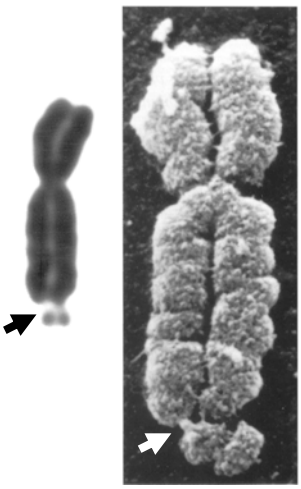
Protein mutations affecting TNRs. No defective protein factors have been implicated in TNR instability in human patients. However, a genetic approach has been taken in *S. cerevisiae* to understand the contributions of different components of the replication machinery, as they relate to TNR expansions and contractions (104, 157). The largest effect on TNR stability was observed upon deletion of the Rad27 (Fen1 in humans) flap exo/endonuclease protein (50). Rad27 is involved in the processing of the 5' flap of Okazaki fragments during DNA synthesis as well as base excision repair (109). Mutations in other proteins involved with Okazaki fragment maturation, namely PCNA and DNA ligase, also cause an increase in TNR instability (82, 154, 166). Specific mutations in Rad27, PCNA, and DNA ligase that disrupted their interactions with each other had additive effects on TNR instability, suggesting that the defect in loading of Rad27 and DNA ligase to Okazaki 5' flaps by PCNA was the cause of TNR expansion (154). Further supporting a role of lagging strand synthesis in TNR instability, mutations in lagging-strand polymerases alpha and delta caused an increase in TNR instability, but mutations in the leading-strand polymerase epsilon had no observed effect (165). Mutation of the proofreading domain of polymerase delta did not affect TNR stability. Rather, the mutations in polymerases alpha and delta should have the effect of slowing the rate of Okazaki fragment synthesis (165). Together, these mutational analyses of

replication proteins in *S. cerevisiae* implicate lagging strand DNA synthesis defects in TNR expansions and contraction.

Secondary DNA structure formation within TNR sequences. Disease-relevant TNR sequences (CTG, CCG, GAA, and GAC) share the common feature that they are capable of forming secondary DNA structures (131). TNRs that do not readily form secondary structures have not been associated with instability and disease. In each case, one strand of the TNR duplex will form a secondary DNA structure more readily than the other (CTG>CAG, CGG>>CCG, GTC>GAC), except in the case of GAA repeats in which both strands cooperate to form a triplex DNA structure (104, 121, 131). DNA secondary structures such as hairpins can be difficult to process during DNA replication, potentially leading to polymerase slippage or stalling (122).

Replication fork progression and fragile sites. Common fragile sites in the genome include, but are not limited to, microsatellite repeats and other di-, tri-, and poly-nucleotide repeat tracts. The fragile site associated with an expanded disease-allele CGG repeat tract next to the promoter of the Fragile X Mental Retardation gene (FMR1) is normally observed only when progression of replication forks is impeded. Aphidicolin, or nucleotide synthesis inhibitors methotrexate and fluorodeoxyuridine, retard the progression of replication forks. Treatment of Fragile X patient cells with either of these drugs results in the appearance of a fragile site at the FMR1 locus, visible on mitotic chromosome spreads (Figure 5) (68).

Figure 5. Metaphase X chromosome with the fragile site at Xq27.3. The Giemsa-stain (left) and scanning electron micrograph (right) images are adapted from Griffiths et al., 1991 (64). Expression of the fragile site was achieved by a deoxycytidine release from a thymidine block. Arrows indicate the fragile site in each image.



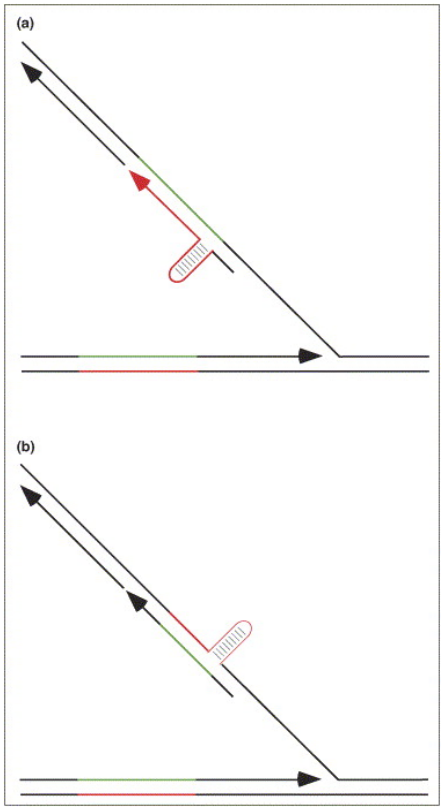


Figure 6. From Mirkin, 2006, Figure 3 (129). In this model, lagging strand synthesis dictates whether contraction or expansions occur. (A) If the hairpin-forming repeat strand is on the newly synthesized Okazaki fragment, expansions in the daughter strand are more likely. (B) If the hairpin forming repeat strand is on the lagging strand template, contractions in the daughter strand are more dominant.

Replication models of repeat expansion. The direction of replication, as well as the distance of an origin of replication from the TNR tracts, has been implicated in affecting TNR stability (31, 130). Because repeat instability occurs via lagging strand replication machinery, the direction of replication determines whether the hairpin-forming strand of the repeats will be on the parent or daughter strand (Figure 6). If the hairpin-forming strand is the lagging-strand template, contractions are favored (Figure 6B); if it is on the daughter strand, expansions are favored (Figure 6A). Supporting this conceptual model, expanded repeats integrated next to origins in *E. coli* (77, 148, 159), *S. cerevisiae* (13, 50, 128, 145), and COS-1 primate cells (29, 137) display orientation-specific expansions and contractions. An *in vitro* SV40 assay using human cell extracts also found that the direction of DNA replication specifies whether expansions or contractions occur (142).

In addition to replication direction, the distance between an origin and TNR tracts can affect their stability (29, 137, 142). This observation has prompted the Okazaki Initiation Zone (OIZ) model of repeat stability (31, 130, 143) (Figure 7). Since processing of the 5' ends of Okazaki fragments is implicated in repeat stability, the placement of Okazaki fragment initiation inside or outside repeats is hypothesized to be important (30, 31, 129, 130). If Okazaki fragment synthesis begins within the repeat tract, a greater chance for repeat instability is predicted. Consistent with this hypothesis, repeat instability is normally observed as the number of trinucleotide repeats expand beyond ~40-50 (120-150 bp), roughly the size of an Okazaki fragment (~135-145 bp) (10, 30, 75). Moreover, as the DNA duplex is unwound with a moving replication fork, the leading strand is replicated almost immediately after becoming single-stranded. In

contrast, a portion of the lagging strand is unwound prior to RNA priming and Okazaki fragment synthesis. The OIZ model predicts that if the unwound DNA contains a length of trinucleotide repeats sufficient for hairpin formation, this hairpin can form and impede DNA synthesis. Thus, the location of the start of the Okazaki fragment can affect how much of the repeats are included, and longer repeats have a greater chance of spanning the entire Okazaki fragment.

The OIZ model also predicts that slowing replication speed would extend the amount of time that the repeat tract remains single stranded and increase the chance of hairpin formation. If the OIZ model is correct, slowing replication fork speed should increase repeat instability. This may explain why the fragile site in Fragile X Syndrome patient cells is only observed when the cells are grown in replication fork inhibitors such as aphidicolin, methotrexate, or fluorodeoxyuridine (36, 52, 62, 135, 182) (Figure 5). Also supporting this model, treatment of muscular dystrophy type 1 (DM1) cells with the drug emetine, which preferentially inhibits lagging strand synthesis and results in large stretches of unwound lagging strand template, resulted in large (up to 170 repeats) expansions in the mutant TNR allele (31, 217).

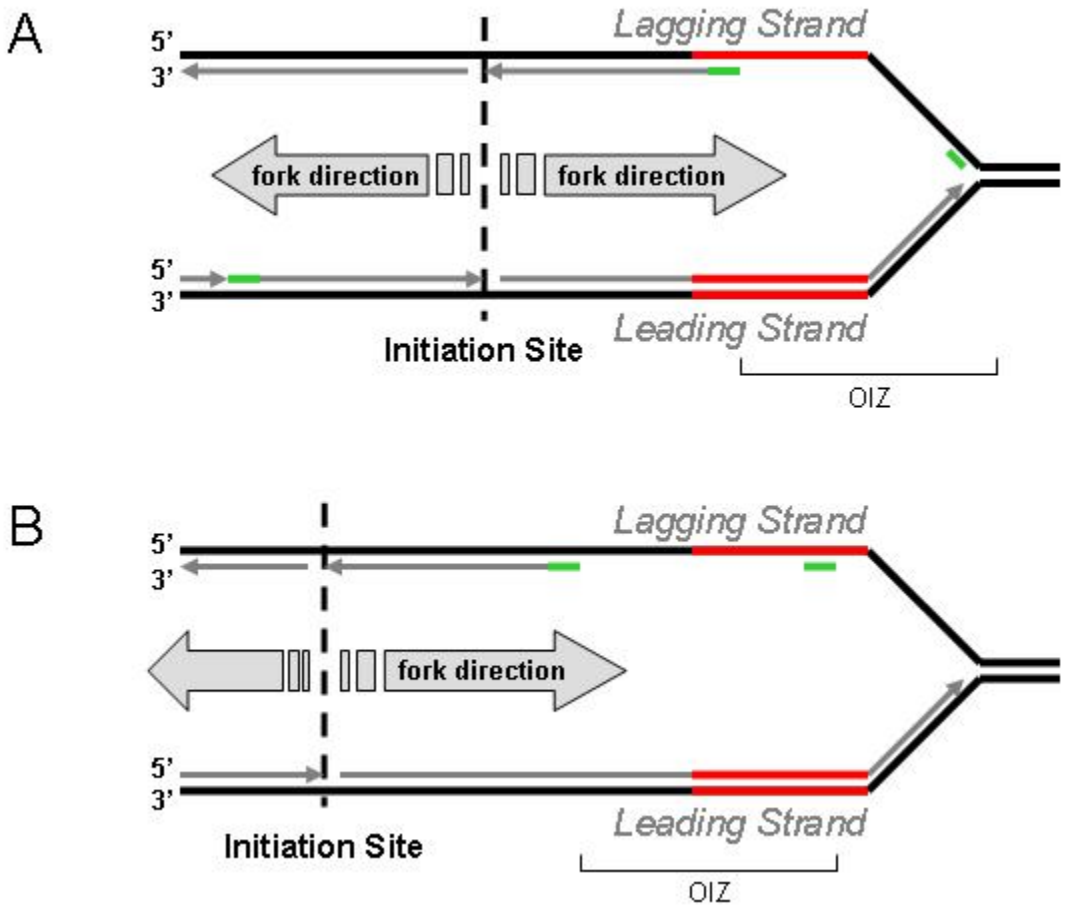


Figure 7. Okazaki Initiation Zone (OIZ) model of repeat instability. (A) Initiation of Okazaki fragments outside the TNR tract (red) results in reduced potential for instability. (B) If initiation of Okazaki fragment synthesis shifts such that it occurs within the TNR tract (red), there is predicted to be a greater chance for repeat instability due to errors in Okazaki fragment 5' processing and a greater probability of secondary structure formation. Black line, template strand; gray lines, daughter strand; green lines, RNA priming site for Okazaki fragment initiation.

CHAPTER II

AN ORIGIN OF DNA REPLICATION IN THE PROMOTER REGION OF THE HUMAN FRAGILE X MENTAL RETARDATION (FMR1) GENE

Introduction

Fragile X syndrome, the most common form of inherited mental retardation in males, is characterized by expansion of a CGG repeat tract in the 5' untranslated region of the human Fragile X Mental Retardation (FMR1) gene on Xq27.3 (139, 186, 191). The repeat number expands from 5-50 in normal males to over 200 in affected patients, accompanied by methylation of CpG dinucleotides and silencing of FMR1 gene expression, giving rise to the disease (139, 186, 191). DNA methylation of CpG dinucleotides in the FMR1 promoter is a necessary step in the inactivation of transcription, since FMR1 transcription can be reactivated in Fragile X affected cells upon treatment with 5-aza-2-deoxycytidine (73, 146). Moreover, rare individuals, whose CGG repeats are expanded but not methylated, still express FMR1 and display normal intelligence (67, 147, 174). The cause of the CGG expansion in the patient's maternal lineage is not known, but *cis*-acting factors are thought to play an important role (30, 104, 129, 130).

More than 20 genetic diseases are caused by the expansion of trinucleotide repeats, including myotonic dystrophy and Huntington's disease (30, 54, 143). Trinucleotide repeat diseases begin with a "pre-mutation" state when the repeat number becomes unstable for unknown reasons. Through successive generations, as the repeats

expand, the symptoms of disease appear and increase in severity (30, 31, 54, 129, 143). Expansion of trinucleotide repeats can thus be divided into two categories: 1) the initial trigger that leads to instability and 2) the mechanism for expansion once the repeat tract is lengthened and unstable. The mechanism that first initiates repeat instability may be different from the mechanism that propagates repeat lengthening once the repeat tract becomes unstable. The expanded repeat length itself may be enough to create instability, as observed when expanded CGG tracts are integrated at different loci in *Saccharomyces cerevisiae* and *Escherichia coli* (13, 77, 104, 168). Instability is thought to occur because expanded repeats form extensive stable secondary structures, such as hairpins, triplex, and quadruplex DNA, that are difficult to process correctly during replication, recombination, or repair (132, 145, 159, 168). In contrast, the initial transition from short, stable repeat tracts to longer, unstable ones remains poorly understood.

Some evidence has implicated replicon organization as a potential factor in the initial expansion of a stable repeat tract. Instability of a wild-type-length CAG repeat tract placed within ~350 bp of an SV40 replication origin was observed during viral DNA replication in monkey cells (29). In these experiments, the distance between the repeat tract and the origin, as well as the orientation of the repeats relative to the origin, affected the stability of the repeats. Similar experiments using CGG repeats demonstrated that replication fork dynamics, repeat length, and CpG methylation can affect repeat stability (137). Orientation-dependent repeat instability was also observed when CGG repeats were placed in the *E. coli* or *S. cerevisiae* genome near an origin of DNA replication (13, 159, 208). CGG repeats in the lagging strand template favored contractions, while CGG repeats in the newly synthesized Okazaki fragment favored expansions (13, 159, 208).

Consistent with the importance of replication direction, trinucleotide repeats can be induced to expand or contract in *S. cerevisiae* by mutating replication proteins involved in lagging strand, but not leading strand, DNA synthesis (104, 145, 157, 165, 166). These observations point to a potential role for origins of DNA replication in initiating and/or exacerbating trinucleotide repeat length instability.

In this chapter, I describe experiments to identify origins of DNA replication at sites surrounding the CGG repeats in the human FMR1 locus and to monitor their initiation activity in cells from normal and Fragile X individuals. An approximately 35 kb region surrounding the FMR1 promoter and the repeats was investigated. Initiation activity localized in the FMR1 promoter region, and the activity of the FMR1 origin was equivalent to that of two previously characterized origins of DNA replication analyzed in parallel as controls. The FMR1 origin was active in untransformed fibroblasts derived from normal male and female adults and from Fragile X-affected adults. The potential role of the FMR1 origin in CGG tract instability is discussed.

Materials and Methods

Cells and culture conditions. Cells used for analysis of origin activity at the FMR1 locus are described in Table 1. HCT116 and Hela were grown in DMEM with 10% heat-inactivated fetal bovine serum (FBS) and 5% CO_2 at 37° C. All untransformed fibroblasts were grown in DMEM with 15% un-inactivated FBS and 5% CO_2 at 37° C, except GM05381 (grown with 20% FBS) and HAF (grown in 10% heat-inactivated FBS).

TABLE 1. Cells

Cell Name	Description	Sex	Reference
HCT116	~diploid colorectal carcinoma	male	(21)
Hela S3	~triploid cervical adenocarcinoma	female	(150)
HAF	primary healthy adenoid fibroblast	male	
GM05381	primary healthy skin fibroblast	male	*
GM08400	primary healthy fibroblast	female	*
GM05848	primary FRAXA fibroblast	male	*
GM04026	primary FRAXA fibroblast	male	(78) *
GM07072	primary FRAXA fetal lung fibroblast	male	(79) *

*These cells were obtained from the Coriell Cell Repository in Camden, NJ. More information can be found at http://ccr.coriell.org/

Southern blot analysis of repeat tract length. Genomic DNA was isolated from cells as described in (8). The isolated genomic DNA was digested overnight with 10 µg RNase A at 4° C and purified by phenol/chloroform extraction followed by ethanol precipitation. Approximately 50 µg of genomic DNA was digested with PstI (NEB #R0140S) and electrophoresed through a 1% agarose gel. The DNA was blotted on a Highbond N+ or Highbond XL membrane (Amersham #RPN203B or RPN203S, respectively) and crosslinked using a Stratagene UV Stratalinker 1800. The blot was prehybridized in 15 mL Church buffer (0.5 M Na_2HPO_4, 1 mM EDTA, 7% SDS, 1% BSA) at 65° C for several hours. Fifty ng of a 1 kb PstI fragment from the FMR1 locus which includes the CGG repeats (Figure 8B) was radiolabeled with 50 µCi [alpha-^{32}P]dCTP (3000ci/mmol, PerkinElmer LAS # BLU513H250UC) using the High Prime labeling reagent (Roche # 11585592001). The radiolabeled probe (~50 ng at 1-2 x10^6 dpm/ng) was heat denatured in the presence of 750 µg salmon sperm DNA, added to the prehybridization solution, and hybridized overnight at 65° C. The blot was washed: twice with 2X SSC (300 mM NaCl, 30 mM NaCitrate) and 0.1% SDS for 15 min each at 25° C,

once with 1X SSC / 0.1% SDS for 15 min at 25° C, and four times with 0.1X SSC / 0.1% SDS for 5 min each at 65° C. The washed blot was exposed to a phosphor imager screen to detect the radioactive signal. PstI fragment lengths were verified with at least two independent Southern blots for each genomic DNA sample.

Reverse-transcriptase PCR. Actively growing cells (~1-3 x 10^6) were collected by trypsinization, and total RNA was extracted using the Qiagen RNeasy Mini kit (cat #74104) according to the manufacturer's instructions. cDNA was synthesized using the Invitrogen SuperScript III reverse transcriptase following the manufacturer's instructions, using 1 µg RNA (determined by absorbance at 260 nm) and 500 ng oligo$(dT)_{15}$ in a 20 µL reaction at 50° C. Samples (0.5 µL) of the reaction product were then used as a template for real-time quantitative PCR. The FMR1 primer set was designed to cross the exon 13/14 junction (Appendix Table 3), so that it specifically amplified cDNA from spliced mRNA, but not genomic DNA which contained the intervening intron. The standard used for this quantitative PCR was serial dilutions of the target PCR amplicon from HCT116 cells. Raw values for real-time PCR quantitation of cDNA for all samples are provided in the Appendix Table 6.

Isolation of nascent-strand-enriched DNA. Nascent DNA was isolated from asynchronously growing cells as described (113). Briefly, ~$7x10^7$ HCT116 cells, ~$3x10^7$ Hela S3 cells, or ~1-$3x10^7$ untransformed fibroblasts were harvested by trypsinization and loaded onto a 1.25% alkaline agarose gel. After a 15-20 minute alkaline lysis of the cells in the well, the DNA was electrophoresed for 16 hours at 30V. Single-stranded DNA of 0.5-1 kb or 1-2 kb was cut out of the gel and purified with a Qiagen gel extraction kit (cat # 28706) following the manufacturer's instructions. The DNA was

eluted in 10 mM Tris-HCl (pH 8.5) and used directly for quantitative PCR. Between 2 and 4 independent preparations of nascent DNA were isolated from each cell type.

Quantitative PCR primer optimization. PCR primers (PAGE-purified, Integrated DNA Technologies) were designed to have similar annealing temperatures and to create amplicons approximately 180-200 bp long. Each primer was used to amplify 1.2 ng of genomic DNA under Quantitative PCR conditions (see "quantitative PCR of nascent DNA") with 2-5 mM Mg^{++} (final concentration) to empirically optimize the Mg^{++} concentration. The annealing temperature of the primers was also varied to obtain a single clean PCR product. Primers were considered optimized when the PCR product melted completely at a single temperature (specific for each product), and when the PCR product gave a single band of the appropriate size when run on an agarose gel. A list of all primers used is provided in Appendix Table 3. PCR products, melting curve analysis, and examples of standard curves are provided in Appendix Figure 21.

Quantitative real-time PCR. Fourteen primer sets for the FMR1 locus (Figure 8), three primer sets for the LaminB2 origin (59) (Figure 11A), and three primer sets for the MCM4 origin (96) (Figure 11B) were used to quantitate target sequences in nascent-strand DNA samples. The LightCycler FastStart DNA master SYBR Green I kit (Roche # 12239264001) was used following the manufacturer's instructions. All reactions were carried out on a Roche diagnostic real-time PCR LightCycler. Magnesium concentration and annealing temperature (T_a) were optimized for each primer (see "quantitative PCR primer optimization," Appendix Table 3 and Appendix Figure 21). For each reaction, the cycling parameters were as follows: 10 minutes at 95°; 5 cycles at 95° for 15 sec, T_a+4° for 5 sec, and 72° for 15 sec; 5 cycles at 95° for 15 sec, T_a+2° for 5 sec, and 72° for 15

sec; 40 cycles at 95° for 15 sec, T_a for 5 sec, and 72° for 15 sec. Some CG-rich amplicons used a melting temperature of 96° or 97°. At the end of each run, a melting curve analysis was performed in which the PCR products were annealed at 72° and the temperature was gradually raised to 99°. In all cases, the PCR products melted in a narrow temperature range, indicating a pure PCR product without detectable non-specific amplification (Appendix Figure 21).

To quantitate the abundance of specific target sequences in nascent DNA, standard curves of amplification were done using two-fold serial dilutions of XhoI/NarI-digested HCT116 genomic DNA (50 ng to 0.1 ng). This range of genomic DNA included copy numbers of target sequence equivalent to those in the nascent DNA samples. Genomic HCT116 DNA digested with NarI only was used as a standard for LaminB2 amplicons since XhoI cuts within two of the amplicons used. Amplification profiles obtained from target sequences in nascent DNA samples were compared to the standard curve of cut HCT116 genomic DNA to quantitate the abundance of each target sequence in the nascent DNA sample.

Each PCR reaction was done in duplicate, and the average quantitation from the two reactions with every nascent DNA sample with every primer set is provided in the Appendix Table 5. For each nascent DNA preparation, relative nascent DNA abundance (initiation activity) was calculated by dividing the amount of target sequence in nascent DNA at each primer set by the amount of target sequence in nascent DNA at the outlying primer set(s) at that locus. Specifically, all values for the control MCM4 primer sets were divided by the EX6b value, all values for control LaminB2 primer sets were divided by the LB2C1 value, and all values for FMR1 primer sets were divided by the average of the

values from primer sets 10, 3, and 20 (see Appendix Table 3 or Figure 11C as an example). Initiation activities for like primer sets in different nascent DNA preparations of the same cell type were averaged, and standard deviation was calculated based on the number (n) of nascent DNA preparations with that cell type.

Chromatin Immuno-Precipitation (ChIP). 1 x 10^8 HCT116 cells were washed with PBS and treated with 1% formaldehyde in prewarmed medium for 5 min at 37°C. Nuclei were prepared based on a protocol by Mendez and Stillman (124) with modifications. Briefly, cells were harvested, washed with PBS, and resuspended in 4 mL hypotonic buffer A (10 mM Hepes pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM DTT, protease inhibitor). Cells were lysed by adding 0.04% Triton X-100 and incubated for 10 minutes on ice. Samples were centrifuged (4 minutes, 1300 *g*, 4°C). Nuclei were washed in ice-cold buffer A supplemented with 200 mM NaCl. After centrifugation (1300 *g*, 5 minutes, 4°C), fixed nuclei were washed with PBS, resuspended in 2.7 ml LSB (10 mM Hepes pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$) and lysed by adding 300 µL 20% Sarkosyl. The chromatin was transferred onto a 40 mL sucrose cushion (LSB plus 100 mM sucrose) and centrifuged (10 minutes, 4°C, 2500 *g*). Supernatant was removed, and the chromatin was resuspended in 4 mL TE and sonicated (Branson sonifier 250-D, 5 output 50 impulse, twice). For partial DNA digests, 3 mM $CaCl_2$ and 10 U micrococcal nuclease (Roche) per 1 mg chromatin were added and incubated for 10 minutes at 37°C. The reaction was stopped by adding 20 mM EDTA. For immunoprecipitation, 1/10 volume of 11x NET (550 mM Tris-HCl pH 7.4, 1.65 M NaCl, 5.5 mM EDTA, 5.5% NP40) was added to the extract followed by 10 µg affinity-purified polyclonal antibodies (HsOrc3p, HsMcm4p), or as control 10 µg rabbit IgG. The

immunoprecipitation and purification of co-precipitated DNA was performed as described (163).

Real-time PCR analysis was performed according to the manufacturer's instructions (Roche) using the same parameters and primer pairs described above and in Appendix Table 3. "Enrichment" of immunoprecipitated DNA is defined as the abundance of target sequence detected in the specific Orc3p or Mcm4p immunoprecipitate minus the abundance of target sequence detected by a non-specific IgG immunoprecipitate, divided by the abundance of target sequence detected in 30 ng of DNA purified from the pre-IP chromatin preparation (162).

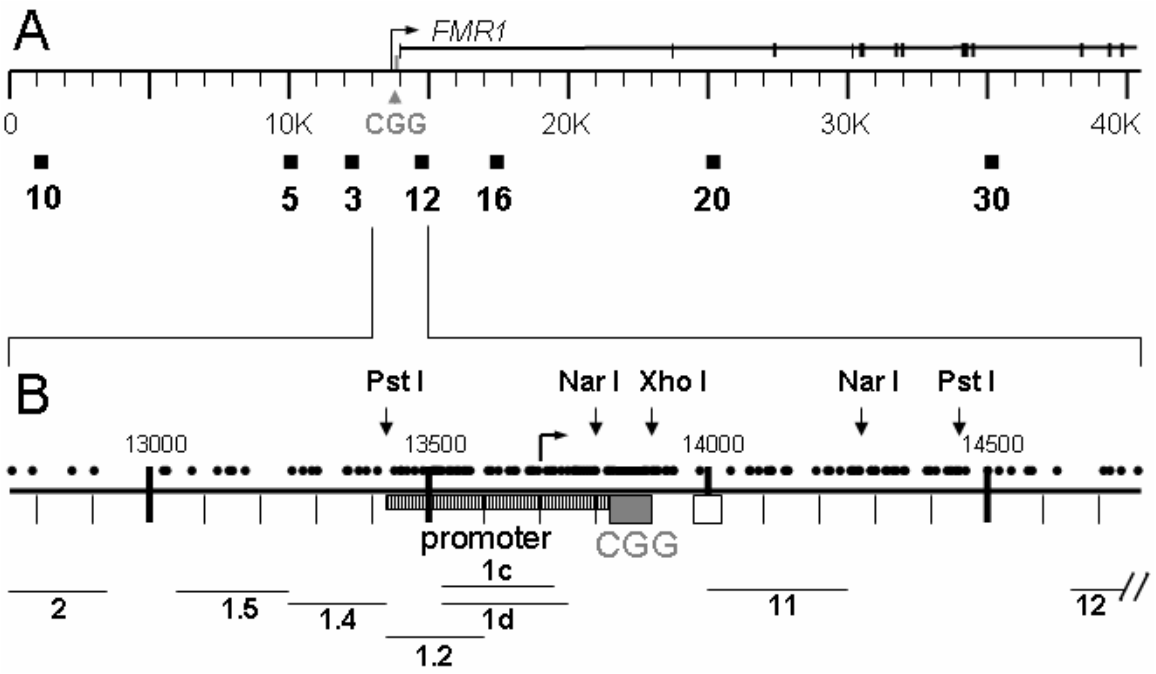


Figure 8. Diagram of the FMR1 gene locus. (A) A 40 kb region containing part of the FMR1 coding sequence and upstream region (Genbank accession # L29074) indicates the start site for transcription of the FMR1 gene (bent arrow), with the CGG repeats highlighted with a gray triangle. Above the scale is the FMR1 open reading frame, with the first 12 of 17 exons shown as vertical bars. Below the scale are the locations of the outlying primer sets (black boxes) used for quantitative PCR (Appendix Table 3). (B) A detailed diagram of the FMR1 promoter region, with hatches every 100 bp. Black dots indicate the locations of CpG dinucleotides in the sequence. The FMR1 promoter (box with vertical bars), CGG repeat tract (gray box), primary start site for transcription (bent arrow), and Exon 1 (white box) are indicated. PCR amplicons used for quantitative PCR (Appendix Table 3) are shown below the scale as horizontal bars, and pertinent restriction sites are shown above the scale.

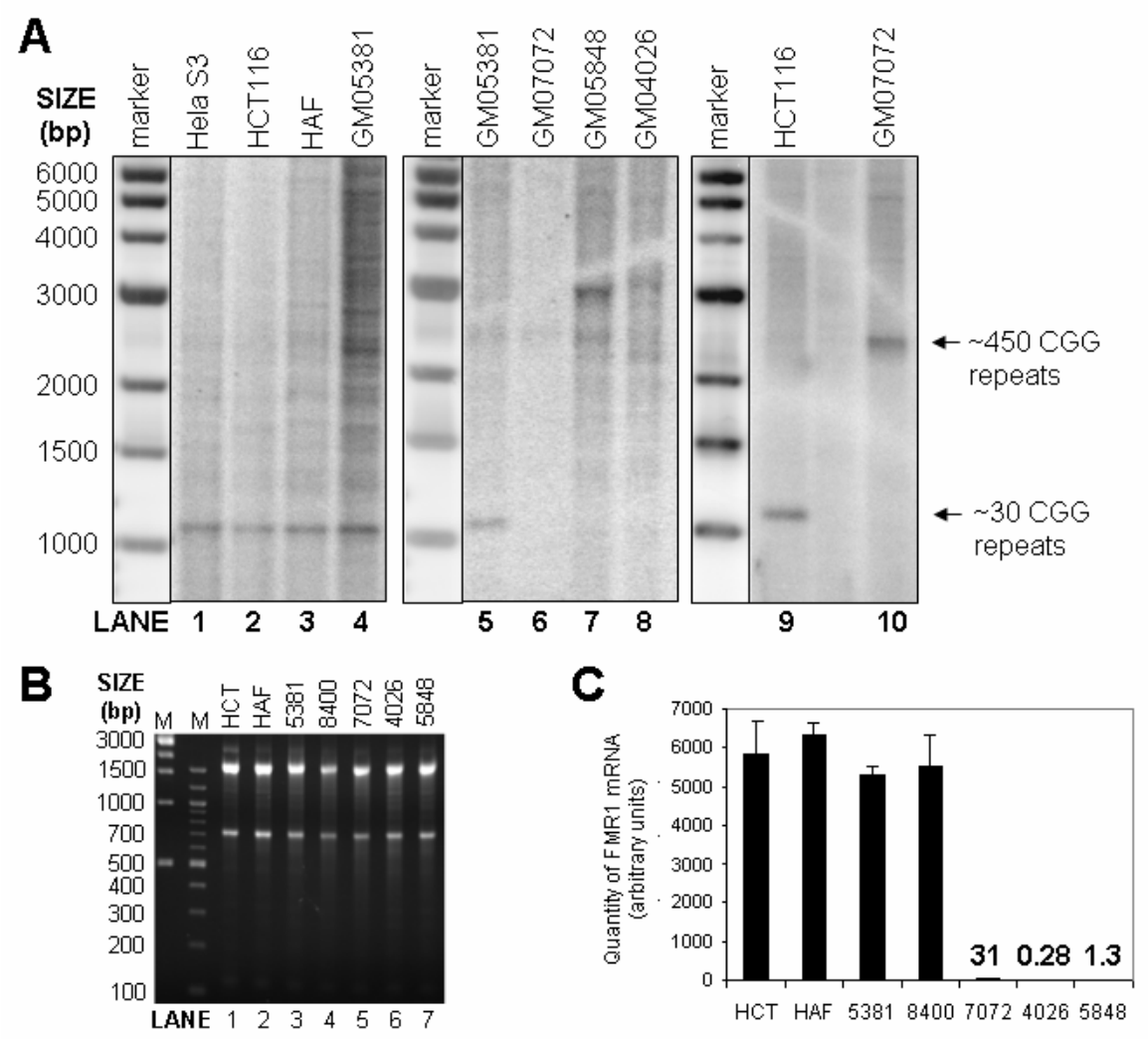


Figure 9. CGG tract length and FMR1 transcription. (A) Total genomic DNA from the indicated cell cultures was digested with PstI and analyzed by Southern blotting, using the radiolabeled 1 kb PstI promoter fragment (Figure 8B) as the probe. Lanes 1-5 and 9, unaffected cells with normal length CGG repeats; lanes 6-8 and 10, Fragile X cells with expanded repeat lengths. (B) One microgram of total RNA from the indicated cells was electrophoresed through an agarose gel and stained with ethidium bromide. (C) Equal amounts of total RNA from the indicated cells was reverse-transcribed to cDNA by RT-PCR, and equal amounts of cDNA were used as template in quantitative, real-time PCR with the primer set FMR1 13/14 (Appendix Table 3). The bars indicate the average quantitation from three independent sample preparations for each cell type. Brackets indicate SD.

Results

CGG repeat tract length and FMR1 transcription. A panel of human tumor cell lines and primary cultures from normal and Fragile X individuals (Table 1) was initially characterized by Southern blot to estimate the length of the CGG repeat tract in the FMR1 locus. A PstI fragment containing the FMR1 promoter region and the normal 5-50 CGG repeats has a length of ~1 kb (Figure 8), whereas a full-mutation allele containing >200 repeats has a length of over 1.5 kb (147, 211). Blots of PstI-digested genomic DNA from Hela S3, HCT116, HAF, and GM05381 cells displayed a PstI fragment of approximately 1 kb that hybridized with FMR1 promoter sequences, indicating a normal repeat length of ~30 repeats as previously reported for HCT116 cells (48) (Figure 9A, lanes 1-5, 9). The Fragile X genomic DNA (GM07072, GM05848, and GM04026) displayed PstI fragments of 2-4 kb (Figure 9A, lanes 6-8, 10). The PstI fragment lengths for GM05848 and GM04026 were not uniform, suggesting incomplete digestion, repeat instability, or mosaicism in the length of the repeat tract in the cell population. Despite the heterogeneity, these results confirm the presence of CGG repeat tract expansions in the Fragile X genomic DNA.

The expanded CGG tract observed in the Fragile X cells (Figure 9A, lanes 6-8, 10) predicts that FMR1 expression should be silenced (139, 186). To monitor FMR1 gene expression in the Fragile X patient cells, we measured the level of FMR1 transcription by quantitative reverse-transcriptase PCR. Equal quantities of RNA from all cells tested were used for cDNA synthesis, as determined by absorbance at 260 nm and visualization by agarose gel chromatography (Figure 9B, Appendix Table 7). FMR1 mRNA was

detected at similar levels in HCT116 cells and in the normal male and female untransformed fibroblasts (Figure 9C). In contrast, FMR1 mRNA levels were suppressed by 2-3 orders of magnitude in the Fragile X cells with expanded CGG tracts relative to the levels in normal cells (Figure 9C), supporting the prediction that FMR1 transcription is silenced (139, 186, 191).

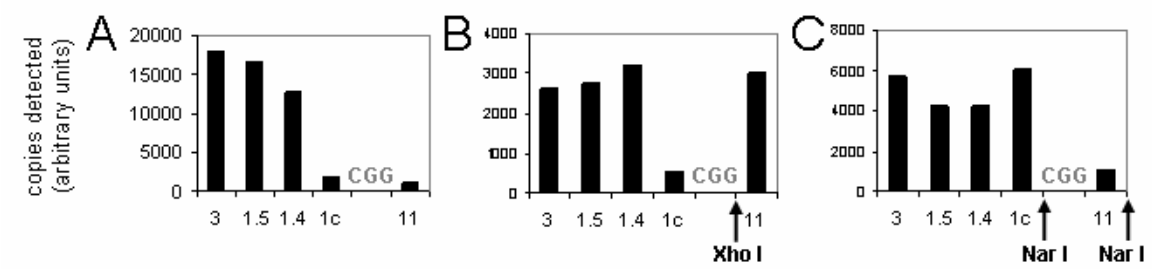


Figure 10. Flanking $(CGG)_n$ repeats reduce PCR amplification efficiency. Genomic HCT116 DNA either undigested (A), digested with XhoI (B), or digested with NarI (C) was used as template for real-time PCR amplification with the indicated primer sets in the FMR1 promoter region. Amplification was calibrated using known amounts of NarI/XhoI-digested genomic DNA to generate the standard curves. Arrows indicate the locations of XhoI (B) and NarI (C) restriction sites relative to primer locations and the CGG repeat tract.

Characteristics of primer amplification in the FMR1 locus. Primers were developed to amplify target sequences over a ~35 kb region of the FMR1 locus (Appendix Table 3). By varying magnesium concentrations and temperature, primer usage was optimized to specifically amplify each target sequence (Appendix Figure 21). However, initial characterization of primer efficiency on genomic DNA revealed that primers located near the CGG repeat tract displayed lower efficiencies of amplification than those located farther from the repeats (Appendix Table 4 and data not shown). These preliminary results raised the question of whether the CGG repeats might suppress the amplification efficiency of flanking sequences, even though the amplicon did not directly

contain the repeats. To test this possibility, we reasoned that if the amplicon template were on a separate DNA molecule, its efficiency of amplification should be greater than when it was flanked by the repeat tract. To separate the repeats from the amplification templates, genomic HCT116 DNA was cut with XhoI (which cuts immediately downstream of the repeats), with NarI (which cuts immediately upstream of the repeats), or with both NarI and XhoI (see Figure 8B). PCR amplification was carried out with five primer sets on each of these genomic DNAs. The NarI/XhoI-digested DNA was used to generate the standard curve for these experiments (Figure 10). PCR amplification using undigested genomic DNA as template with primer sets 1c and 11 was clearly lower than that with the other three primer sets (Figure 10A). When XhoI-digested DNA was used as the template, separating the primer set 11 sequence from the CGG tract, amplification with primer set 11 increased to a level comparable to that seen with the other three CGG-distal primer sets (Figure 10B). Conversely, when NarI-digested DNA was the template, separating primer set 1c from the CGG tract, amplification with primer set 1c increased (Figure 10C). These results demonstrate that the physical linkage of the CGG tract with the template DNA reduced the amplification efficiency of the flanking target sequences, and that eliminating the linkage with the CGG repeats alleviated the reduction in amplification efficiency. Based on these results, HCT116 genomic DNA digested with XhoI and NarI was chosen as the calibration standard to quantitate the amplification of target sequences in nascent DNA samples.

Detection of the FMR1 origin of DNA replication in transformed cells. To test for initiation activity near the CGG repeats of the FMR1 locus, a fraction enriched in short, newly synthesized, single-stranded DNA was prepared from HCT116 cells, which

are derived from a male individual and therefore have only one copy of the FMR1 locus per cell. As internal controls for each preparation of nascent DNA, real-time quantitative PCR was used to measure relative levels of target sequences at the well-characterized MCM4 and LaminB2 origins of DNA replication (2, 33, 59, 88, 90, 96, 141, 169, 184, 190) (Figure 11A, B). The MCM4 and LaminB2 origins showed easily detectable initiation activity in HCT116 cells at origin-proximal sequences, well above that at distal sequences (Figure 11C).

Real-time quantitative PCR was used with multiple FMR1 primer sets to measure relative levels of target sequences in the same preparations of nascent DNA at sites across the FMR1 locus. In HCT116 nascent DNA, a peak of abundance was detected at primer set 1.4 adjacent to the FMR1 promoter relative to outlying primer sets, using both the 0.5-1 kb and the 1-2 kb nascent DNA preparations as templates (Figure 11C). The level of initiation activity at the FMR1 promoter was comparable to that at the MCM4 and LaminB2 origins in the same preparations of nascent DNA. The pattern of initiation activity at the FMR1 locus in HCT116 nascent DNA was identical whether XhoI/NarI-digested genomic DNA (Figure 11C) or a XhoI/NarI-digested plasmid, containing the cloned HCT116 FMR1 sequence, was used as the real-time PCR calibration standard (data not shown).

Measurements of the PCR efficiency of the HCT116 nascent DNA compared to that of the XhoI/NarI cut genomic HCT116 DNA standard showed that these two templates had nearly identical efficiencies at primer sets EX6b, UPR4, 3, 1.4, and 20, validating the peaks of nascent DNA abundance at the MCM4 and FMR1 origins (Appendix Table 4). However, the amplification efficiency of primer sets within the

FMR1 promoter (1.2, 1c, and 1d) was lower-than with the cut genomic DNA used as a quantitation standard (Appendix Table 4), resulting in an underestimate of the 1.2, 1c, and 1d target sequences in the nascent DNA. Thus, the peak of FMR1 initiation activity detected with primer set 1.4 may actually extend into the promoter, closer to the CGG repeats (Figure 8).

To determine whether replication initiated at the FMR1 origin in other cells, Hela S3 nascent DNA was also tested. Since the peak of nascent DNA at primer set 1.4 was observed with both HCT116 nascent DNA size classes, only the 0.5-1 kb nascent DNA fraction was tested. The initiation activity at MCM4 and LaminB2 origin-proximal sequences in two Hela S3 nascent DNA preparations was lower than that in HCT116 nascent DNA preparations, but still greater than that at distal sequences (Figure 11D). Initiation activity was detected with several closely spaced primer sets in the FMR1 promoter region (Figure 11D). Even though the level of initiation activity in the FMR1 origin was lower in Hela S3 than in HCT116 cells, it was similar to the initiation activities seen at the MCM4 and LaminB2 control origins in the same preparations of nascent DNA. Relative to the control origins, the FMR1 initiation activity was equivalent in that in HCT116 and Hela S3 cells.

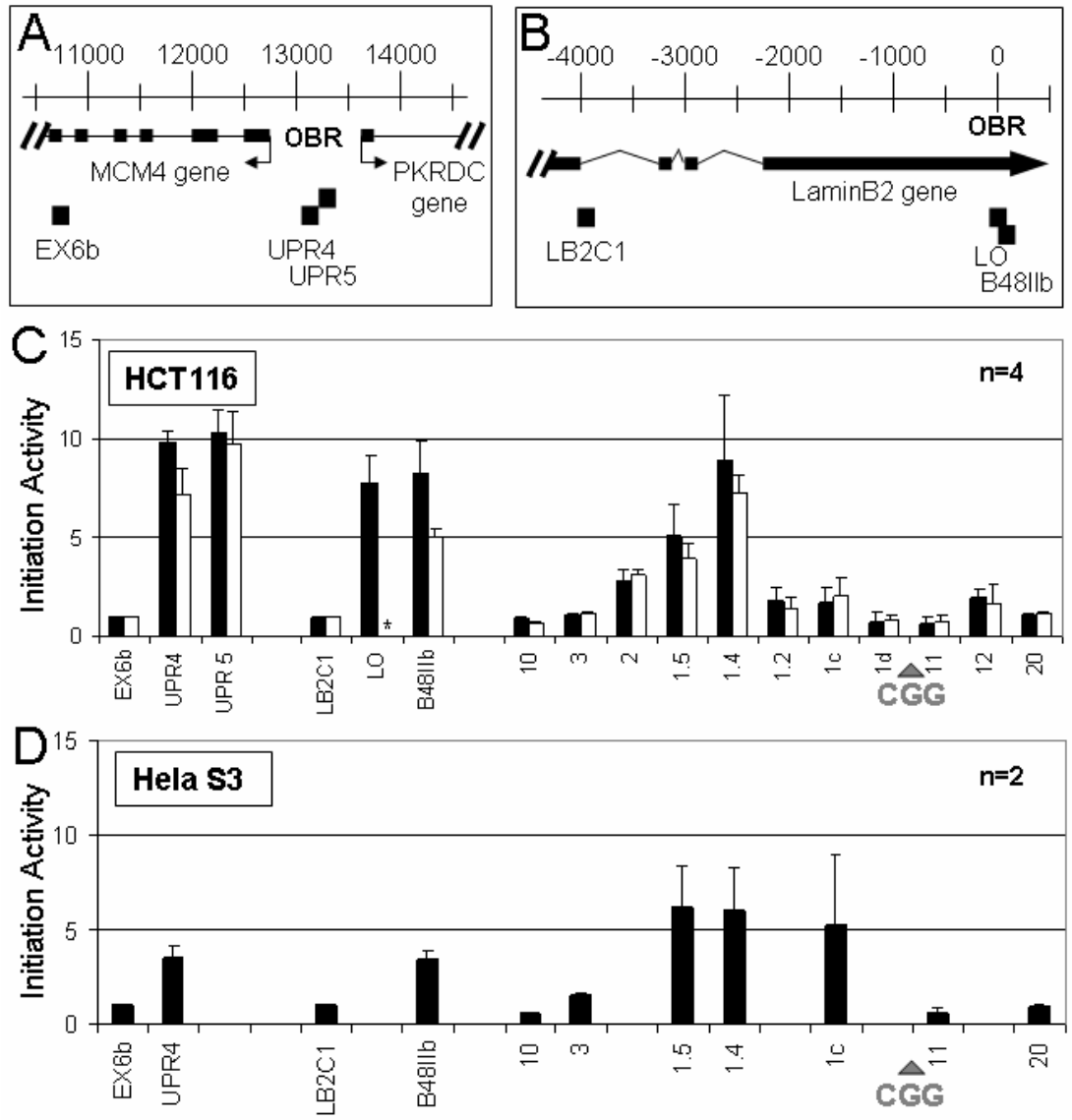


Figure 11. Replication initiates at the FMR1 locus in human cell lines. (A, B) As internal standards in each preparation of nascent DNA, initiation activity was measured at the human MCM4 and LaminB2 loci. Gene exons (thick lines), introns (thin lines), transcriptional start sites (bent arrows), and primer set locations (black boxes; Appendix Table 3) are shown below the scales. The location of the Origin of Bidirectional Replication (OBR) has been mapped approximately in the MCM4 locus (96) (A) and at high resolution in the LaminB2 locus (2) (B). Locations in bp are shown on the scale relative to Genbank accession # U63630 for the MCM4 locus (A) and to the OBR for the LaminB2 locus (B). (C) DNA sequence abundance at the indicated regions of the FMR1 locus in 4 independent HCT116 nascent DNA samples (n=4). "Initiation activity" calculations are described in Materials and Methods. Black bars indicate 0.5-1 kb nascent DNA, and white bars indicate 1-2 kb nascent DNA. *, not tested. (D) Abundance of indicated FMR1 sequences in 0.5-1 kb nascent DNA from 2 independent Hela S3 cell samples (n=2). Brackets in panels C and D indicate the SD for the indicated number of nascent DNA preparations (n).

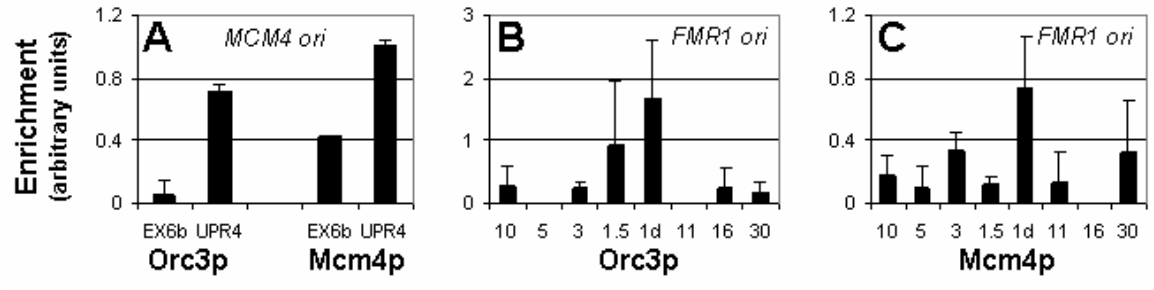


Figure 12. Orc3p and Mcm4p localize to the FMR1 promoter *in vivo*. A ChIP assay was done on 3 independently prepared samples from asynchronously growing HCT116 cells for Orc3p (A, B) and Mcm4p (B, C). (A) "Enrichment" for Orc3p and Mcm4p, as described in Materials and Methods, was tested using primer sets at the MCM4 origin (Figure 11B, Appendix Table 3), Enrichment was measured at the MCM4 origin (primer set UPR4) for Orc3p and Mcm4p relative to the outlying primer set EX6b. Enrichment for Orc3p (B) and Mcm4p (C) was tested using primer sets at the FMR1 locus (Figure 8, Appendix Table 3), with peaks of enrichment detected at primer set 1d in the FMR1 promoter. Brackets indicate SD of multiple independent immunoprecipitates.

Orc and Mcm proteins associate with the FMR1 origin in vivo. During the G1 phase of the cell cycle, licensed replication origins are bound by pre-replication proteins such as Orc1-6p and Mcm2-7p (1, 89, 96, 162, 190). If an origin of DNA replication exists in or near the promoter of the FMR1 gene, as suggested by nascent DNA abundance assays (Figure 11), pre-replication complexes would be expected to assemble in this region. To test this prediction, we used a Chromatin Immuno-Precipitation (ChIP) assay. As a control, we first confirmed enrichment of Orc3p and Mcm4p at the MCM4 origin (UPR4 primer pair) relative to a distal control region (Ex6 primer pair) (Figure 12A). In the FMR1 locus, the greatest enrichment of Orc3p was detected with primer set 1d at the FMR1 promoter region and slightly lower enrichment with primer set 1.5 (Figure 12B). Orc3p bound to origin-distal chromatin was detected with the outlying primer sites 3, 10, 16, and 30, but at significantly lower levels. With primer pairs 5 and

11, no Orc3p binding to chromatin was observed. Consistent with these findings, the greatest binding of Mcm4p in the FMR1 locus was detected at primer set 1d, with lower enrichment at the other primer sets (Figure 12C). The specific binding of Orc3p and Mcm4p within the FMR1 promoter is consistent with the existence of an origin of DNA replication in this region.

FMR1 origin activity in untransformed fibroblasts. The activity of the FMR1 origin was measured in nascent DNA isolated from untransformed human fibroblasts to test whether the initiation site detected in the HCT116 tumor cell line is also utilized in normal human cells. The peak initiation activity detected at the FMR1 origin in either HAF or GM05381 nascent DNA was equivalent to that at the MCM4 and LaminB2 control origins (Figure 13A and 13B), suggesting that the FMR1 origin is fully active in these normal male cells. The FMR1 origin in normal female cells (GM08400) showed initiation activity at closely spaced primer sets 1.5, 1.4, and 1c in the promoter region (Figure 13C). The level of FMR1 initiation activity in the female cells was nearly as high as that of the MCM4 control origin.

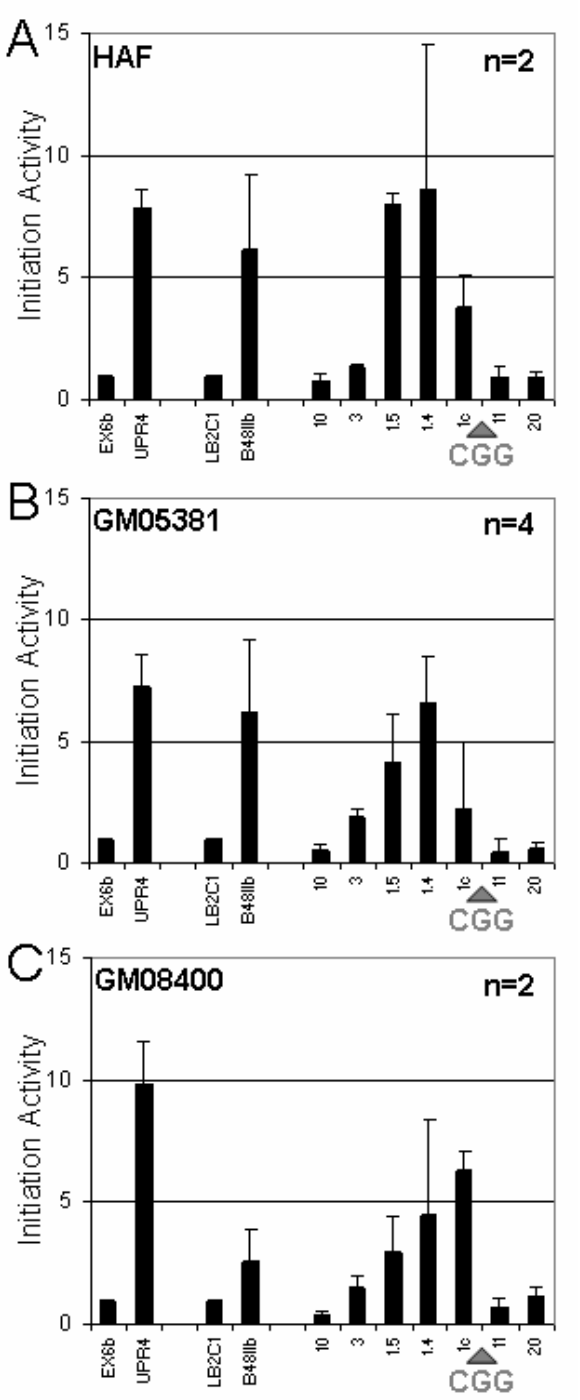


Figure 13. Replication initiates at the FMR1 origin in normal male and female fibroblasts. As described in Figure 11 and in Materials and Methods, initiation activity was tested with primer sets at the FMR1 locus in independent nascent DNA preparations from HAF male cells (panel A, n=2), GM05381 male cells (panel B, n=4), and GM08400 female cells (panel C, n=2). Initiation activity was tested in the same nascent DNA preparations using primer sets from the MCM4 and LaminB2 control origins. Brackets indicate the SD for the indicated number of nascent DNA preparations (n).

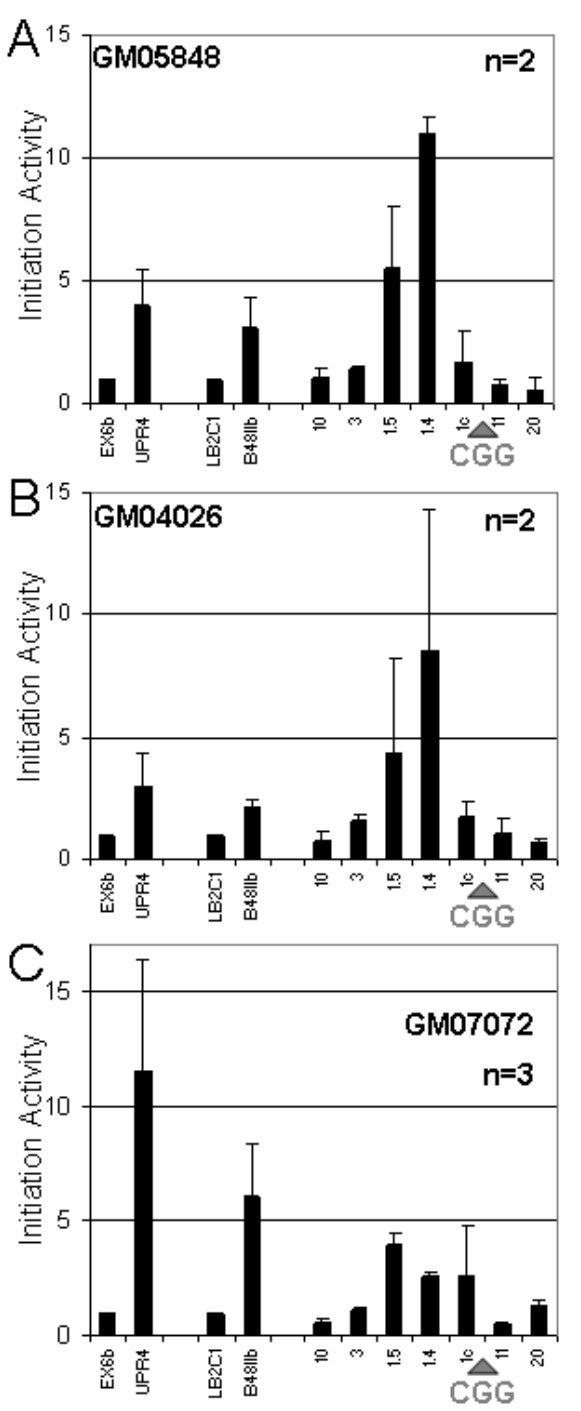


Figure 14. Replication initiates at the FMR1 origin in Fragile X cells. As described in Figure 11 and under Materials and Methods, initiation activity was tested with primer sets at the FMR1 locus in independent nascent DNA preparations from GM05848 male patient cells (panel A, n=2), GM04026 male patient cells (panel B, n=2), and GM07072 male fetal patient cells (panel C, n=3). Initiation activity was tested in the same nascent DNA preparations using primer sets from the MCM4 and LaminB2 control origins. Brackets indicate the SD for the indicated number of nascent DNA preparations (n).

FMR1 origin activity in Fragile X affected cells. FMR1 origin activity was tested in cells from three individuals affected with Fragile X syndrome to determine whether the site or the level of activity differed from that observed in normal cells. Importantly, the same site of initiation of DNA replication in the FMR1 locus was used in cells from all three affected individuals (Figure 14). The levels of initiation activity at the FMR1 origin in the two adult patient cells, GM05848 and GM04026, were similar (Figure 14A, 14B) to that observed in normal male fibroblasts. However, initiation activity at the MCM4 and LaminB2 origins was lower in the Fragile X cells than in the normal male cells. These comparisons suggest that the FMR1 origin is at least as active in the adult Fragile X cells as in the normal male cells, and perhaps more active.

Unexpectedly, in three independent preparations of nascent strand-enriched DNA from fibroblasts from a 22-week old fetus with Fragile X, GM07072, initiation activity at the FMR1 origin was reproducibly lower than in any of the other cells tested (Figure 14C). In contrast, GM07072 cells displayed strong initiation activity at the MCM4 and LaminB2 origins, comparable to that observed in normal cells (compare Figure 14C to Figure 13). These controls argue against the possibility that the low activity of the FMR1 origin in these nascent DNA fractions was due to poor nascent strand enrichment or slow cell growth. We conclude that the FMR1 origin activity in the Fragile X fibroblasts from this individual was low.

Discussion

Detection of a novel origin of DNA replication at the FMR1 promoter. We have mapped a peak of replication initiation in the promoter region of the human FMR1 gene that displays initiation activity comparable to that of known origins tested in the same nascent DNA preparations (Figure 11, 13). The co-localization of Orc and Mcm proteins (Figure 12) with the peak of initiation activity further supports the identification of the FMR1 origin of DNA replication. The lack of nascent DNA enrichment at primer sets distal from the FMR1 origin suggests that this is the primary initiation site within a 35 kb region flanking the FMR1 promoter and CGG repeats (Figure 11C and data not shown). These features of the FMR1 origin identify it as a new member of a class of origins that reside near CpG islands and the promoters of housekeeping genes (12, 34, 38, 63, 116, 189).

The low efficiency of PCR amplification at target sequences flanking the CGG repeat tract limited the region of accurate PCR quantitation of target sequences in genomic DNA, but was alleviated by cleaving the template away from the repeats (Figure 10). Since the same limitation applies to quantitation of CGG-proximal target sequences in the single-stranded nascent DNA fraction (Appendix Table 4), the abundance of target sequences detected by primer sets 1.2, 1c, and 1d is somewhat underestimated. Hence, the peak of initiation activity may include these target sequences in the promoter (Figures 11C, D, 13, 14). Our results illustrate the effects that unequal PCR efficiencies can have on the accuracy of quantitation, warranting caution in interpreting peaks of nascent DNA abundance near sequences such as trinucleotide repeats.

Differential regulation of the FMR1 promoter and FMR1 origin in Fragile X patient cells. The FMR1 promoter region has a complex architecture of bent DNA, CG-rich tracts, a $(CAAAC)_6$ tract, occupancy by multiple transcription factors (53, 73, 94, 110, 175, 176) and replication factors (Figure 12). The contributions of these features to FMR1 origin activity remain to be determined. Methylation of CpG dinucleotides (53, 139, 191), remodeling of chromatin (53, 73), and loss of transcription factor binding (175, 176) have been reported in the promoter of Fragile X somatic cells. The FMR1 origin is as active in adult Fragile X patient cells as in normal male cells (Figure 13 and 14). Since transcription of FMR1 is silenced in the patient cells (Figure 9B), the FMR1 origin activity is maintained regardless of transcriptional activity.

The transcriptionally silent FMR1 locus in Fragile X patients replicates later in S phase than does the active FMR1 locus in unaffected individuals (70-72, 181, 218). Since our data show that the FMR1 origin was active in GM04026 cells (Figure 14B), and these cells were previously shown to have delayed replication in the Fragile X locus (71), transcriptional silencing is correlated with a delay of replication timing but not FMR1 origin usage.

The position and usage of the FMR1 origin has implications for CGG repeat stability. A replication fork emanating from the FMR1 origin would replicate the CGG repeats such that the CGG sequence is the lagging strand template and the CCG complement would be in the newly synthesized Okazaki fragment (Figure 15). The CGG repeats more readily form stable DNA secondary structures such as hairpins than their CCG complement (104, 129-132, 219). Hairpin-forming trinucleotide repeats appear to be more prone to instability when located in the lagging strand template, favoring

contractions over expansions (30, 104, 129, 145, 156, 157, 159, 193, 208). Thus, the position of the FMR1 origin relative to the CGG repeat tract suggests that it might normally favor contractions when replicating unstable repeats.

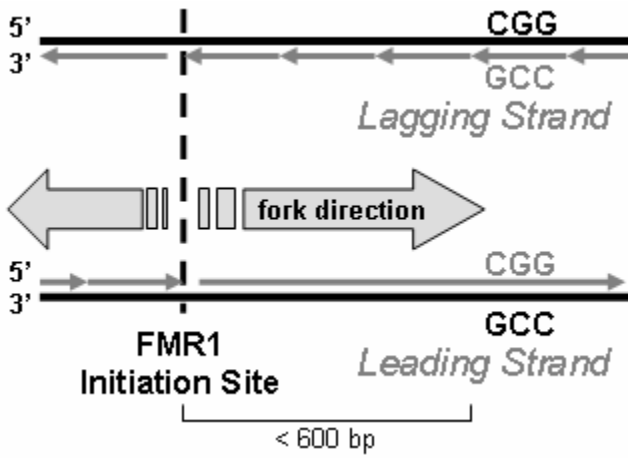


Figure 15. Model of replication fork movement originating from the FMR1 origin. Black lines represent the template strand, and gray arrows indicate the daughter leading and lagging strands. A dashed line indicates the transition point between leading and lagging strand synthesis. The lagging strand template thus contains the CGG repeats within 600 bp downstream from the initiation site. CGG more readily forms secondary structures than CCG, and DNA secondary structure formation in the lagging strand template is implicated in repeat instability, favoring contractions (104, 129, 130).

If the direction of DNA replication plays a role in expansion of the CGG repeats, one can hypothesize that a change in the pattern of origin usage, such that the CGG repeats are not replicated from the normal somatic FMR1 origin mapped here, could lead to repeat instability (130). The expansions seen in Fragile X patients occur through maternal transmission, and the repeat instability associated with transmission of the Fragile X allele is mostly specific to gametogenesis and/or early embryogenesis (30, 119, 134). During early development, the pattern of origin usage in humans is not known, but in model systems such as *Xenopus* and *Drosophila*, origin usage is relatively non-specific

during early development (81, 103, 161, 201). A different pattern of origin usage during early human development might provide a time window in which the repeats would escape the stable or potentially contraction-favoring pattern of replication offered by the somatic FMR1 origin reported here, and therefore become more susceptible to repeat expansion. One of three patient samples used in this study, GM07072, was derived from a 22-week fetus and showed reduced usage of the FMR1 origin, in relation to the MCM4 and LaminB2 control origins (Figure 14C). Although only one prenatal Fragile X individual has been examined, these results hint at a potential developmental change in origin usage. Analysis of normal and Fragile X cells from several individuals at different prenatal developmental stages will be needed to explore this possibility.

Other mechanisms known to cause transient changes in the pattern of origin usage are depletion of nucleotide pools during DNA replication (11, 37), global changes in histone acetylation (90), and agents that induce replication stress (47, 199). If a dormant origin on the downstream side of the CGG repeat tract became transiently activated in response to DNA replication stress, its position would tend to favor repeat expansion.

Since the position of the FMR1 origin would be predicted to favor CGG tract contractions in FMR1 patients (31, 104, 129, 130, 145, 157), why are they not more commonly observed in somatic cells of patients? An emerging role for CpG methylation in the stabilization of CGG repeats may give clues to this question. Methylation stabilizes CGG repeats in *E. coli* (138) and in primate cells in culture (137). The timing of CGG repeat instability, prior to germline segregation in early human development, coincides with the period in which the DNA lacks epigenetic modifications such as CpG methylation (30, 119). After this time, the repeats are stable and methylated, except in

the case of the male testes and sperm, which are unmethylated and display contractions (119, 185). An exception occurs in some high-functioning Fragile X patients who display mosaicism for repeat length and methylation (67, 174). Hypermethylated Fragile X cells display homogeneous and stable repeat lengths, but unmethylated repeats are heterogeneous for repeat length and stability (30, 56, 211, 212). Since the position of the FMR1 origin is predicted to favor contraction events, the hypomethylation and consequent instability of the expanded repeats in spermatocytes may provide a mechanism by which the repeats contract in these cells. This phenomenon could occur through normal DNA replication from the FMR1 origin. Postnatal somatic cells in the same individual would have hypermethylated, stable repeats that would resist contraction during normal DNA replication.

In summary, the results presented here indicate that an origin of DNA replication exists in the promoter region of the FMR1 gene, and that this origin is active in somatic cells with either normal-length or expanded CGG repeats, regardless of FMR1 transcription. The discovery and characterization of the FMR1 origin opens the possibility of exploring the relationship of a trinucleotide repeat sequence with its native origin of DNA replication.

CHAPTER III

DISCRETE FUNCTIONAL ELEMENTS REQUIRED FOR INITIATION ACTIVITY OF THE CHINESE HAMSTER DIHYDROFOLATE REDUCTASE ORIGIN BETA AT ECTOPIC CHROMOSOMAL SITES

Introduction

More than 40 years ago Jacob *et al*. (83) proposed a DNA replicon model which led to the discovery of replicators from bacteria to mammals (60, 61). A replicator is a *cis*-acting genetic element that directs replication initiation to occur at a specific location recognized by a *trans*-acting initiator. Budding yeast *Saccharomyces cerevisiae* use DNA replicators that contain a short consensus sequence that interacts with the origin recognition complex (ORC) (15). Genetic, biochemical, and physical mapping of origins of DNA replication on mammalian chromosomes suggests the existence of replicators that may specify DNA replication initiation sites in mammalian cells (5, 7, 8, 92, 113, 118). Moreover, ORC and its role in the initiation of DNA replication are conserved from yeast to mammals, indicating that mammalian replicators might share some features with those of budding yeast (1, 14, 35, 89, 96, 107, 117, 125, 190). However, unlike replicators in budding yeast, the *cis*-acting sequence elements that contribute to initiation activity of mammalian replicators have no consensus sequence except for a small AT-rich sequence bias (60, 61, 203).

To identify a chromosomal replicator in mammalian cells, an origin DNA fragment is placed at an ectopic chromosomal site and assayed for its capacity to direct

initiation of replication at the ectopic site. To achieve this, two general strategies have been used. Using a non-specific integration system, the 5.8 kb Chinese hamster DHFR origin beta (ori-beta) (7, 8) and the 1.2 kb human LaminB2 origin (141) showed replication initiation activity at multiple chromosomal sites in mammalian cells. These functioned as active replicators in pooled stably transfected cells or individual cell clones, although DNA replication activity varied at different chromosomal sites. The second strategy employed FLP- or Cre-mediated specific recombination integration systems to introduce the 2.4 kb human c-myc origin (113, 118) and the 2.6 and 3.2 kb human beta-globin origins (5, 203, 204) into unique ectopic chromosomal sites in a human cell line and a mouse cell line, respectively. The c-myc and beta globin origins serve as replicators at their specific ectopic chromosomal sites. Whether the two different strategies are equally effective in identifying mammalian replicators has so far not been validated by directly comparing them with the same mammalian origin.

One of the obstacles to understanding mammalian origins is the lack of identifiable sequence homology between different origins, but even in budding yeast replicators, some elements are not conserved, e.g. the B2 element in ARS1 (210) and the binding site for the Abf1 transcription factor in the ARS1 replicator (43, 120, 152, 179, 214). Interestingly, an Sp1 binding site can functionally replace the Abf1 site in ARS1 (215), suggesting that yeast replicators lacking an Abf1 site contain other elements that may serve the same functional role in directing initiation of replication. Moreover, these results suggest a hypothesis that may explain the DNA sequence diversity among mammalian replicators and that can be experimentally addressed.

The mammalian replicators characterized so far are composed of multiple sequence elements, of which several are required for initiation activity of the replicator. Five sequence elements identified within the 5.8 kb DHFR ori-beta fragment are necessary for full initiation activity of ori-beta at ectopic locations in Chinese hamster cells (7, 8). These include a 4 bp GGCC within a GGGCCC palindrome within the peak of initiation activity, an AT-rich element (AT), a CA+GA dinucleotide repeat element (DNR), a region of bent DNA, and a binding site for the 60 kDa Replication Initiation Protein (RIP60) (Figure 16). A sixth element (not shown) was dispensable for activity (7). Importantly, the hamster ori-beta AT element could be functionally replaced by a non-homologous sequence from the human laminB2 origin (8), consistent with the hypothesis that dissimilar DNA sequences from different mammalian replicators may serve similar functional roles. Further support for the concept that replicator elements that differ in DNA sequence can functionally substitute for each other comes from the ability of a Gal4-binding site to replace a 1.4 kb element in the 2.4 kb ectopic c-myc replicator in Hela cells expressing Gal4-CREB, but not Gal4 alone (58).

In this chapter, I will describe DNA sequences in the DNR element of hamster ori-beta that are crucial for full initiation activity of the ori-beta replicator at random ectopic chromosomal sites in hamster cells. We present evidence that either the *Xenopus* 5S ribosomal RNA gene or an element from the murine 28S ribosomal RNA regulatory region can replace the function of the DNR element to restore ori-beta initiation activity. Lastly, we demonstrate the requirement for DNR and the other ori-beta elements for replicator activity at ectopic chromosomal sites in human cells, using both the specific and non-specific integration strategies.

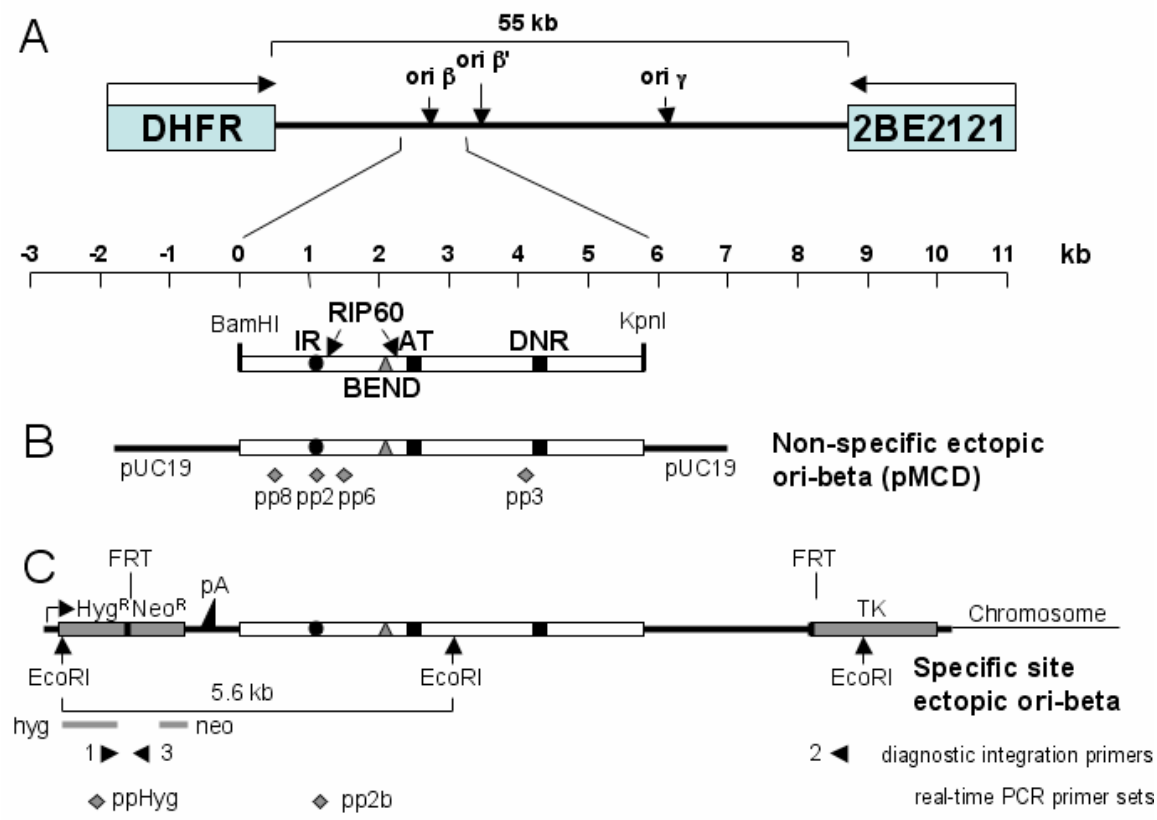


Figure 16. The DHFR origin beta at endogenous and ectopic locations. (A) Diagram of the endogenous DHFR ori-beta IR in hamster cell. Preferred start sites of DNA replication (ori-β, ori-β', and ori γ) within a 55-kb initiation zone between the genes DHFR and 2BE2121 are indicated. The 5.8 kb DNA fragment containing ori-beta extends from the BamHI (nucleotide position 1) to the KpnI restriction enzyme site. Previously identified (7, 8) functional elements of ori-beta are indicated: IR, initiation region; RIP60, 60 kDa replication initiation protein binding sites; BEND, sequence-induced stable bend in the DNA; AT, $(AT)_n$ repeats and AT-rich sequences; DNR, GA+CA dinucleotide repeat element. (B) 5.8 kb ori-beta sequence cloned into pUC19 (pMCD) and integrated at non-specific ectopic locations in DR12 and Hela cells. Locations of pp8, pp2, pp6 and pp3 PCR primer sites are indicated by gray diamonds. (C) 5.8 kb ori-beta sequence integrated at the FRT site in Hela 406 cells (113). The locations of the FRT sites, hygromycin-neomycin resistance fusion gene (Hyg^R Neo^R) with transcription start site (bent arrow), poly-adenylation and transcription termination site (pA), and thymidine kinase (TK) gene are indicated. The 5.6 kb EcoRI fragment used for Southern blot analysis, and diagnostic PCR primers 1, 2, and 3 for specific site integration are shown. The target sites for ppHyg and pp2b are indicated with gray diamonds.

Materials and Methods

Plasmid construction. The plasmid pMCD containing DHFR ori-beta, a 5.8-kb BamHI-KpnI fragment in pUC19, was the kind gift of N. Heinz (25). Mutant constructs pMCDΔDNR, pMCDΔAT, pRIP DSx, and pBENDx were described previously (7, 8). Full-length DHFR ori-beta and its mutant derivatives were subcloned into the plasmid pFRT.Myc (113) to replace the BamHI-NotI fragment including the 2.4 kb c-myc replicator, and named pFRT.ori-beta, pFRT.DNR, pFRT.AT, pFRT.RIP, and pFRT.BENT, individually. The DNR reversed mutant (pMCD-DNRrev) was generated by digesting pMCD with NheI and partially with XbaI to release the DNR fragment, then ligating its compatible cohesive ends back into the vacant DNR location in pMCD and screening for integrations in the opposite orientation. The resulting DNRrev mutation was subcloned from pMCD into the pFRT vector to make pFRT.DNRrev. The XK deletion mutant was made by digesting pMCD fully with KpnI and partially with XbaI, to delete a 1.3 kb fragment from nucleotide 4454-5781 (relative to the BamHI site) to make the plasmid pMCDΔXK.

Construction of the following DNR mutants used the same strategy: pMCD-0.5DNR1, pMCD-0.5DNR2, pMCD-DNRspcr, pMCD-SB2r, pMCD-5SRNA, pMCD-5SNPEf, and pMCD-5SNPEr. The SB2 insert (57, 114) came from the plasmid pUC-SB2+ and was the kind gift of F. Grummt. The 5S RNA insert (188) came from the plasmid pXP10 and was the kind gift of J.J. Hayes. Each mutant construct used a derivative of pMCD (pMCD-X) in which the TCTAGA XbaI site at the downstream border of DNR was changed to a CCCGGG XmaI site for easier cloning. For each

construct, the DNR element was removed by NheI/XmaI digestion, and the vacant gap was filled by ligation of a PCR-amplified fragment flanked with NheI- and XmaI-compatible cohesive ends. PCR primers used to amplify the replacement fragments are provided in Appendix Table 3. All changes to the normal ori-beta sequence were verified by sequencing. The DNRspcr mutation was subcloned from pMCD into the pFRT vector to make pFRT.DNRspcr.

Cell culture and stable transfection. Hela S3 (ATCC CCL-2.2) and DR12 (86) cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FCS at 37°C with 10% CO_2. Hela 406 cells, a kind gift from M. Leffak, were previously derived from Hela cells by integration of an FLP recombinase target site (FRT) (118) and maintained in Dulbecco's modified Eagle medium with 10% FBS and 50 µg of gentamicin/ml at 37°C with 5% CO_2.

Four µg of pMCD plasmid DNA or a mutant derivative plasmid mixed in a 3:1 molar ratio with PvuI-linearized pSV2neo DNA, were electroporated into either 5 x 10^6 Hela or DR12 cells using a Bio-Rad Gene Pulser at 360 V and 650 µF. After 3-4 weeks of growth under G418 (0.5 mg/ml) selection, ~100-200 G418 resistant clones per transfection were pooled for further analysis.

Hela 406 cell transfections were performed using Lipofectamine 2000 (Gibco-BRL) according to the manufacturer's instructions. For each transfection in a 24-well plate the total amount of DNA was 1 µg; the molar ratio of the donor plasmid to the cotransfected, FLP recombinase-expressing plasmid, pOG44, was 1:8. At 24- to 48-hrs post-transfection, selection for G418 resistance (0.5 ng of active component per ml of culture medium) was initiated and continued for 15 days. After single colonies formed, 20

μM ganciclovir was supplied for 2 to 3 days. Colonies resistant to hygromycin, G418, and ganciclovir were used for further analysis.

Diagnostic PCR screening for site-specific integration in the Hela genome. After using the site-specific FRT-mediated integration strategy, single-clone colonies resistant to hygromycin, G418, and ganciclovir were screened for site-specific integration by diagnostic PCR as previously described (113). Briefly, genomic DNA isolated from resistant colonies was amplified by PCR with the diagnostic primer sets 1+2 or 1+3 (Figure 16C). A PCR amplicon that spans the FRT acceptor site using the 1+2 primer set will be detected when no DNA is integrated. Upon integration of DNA at the FRT site the amplicon size becomes too long to amplify under the PCR conditions used. A PCR amplification product from primers 1+3 can be detected only when the Neo^R gene has integrated next to the Hyg^R gene at the FRT site.

Southern blots. For the non-specific integration system, 10 or 40 μg genomic DNA from uncloned pools of pMCD-transfected cells was digested with BamHI/KpnI and electrophoresed through a 1% agarose gel. The DNA was blotted on a Highbond N+ (Amersham #RPN203B) and crosslinked using a Stratagene UV Stratalinker 1800. The blot was prehybridized in 15 mL Church buffer (0.5 M Na_2HPO_4, 1 mM EDTA, 7% SDS, 1% BSA) at 65° C for several hours. The ectopically integrated ori-beta fragment was detected using two ori-beta probes, probe 1 and probe 2. Probe 1 is a 486 bp BamHI/StuI fragment at the 5' end of 5.8 kb ori-beta, and Probe 2 is a 484 bp NsiI/KpnI fragment at the 3' end of 5.8 kb ori-beta. The GNAI3 probe is a 505 bp PCR fragment from the GNAI3 locus in Chinese hamster cells (nt 4623 to 5128, Genbank Accession # X79282) amplified by the following primers: 5' ATGCTAATTGTAGTAGTGATCC and

5’ CCTCAAAAGGCACTGCTCC. The GNAI3 probe hybridizes to a 3.5 kb BamHI/KpnI fragment. Fifty ng of each probe fragment was radiolabeled with 50 µCi [alpha-^{32}P]dCTP (3000ci/mmol, PerkinElmer LAS # BLU513H250UC) using the High Prime labeling reagent (Roche # 11585592001). The radiolabeled probe (~50 ng at 1-2 x10^6 dpm/ng) was heat denatured in the presence of 750 µg salmon sperm DNA, added to the prehybridization solution, and hybridized overnight at 65° C. The blot was washed: twice with 2X SSC (300 mM NaCl, 30 mM NaCitrate) and 0.1% SDS for 15 min each at 25° C, once with 1X SSC / 0.1% SDS for 15 min at 25° C, and four times with 0.1X SSC / 0.1% SDS for 5 min each at 65° C. The washed blot was exposed to a phosphor imager screen to detect the radioactive signal.

For the specific integration system, Southern blot analysis was performed using standard methods with 10 µg of genomic DNA and hybridizing with either the 756 bp EcoRI/XbaI fragment of the Hyg gene from plasmid pFRT.Hyg.TK or the 405 bp NcoI/SmaI fragment of the Neo gene from pFRT.myc, as previously described (113).

Nascent DNA isolation and PCR-based nascent DNA strand abundance assay. For the non-specific integration system in DR12 or Hela cells, ~0.4-2 kb nascent DNA isolation using neutral sucrose gradient size fractionation was carried out as previous described (7, 8). Target sequence quantitation in nascent DNA-enriched fractions was done using competitive PCR for Hela samples as described (7, 8) or by Real-time PCR for DR12 samples. To test the initiation activity of origin constructs at the specific integration site at Hela chromosome 18, 1-2 kb nascent DNA was isolated by alkaline agarose gel size chromatography as previously described (113), and target

sequence quantitation of a nascent-enriched DNA fraction was done using Real-time PCR. Primer sequences are provided in Appendix Table 3.

The LightCycler FastStart DNA master SYBR Green I kit (Roche # 12239264001) was used for Real-time PCR quantitation following the manufacturer's instructions. All reactions were carried out on a Roche diagnostic real-time PCR LightCycler. The total volume of each PCR reaction was 10 µl with 3.5 mM Mg^{++}, 500 nM of each primer (Integrated DNA Technologies, Inc), and 3 µl of sample DNA. Reactions were started with 10 min at 95 ^{o}C, 3 cycles using the highest annealing temperature, and 3 cycles using the intermediate annealing temperature, followed by 39 cycles using the lowest annealing temperature. Fluorescence measurements were taken only during the last 39 cycles. During the final 3-sec step of each cycle when fluorescence was measured, the temperature was elevated to minimize any signals emitted by primer dimers and non-specific PCR products with melting temperatures below that of the target. Calibration curves were generated using serial dilutions of a known concentration of pre-sheared genomic DNA from the 406/pFRT.ori-beta cell line when using the specific integration system, or *ScaI*-linearized pMCD when using the non-specific integration system.

Table 2. PCR primer sequences for the amplification of DNR replacement fragments

Replacement	nt.[a]	Sequence (5' to 3')[b]	Template	Insert (bp)
DNRspcr	1202	AGTCTAGACAGCTTTTTCCTTTGTGGTGT	pSV2-neo[c]	235
	1429	CATCCCGGGACTGGTGGAATGCCTTTAATG		
0.5DNR1	3558	AAGGACCTCAGCCTCTGAAAC	pMCD[d]	132
	4351	ATGCCCCGGGTCTCTGCCTCTCTCCCTCTG		
0.5DNR2	4332	CTAGCTAGCACAGAGGGAGAGAGGCAGAG	pMCD[d]	126
	5075	ACCCTGTTCTCTGCTAAGCAG		
5S RNA gene	-79	ATATGCTAGCAATTCGAGCTCGCCCCGG	pXP10[e]	282
	+212	CGAGGTCGACTCTAGAGGA		
5S NPEr	-79	ATATCCCGGGAATTCGAGCTCGCCCCGG	pXP10[e]	139
	+74	ATGCGCTAGCTAACAGGCCCGACCCTGC		
5S NPEf	-79	ATATGCTAGCAATTCGAGCTCGCCCCGG	pXP10[e]	153
	+74	TATACCCGGGTAACAGGCCCGACCCTGC		
SB2r	MCS	CTAGATCCCGGGCGACTCTAGAGGATCCCC	pUC.SB2+[f]	161
	599	CTAGCGGCTAGCACTCCGGGCGACACTTTG		

[a] refers to the 5' end of each primer

[b] restriction sites for cloning into the vacant DNR site are underlined. "0.5DNR1" used the DNR-flanking NheI site (nt # 4219) for the upstream primer. "0.5DNR2" used the DNR-flanking XmaI site (previously XbaI in pMCD, nt # 4454) for the downstream primer.

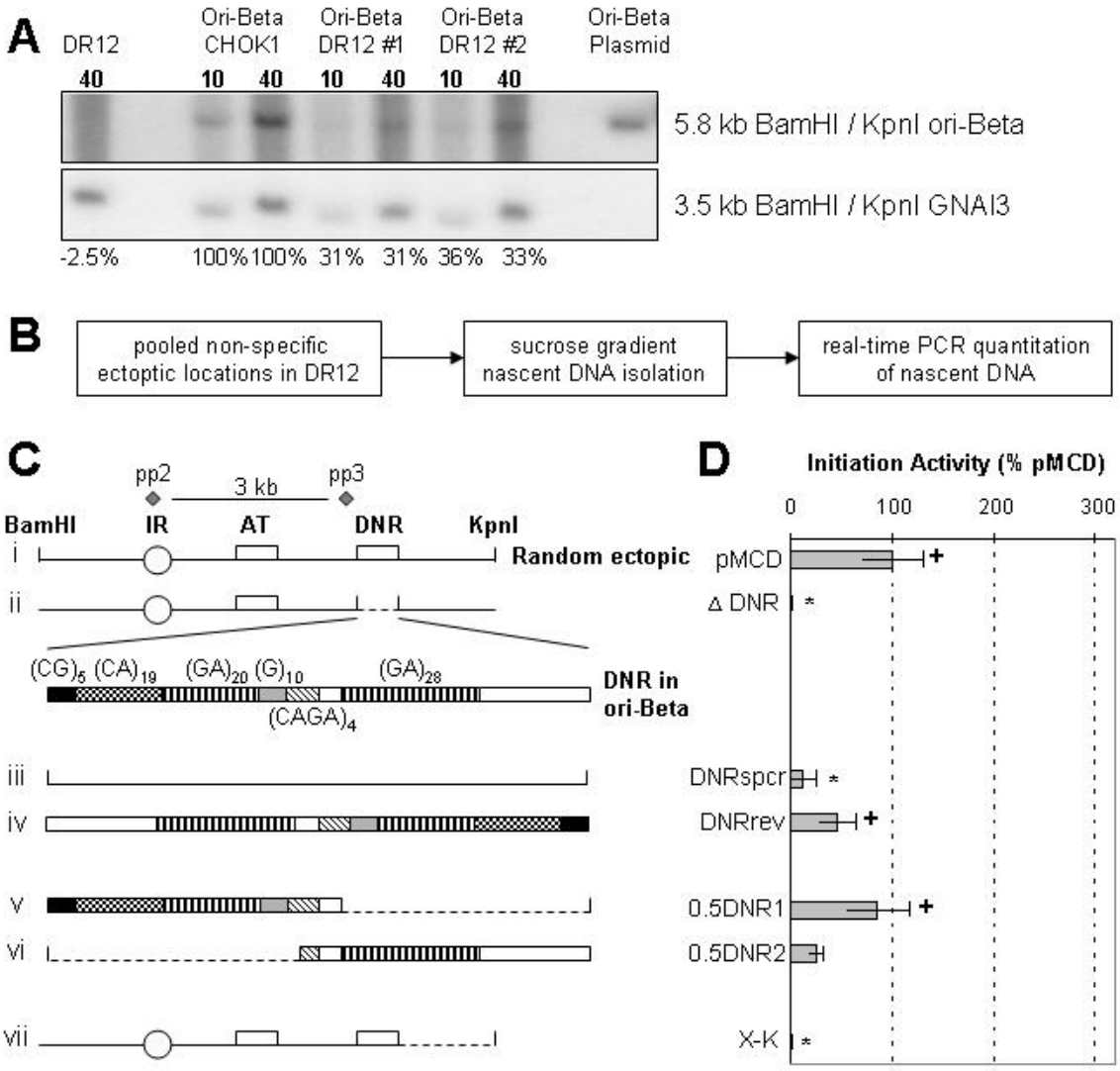
[c] SV40 intron amplicon, Genbank Accession # U02434

[d] Genbank accession # Y09885, positions relative to the BamHI site.

[e] nt numbers are relative to the 5S RNA gene transcriptional start site.

[f] nt numbers refer to Genbank accession #M12074 or the pUC18 multiple cloning site (MCS).

Figure 17. Initiation activity of ori-beta constructs at multiple ectopic chromosomal sites in pooled stably transfected DR12 cells. (A) Integration of full-length ori-beta into DR12 cells was assayed by Southern blot analysis of BamHI/KpnI-digested genomic DNA hybridized to ori-beta probes 1 and 2, or probe GNAI3 as a loading control. "DR12" refers to untransfected cells, "ori-beta CHOK1" refers to the endogenous ori-beta, "ori-beta DR12" refers to stably transfected ori-beta sequences in DR12 cells, and "ori-beta plasmid" refers to 30 pg of pMCD. Densitometry quantitation of ectopic ori-beta was first calculated in each lane as [(intensity of 5.8 kb band) / (intensity of 3.5 kb band)], then expressed as a percentage (shown below the blot) of the same densitometry quantitation done in the CHOK1 lane corresponding to the same μg quantity of DNA. (B) Summary of the methodology used to prepare and analyze nascent DNA samples. (C) Diagram of ori-beta constructs, with primer sets pp2 and pp3 shown (gray diamonds). (C, rows I, ii, and vii) Cartoon representation of the DHFR ori-beta, with deleted regions shown as dashed lines. (C, rows iii-vi) DNR replacement and partial deletion constructs, in the context of the entire ori-beta fragment, as described in the Materials and Methods. The DNR region and replacement elements are drawn to scale +/- 5 bp. All construct diagrams in (C) correspond to the adjacent name and bar graph in (D). In panel D, "initiation activity" is calculated as the abundance of target sequences detected with primer set pp2 in nascent-enriched DNA fractions, divided by those detected with primer set pp3. For comparison, all initiation activities are expressed as a percentage of the initiation activity of pMCD (unmodified 5.8 kb ectopic ori-beta, row i). Each bar is the average initiation activity measured from at least 3 independent transfections and nascent DNA preparations. Brackets indicate SEM. *, significantly lower than the pMCD construct; +, significantly greater than the ΔDNR construct (student's t-test, p value<0.05).


A
DR12
Ori-Beta CHOK1
Ori-Beta DR12 #1
Ori-Beta DR12 #2
Ori-Beta Plasmid
40
10
40
10
40
10
40
5.8 kb BamHI / KpnI ori-Beta
3.5 kb BamHI / KpnI GNAI3
-2.5%
100%
100%
31%
31%
36%
33%
B
pooled non-specific ectoptic locations in DR12
sucrose gradient nascent DNA isolation
real-time PCR quantitation of nascent DNA
C
pp2
3 kb
pp3
BamHI
IR
AT
DNR
KpnI
i
Random ectopic
ii
(CG)5 (CA)19
(GA)20 (G)10
(GA)28
(CAGA)4
DNR in ori-Beta
iii
iv
v
vi
vii
D
Initiation Activity (% pMCD)
0
100
200
300
pMCD
Δ DNR
DNRspcr
DNRrev
0.5DNR1
0.5DNR2
X-K

Results

Characterization of the DNR element. Deletion of DNR from the 5.8 kb DHFR ori-beta at random ectopic chromosomal sites in DR12 hamster cells resulted in a 9-fold decrease in initiation activity (7). To assess whether all or part of the DNR sequence was specifically required for full ori-beta activity or whether the DNR element simply maintained critical spacing between neighboring elements, we created further deletion and substitution mutants of DNR. Each construct was stably integrated at random locations in DR12 cells, and total genomic DNA from pooled transfected cells was isolated. Integration of the intact 5.8 kb BamHI/KpnI ori-beta fragment into the DR12 hamster genome was monitored by Southern blot using a fragment at the GNAI3 locus as a loading control. In two independent transfections, full-length ori-beta was detected at about 1/3 the level of the endogenous ori-beta in CHO-K1 cells (Figure 17A). Nascent DNA of ~0.4-2 kb was obtained by size fractionation of denatured genomic DNA from asynchronously growing, pooled transfected DR12 cells. The abundance of target sequence in the nascent DNA fractions was determined by real-time PCR using primer sets at the ori-beta initiation site (pp2) and at an outlying site ~3 kb away (pp3) (Appendix Table 3, Figure 17C).

Using 5.8 kb wild type (WT) ori-beta (pMCD) as a positive control and the DNR deletion mutant as a negative control, we measured the initiation activity of ori-beta mutant constructs (Figure 17D). Replacement of the 235 bp DNR element with a 235 bp fragment from an SV40 intron (DNRspcr) was not able to rescue initiation activity. Reversing the original orientation of DNR (DNRrev) reduced the initiation activity at ori-

beta, but it remained easily detectable (Figure 17D). These results indicate that some sequence-specific feature of DNR is critical for its role in ori-beta initiation activity at random ectopic locations in DR12 cells.

The 235 bp NheI-XbaI DNR sequence is composed of $GCTA(CG)_5(CA)_{19}(GA)_{20}(G)_{10}(CAGA)_4GGGAGAGAGGCAGAGAGGG(GA)_{28}$ followed by 44 bp of sequence without any repeats. *In vitro* studies of the DNR sequence showed that the $(CG)_5(CA)_{19}$ sequence can adopt a left-handed Z-DNA conformation, and the two $(GA)_n$ sequences can adopt triplex DNA structures (18). To further define functional regions within the DNR element, we deleted either the first half (0.5DNR2) or the second half (0.5DNR1) of the DNR sequence and tested the initiation activity of each of these ori-beta constructs in DR12 (Figure 17D, rows v-vi). The 0.5DNR1 deletion, which removed the $(GA)_{28}$ and downstream sequences, had only a minor reduction of initiation activity as compared to the WT ori-beta construct. In contrast, the 0.5DNR2 deletion, which removed the upstream half of DNR up to the CAGA repeats, reduced initiation activity to 25% of WT ori-beta. These results indicate that most of the functionality of DNR resides within the first half of its sequence, but that there is a smaller contribution from the second half.

To test if additional functional elements exist downstream of DNR, the entire ~1.3 kb XbaI-KpnI (XK) downstream region was deleted. This XK mutant showed initiation activity levels equivalent to the DNR deletion, indicating little or no initiation at ori-beta (Figure 17D, row vii). This result suggests that at least one additional ori-beta functional element is located downstream of DNR.

Two dissimilar transcriptional elements can independently replace the function of the DNR element. The importance of DNR in the initiation activity of ectopic ori-beta might also be explained by the activity of DNR as a replication fork barrier (RFB), since early studies of the DNR region revealed that it impedes the progression of replication forks emanating from an SV40 origin in an orientation-dependent manner (22). To test this possibility, the DNR element was replaced with a known RFB from the murine 28S rDNA locus (57, 114), such that it would impede replication forks moving toward ori-beta (Figure 18B, row ii). This 161 bp element contains a binding site for the Transcription Termination Factor 1 (TTF-1), which binds to the second Sal box (SB2) motif at the rDNA locus (65). In addition to its role in termination of transcription and replication elongation, TTF-1 also acts to remodel chromatin and activate transcription (98, 99). When the initiation activity of the SB2 replacement mutant was evaluated at ectopic sites in DR12 cells, ori-beta initiation activity was above that of WT ectopic ori-beta (Figure 18C, row ii), indicating that the SB2 element is able to functionally replace DNR.

The non-conserved distal Abf1-binding element in yeast ARS1 is known to limit the ability of nucleosomes to mask the ORC-binding site, thereby inhibiting ARS1 activity (112). Given the distal location of DNR in ori-beta, we wondered whether DNR might also position nucleosomes to facilitate protein binding to elements necessary for ori-beta activity, such as AT. To address this possibility, the DNR sequence was replaced with the 282 bp *Xenopus* 5S RNA gene, which contains a well-characterized nucleosome-positioning element (NPE) (188). The orientation of the 5S RNA gene replacement sets the direction of transcription away from ori-beta. The 5S RNA gene (5Sf) replacement of

DNR displayed initiation activity above that of WT ori-beta (Figure 18C, row iii), indicating that the 5S gene was able to functionally replace DNR.

We next sought to gain more insight into the mechanism of 5S RNA gene function as a DNR replacement. The 5S gene could stimulate ori-beta by positioning nucleosomes, by recruiting transcription factors via its Transcription Factor IIIA (TFIIIA) binding site, or by providing a specific DNA structural conformation (216). To distinguish among these possibilities, new ori-beta constructs were designed in which DNR was replaced by a 5S RNA gene lacking the TFIIIA binding site and downstream sequences, but with the NPE left intact. This partial 5S RNA gene was inserted in both orientations in place of DNR, and the initiation activity of the resulting ori-beta mutant constructs at random ectopic sites in DR12 cells was determined. Both constructs lacking the TFIIIA binding site had sharply reduced initiation activity relative to the complete 5S RNA gene replacement, with minor initiation activity detected in the forward, but not reverse, orientation (Figure 18C, rows iv-v). These results suggest that the functional replacement of DNR by the 5S RNA gene is dependent on its TFIIIA binding site, and not solely on its ability to position nucleosomes.

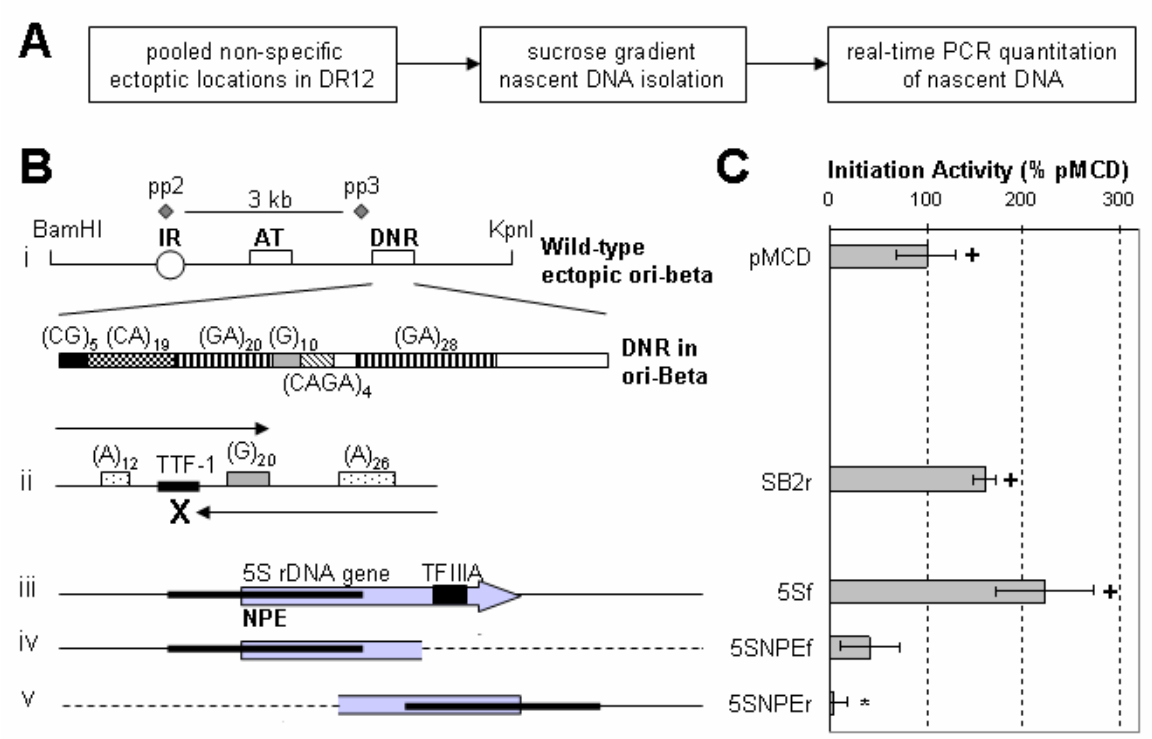


Figure 18. Loss of initiation activity in DNR-deleted ori-beta can be restored by transcriptional elements. (A) Summary of the methodology used to prepare and analyze nascent DNA samples. (B) Diagram of ori-beta constructs, with primer sets pp2 and pp3 shown (gray diamonds). The DNR region and replacement elements are drawn to scale +/- 5 bp. (B, row i) Wild type ectopic ori-beta in DR12 cells. (B, row ii) Arrows above and below the SB2 element indicate the polar impediment of replication forks moving toward ori-beta (X) mediated by TTF-1 at its binding site (thick line). (B, row iii) A gray arrow indicates the transcription direction of 5S RNA gene, with the nucleosome positioning element (NPE, thick line) and TFIIIA binding site (black box) highlighted. (B, row iv-v) Partial 5S RNA gene replacements containing the NPE (black bar) without the TFIIIA binding site, in either orientation. Dashed lines indicate deleted regions. All construct diagrams in (B) correspond to the adjacent name and bar graph in (C). In panel C, "initiation activity" and notations are as described in the legend of Figure 17.

The DNR element is required for ori-beta activity at a specific integration site in human cells. Our characterization of DNR used stably integrated ori-beta constructs at random chromosomal locations, but we cannot rule out the possibility that the requirement for various sequence elements in ori-beta activity depends on the random integration strategy that we used or on the hamster cell environment. In an attempt to

confirm the results of Figures 17 and 18 with ori-beta integrated at a specific chromosomal site in a different mammalian species, we generated a panel of ectopic ori-beta constructs at the same specific chromosomal site in the Hela acceptor cell line 406. Ectopic wild type and mutant ori-beta cell lines with the DNR deletion, DNR orientation reversal (DNRrev), and DNR substitution with an identically-sized DNA fragment spacer from an SV40 intron (DNRspcr) (Figure 19) were integrated at the same chromosomal site in the Hela acceptor cell line 406, so the resulting lines were completely isogenic except for the mutation introduced in each construct. After G418 and ganciclovir screening, the resistant clones were verified by diagnostic PCR analysis (Figure 19A). The acceptor cell line 406 showed the specific band using diagnostic primers 1 and 2. Conversely, when the FRT site was occupied with a large ori-beta DNA fragment mediated by FLP recombinase, primers 1 and 2 did not amplify the corresponding band. But primers 1 and 3, located within the Neomycin resistance gene, amplified the expected band in all positive ori-beta clones. Southern blot analysis of EcoRI-digested genomic DNA from of these clones used two probes, Hyg and Neo, to verify the integration location and integrant copy number (Figure 19B). Hybridization of the Hyg probe to a 5.6 kb EcoRI fragment confirmed that integration occurred at the intended chromosomal location (Figure 19B, left). Hybridization of the Neo probe solely to the 5.6 kb band in the ori-beta lines indicated that specific integration in the acceptor FRT site had taken place (Figure 19B, right).

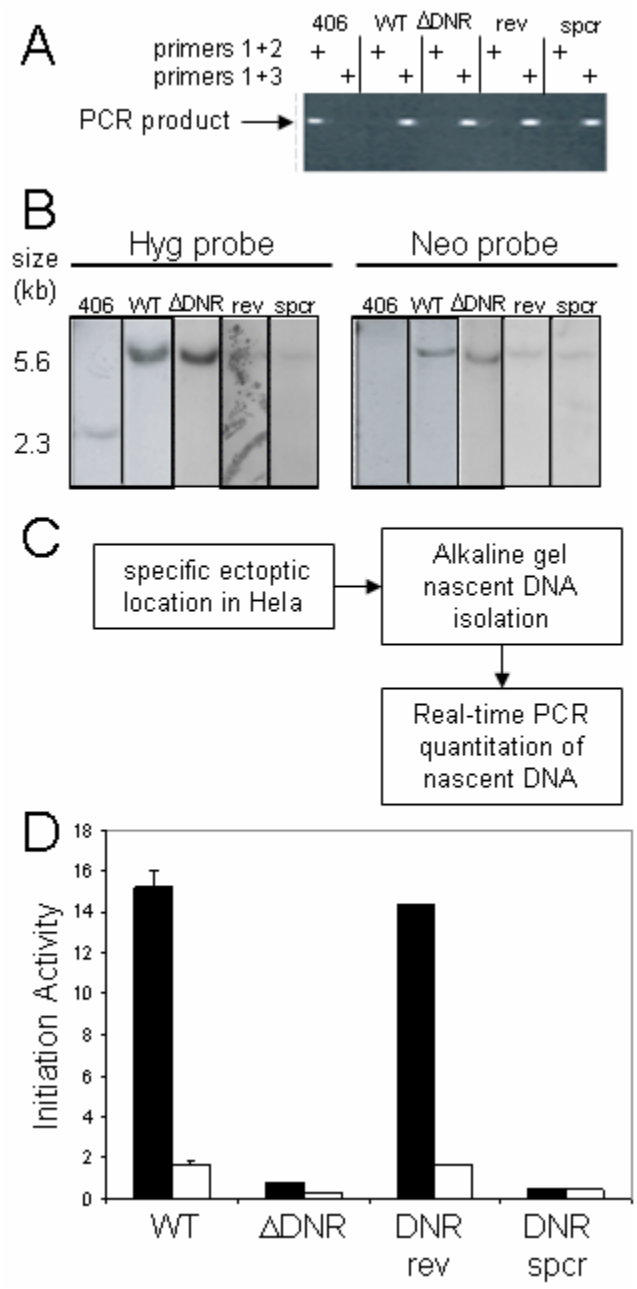


Figure 19. Initiation activity of mutant DNR ori-beta constructs at the specific chromosomal site in Hela. (A) Following transfection with ori-beta constructs and drug selection, genomic DNA from clonal cell lines was tested for site-specific integration by PCR. A PCR product using primers 1+2 indicates no integration at the FRT site, but a product using primers 1+3 indicates integration in the correct orientation. (B) Southern blots EcoRI-digested genomic DNA from clonal lines was hybridized to either the Hyg probe or Neo probe as described in the Materials and Methods. The Hyg probe hybridizes to a 2.3 kb fragment in the acceptor cell line and to a 5.6 kb fragment when ori-beta is integrated. The Neo probe hybridizes to a 5.6 kb fragment only in the integrated ori-beta. (C) Summary of the methodology used to prepare and analyze nascent DNA samples as previously described (113). (D) Analysis of the initiation activity of ori-beta constructs at the specific integration site in Hela. 406, Hela acceptor cell line; WT, full-length ori-beta fragment; ΔDNR, DNR deletion construct; DNRrev, construct in which the DNR orientation was reversed; DNRspcr, replacement of DNR with a 235 bp SV40 intron fragment. "Initiation activity" is defined as the abundance of target sequences in nascent DNA-enriched fractions detected at primer sets pp2b (black bars) or ppHyg (white bars), divided by that at primer set ppGlobin. Brackets indicate SEM for WT.

Once each ori-beta construct was successfully integrated at the specific site in Hela 406, the initiation activity of ori-beta in each cell line was measured by real-time quantitative PCR using three primer sets, pp2b, ppHyg, and ppGlobin (Figure 16, Appendix Table 3). Primer set ppGlobin, located at the human beta-globin locus at

chromosome 11, was used as an internal reference to compare different cell lines and nascent DNA preparations. In the WT ori-beta line, the abundance of target sequences in nascent DNA detected with pp2b was ~15-fold higher than that detected with ppGlobin and ~10-fold higher than that with the outlying ppHyg (Figue 4D), indicating that ori-beta is active at the specific ectopic site in Hela 406 cells. Deletion of the DNR element or substitution with a spacer abolished ori-beta initiation activity, confirming that the specific sequence of DNR is critical to its role in regulating DNA replication initiation (Figure 19D). Reversing the original orientation of DNR did not diminish replication initiation at ori-beta (Figure 19D).

Other ori-beta elements contribute to initiation activity in human cells. Full initiation activity of the ectopic ori-beta replicator in DR12 hamster cells was shown to require specific DNA sequences in an AT rich element (AT), in the downstream RIP60 binding site (RIP), and in the bent DNA element (BENT) (8). To assess the importance of these elements for the initiation activity of ori-beta at random ectopic chromosomal sites in human cells, Hela cells were stably transfected with each of the mutant constructs (Figure 20). Fractions enriched in 0.4-2 kb nascent DNA were prepared from uncloned pools of these transfected cells and the abundance of four ori-beta target sequences in each fraction was determined by competitive PCR. Initiation activity was calculated by normalizing the abundance of each target to that of an outlying target 3 kb away from the initiation site (pp3). In Hela cells, initiation activity of wild-type ectopic ori-beta at pp2 was about 40-fold greater than that pp3 (Figure 20C). In contrast, the initiation activity of ectopic ori-beta lacking the AT element was greatly reduced. Moreover, the mutations of downstream RIP60 binding site (RIPx) and the bent DNA structure (BENDx)

decreased the initiation activity at pp2 to a level similar to that at pp3. These results are consistent with those obtained using a similar strategy in DR12 hamster cells (8).

We then examined the activity of three ori-beta mutants integrated at a specific ectopic chromosomal site in Hela 406 cells. Integration of these ori-beta constructs at the specific integration site was confirmed by PCR and Southern blot analysis (Figure 21A, 6B). Compared to the WT ori-beta control, the AT element deletion, RIPx mutation, and BENDx mutation each virtually eliminated ori-beta initiation activity at the pp2b site, with no enrichment seen above that of the ppHyg and ppGlobin sites (Figure 21D). The results confirm that these three elements are essential for ori-beta activity in human and hamster cells, at either random or specific ectopic chromosomal sites.

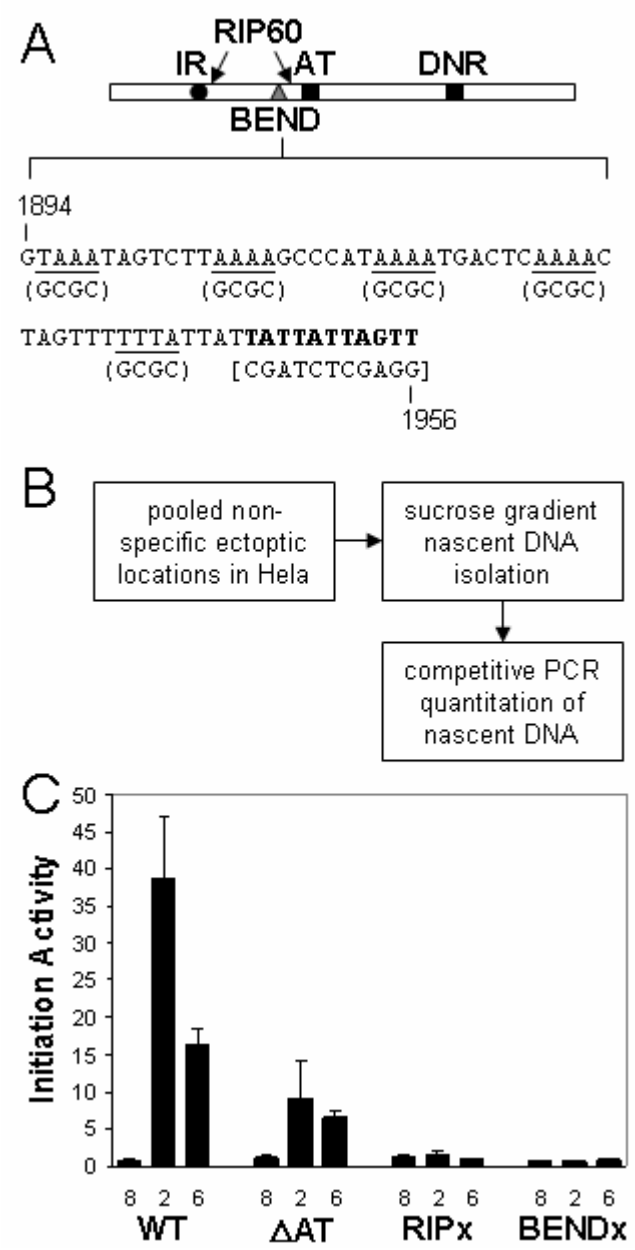


Figure 20. Initiation activity of ori-beta constructs at multiple ectopic chromosomal sites in pooled stably transfected Hela cells. (A) Diagram of WT ori-beta and mutated residues. The sequences necessary for the bent DNA region (underlined) and downstream RIP60 binding site (bold) are shown. Nucleotide substitutions for the BENDx and RIP60 mutants are in parentheses and brackets, respectively. (B) Summary of the methodology used to prepare and analyze nascent DNA samples as previously described for hamster cells (8). (C) "Initiation activity" was calculated by dividing the quantity of target sequences in nascent-enriched DNA fractions detected at primer sets pp8, pp2, or pp6 by that detected at primer set pp3. Primer set locations are given in Figure 16B. Construction of these ori-beta plasmids is described in (7, 8). Bars are the average quantitation of target sequences in at least 2 independent transfections and nascent DNA isolations. Brackets indicate SEM.

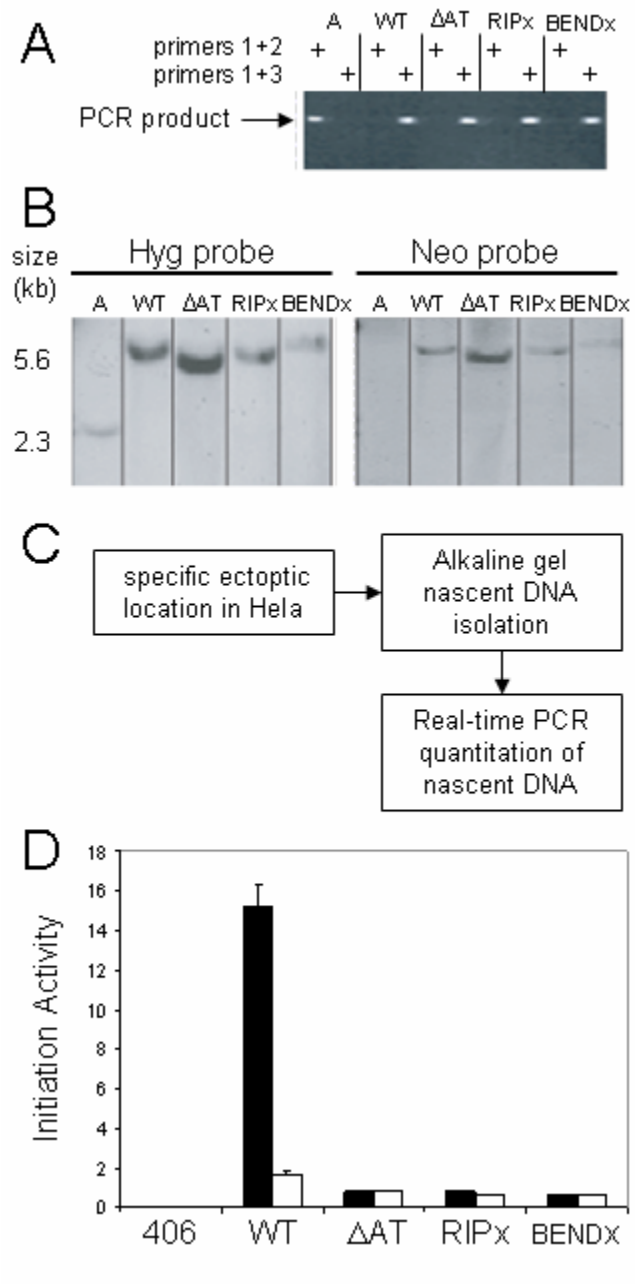


Figure 21. Initiation activity of ori-beta constructs at the specific chromosomal site in Hela cells. (A) PCR-based analysis of clonal cell lines containing ori-beta constructs, as described in the Materials and Methods section and Figure 19A. (B) Southern blot analysis of EcoRI-digested genomic DNA from clonal cell lines, as described in the Materials and Methods section and Figure 19B. (C) Summary of the methodology used to prepare and analyze nascent DNA samples as previously described (113). (D) Analysis of the initiation activity of ori-beta constructs at the specific integration site in Hela. 406, Hela acceptor cell line; WT, full-length ori-beta fragment. Mutant constructs are described in the Materials and Methods and Figure 20. "Initiation activity" is defined as the abundance of target sequences in nascent-enriched DNA fractions detected at primer set pp2b (black bars) or ppHyg (white bars), divided by that at primer set ppGlobin. Brackets indicate SEM for WT.

Discussion

Ori-beta constructs can be evaluated using different ectopic integration strategies. In this report, we have measured the Chinese hamster DHFR replicator ori-beta DNA initiation activity in Hela and DR12 cells at specific and at random ectopic chromosomal sites. Consistent with previous results (7, 8), our investigation showed that the 5.8 kb DHFR initiation region ori-beta is active at ectopic sites not only in hamster cells, but also in human cells (Figures 17 and 20). Mutations that disrupted ori-beta activity in DR12 or Hela using the non-specific integration system also disrupted initiation in the specific site in Hela (compare Figures 17 and 19, 20 and 21). These results provide evidence that the potential pitfalls of non-specific integration, including position effects, do not obscure or falsify the determination of ectopic origin function.

Transcription through the endogenous ori-beta has been shown to inactivate the origin (126, 160). However, transcription of the neomycin resistance gene upstream of the site-specific integrated origin constructs in the Hela 406 line should not be a concern, since a strong SV40 transcription terminator is located between the neomycin resistance gene and the integrated origin fragment (16, 118). Therefore, transcription elongation across the ectopic ori-beta would not be expected to suppress its activity, and this is supported by the ori-beta initiation activity seen in the integrated WT ori-beta construct (Figure 19).

The 3' end of the 5.8 kb ori-beta fragment contains at least two sequence elements necessary for ori-beta initiation activity. Deletion of either the 235 bp DNR element or the 1.3 kb XK region downstream of DNR led to a loss of ori-beta initiation

activity at non-specific chromosomal sites in DR12 cells (Figure 17D), and the effect of the DNR deletion was recapitulated at a specific locus in Hela cells (Figure 19D). The loss of initiation activity upon deletion of the 5' half of DNR, but not the 3' half (Figure 21D), suggests that the main functional component(s) of DNR resides within the first 132 bp of its 235 bp length. It also suggests that the loss of initiation activity upon deletion of the XK region is not due to the disruption of a necessary element that overlaps the DNR and XK, but rather because at least one separate and novel functional element exists within the 1.3 kb XK element. The failure to functionally replace the 235 bp DNR element with a 235 bp DNA fragment from an SV40 intron (Figure 17D and 19D) indicates that DNR does not serve just to maintain a critical spacing between flanking elements, but that its DNA sequence is important.

Transcription factor binding sites can functionally substitute for DNR. The specific sequence of DNR, in particular the first 132 bp of its 235 bp length, serves some functional role in the initiation of DNA replication at ori-beta from over 3 kb away from the initiation site. What functional role might it serve? This 132 bp sequence is as follows: GCTAG$(CG)_4(CA)_{19}(GA)_{20}(G)_9(CAGA)_4$GGGAGAGAGGCAGAGA. The first 6 bp are a palindromic NheI site, and the rest of the sequence, except for the final 16 bp, is repetitive. These final 16 bp, however, were not sufficient for full initiation activity of an ori-beta mutant lacking the repetitive sequences (Figure 17D, 0.5DNR2 mutant). GA repeats have been shown to impede replication forks (22, 80, 153) and affect the assembly of DNA into nucleosomes (46), so both of these potential roles were investigated separately (Figure 18). Importantly, a replication fork barrier (SB2) and an

element that contains a nucleosome-positioning element (5S rDNA) were each able to independently replace the function of DNR (Figure 18C).

The results implied that these two elements share some functional characteristic with the DNR element. The protein TTF-1, which binds to the SB2 element, is capable of actively re-positioning nucleosomes (98), leading us to suspect that SB2 and the 5S rDNA substituted for a function of DNR in positioning nucleosomes. However, since the NPE from the 5S gene could not, without the TFIIIA binding site, functionally replace DNR (Figure 18C), we conclude that SB2 and 5S rDNA must have additional or alternative functions needed to substitute for DNR in ori-beta activity.

The general role of transcription factors is to regulate the formation of various functional complexes at specific chromosomal sites assembled from elements recognized by these factors (93). The functions of transcription factor binding sites may thus resemble those of replicator elements needed to assemble ORC and pre-replication complexes. Comparison of the two transcription elements that substitute for DNR in the ectopic ori-beta replicator may reveal the function of DNR. The SB2 element contains binding sites for TTF-1 and Ku 70/80 heterodimer. These proteins participate in initiation of rDNA transcription (69, 99), termination of transcription and replication (151, 200), and TTF-1 can also remodel chromatin (98, 99). The 5S gene contains a binding site for TFIIIA, which recruits TFIIIC and TFIIIB to form a stable transcription activation complex on the 5S gene (100, 216). Thus both transcription elements able to substitute for DNR share the ability to gain access to a specific site in chromatin and nucleate assembly of a multiprotein complex at that site. This type of transcription element-dependent chromatin remodeling, in the absence of transcription, facilitates assembly and

function of the V(D)J recombination complex at specific times in lymphocyte development (140, 170). We speculate that both the 5S gene and the SB2 element may substitute for DNR by creating a suitable chromosomal environment that is important for the essential functions of the flanking AT or XK elements in ori-beta replicator activity.

The function, but not necessarily the DNA sequence, of *cis*-acting elements in a mammalian replicator promotes initiation of DNA replication. One of the obstacles to understanding mammalian origins has been the lack of identifiable sequence homology between different origins. As shown here, the repetitive sequences in DNR are critical for initiation activity of ectopic ori-beta (Figure 17 and 19, (7)), but CA and GA repeats are not commonly associated with replication origins. Moreover, we have shown that the DNR element can be functionally substituted by two different unrelated sequences, and not by a third sequence (Figure 17, 18). These data provide additional support for the hypothesis that *cis*-acting elements that together comprise a mammalian chromosomal replicator can have different DNA sequences in different replicators, yet fulfill the same function in origin activity. These data also extend our previous finding that the AT-rich element of ori-beta can be functionally replaced by a non-homologous sequence from the LaminB2 origin (8) by showing that mammalian transcriptional control elements can also substitute for a mammalian replicator element. Detailed characterization of additional mammalian replicators, mutant replicators, and the proteins that interact with them will be needed in the future to uncover the common functions of replicator elements that differ in sequence.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

The Implications of an Origin of DNA Replication at the Human FMR1 Promoter

Although several other origins have been reported near TNR disease loci, the work presented in this thesis is to date the most comprehensive and definitive characterization of an origin associated with a trinucleotide repeat tract (28, 136). Our results also demonstrate that the CGG repeats affect PCR efficiency at neighboring sequences, and this may be a feature common to TNRs. If uncut genomic DNA is used as a standard for real-time quantitative PCR, NarI/XhoI-digested genomic DNA will show an artificial "peak" of abundance near the repeats (data not shown). This occurs because the XhoI/NarI-digested genomic DNA without the repeats amplifies with higher efficiency near the repeats when compared to the uncut genomic DNA standard. Since the experimental techniques used to identify the other candidate TNR origins employed PCR-based assays to detect initiation events, their results should be verified with this consideration in mind.

Recombination- or repair-mediated repeat instability. Models of repeat instability involving DNA replication are not exclusive of models involving DNA repair and/or recombination after fork stalling or collapse (97, 104). Polymerase stalling within the repeats, or DNA secondary structures left after the passage of the replication fork, could trigger a DNA damage response and recruit DNA repair machinery. Secondary

structures formed within the repeats might then inhibit proper processing of the broken or damaged DNA. These errors in processing could eventually lead to single-stranded gaps or double-stranded breaks (62, 104). Alternatively, the delay in replication timing seen at Fragile X alleles (71, 72, 181, 218) could cause the locus to sometimes remain unreplicated by the end of S-phase. In this case, a DNA damage response would be triggered to replicate the region, possibly leading to the same problem of improper processing of the CGG repeats.

The potential role of the FMR1 origin as a safeguard against repeat expansion. Investigations into the role of origins in affecting TNR stability have relied on model systems to determine what might occur at TNR loci. The discovery of an origin of DNA replication near the Fragile X CGG repeats reinforces the relevance of these investigations in model systems and expands our understanding of TNR stability beyond hypothetical models to include an actual disease locus. The extensive knowledge gained from model systems can now be applied to the relationship between the FMR1 origin and CGG repeats.

As discussed in chapter 2, the position of the FMR1 origin should favor CGG contractions over expansions, according to numerous investigations in model systems (reviewed in (104). Both expansions and contractions occur when unstable CGG repeats are integrated next to an origin in *S. cerevisiae* and COS-1 cells, but contractions are favored when CGG is the template of lagging strand synthesis (13, 137). Random expansions and deletions may occur periodically at the FMR1 CGG repeats through normal DNA metabolism, but the location of the FMR1 origin relative to the repeats should limit the number of expansions. Thus, the FMR1 origin placement may have

evolved as a safeguard mechanism to maintain genomic integrity by limiting expansion events at the CGG repeats. The reason for the failure of this mechanism in Fragile X patients will require further investigation. Also, validation of the hypothesis of origin placement as a regulator of repeat stability will require the discovery of additional origins near other TNR loci.

Future Directions for the FMR1 Origin

The comprehensive and conclusive identification of the FMR1 origin provides a solid example of an origin normally associated with a disease-relevant TNR. The FMR1 origin also provides a way to move beyond model replication systems and study how the CGG repeats are normally replicated in their natural setting. Cells can be grown under conditions of replicative stress, such as reduced nucleotide pools (11), to determine whether dormant secondary origins might exist at other locations near the CGG repeats. This approach may give a clue to how the CGG tract might be replicated during early development when replication initiation is hypothesized to be relatively non-specific.

Cloning the origin and moving it and the CGG tract to an ectopic site could provide a wealth of information about how replication from the FMR1 origin affects repeat stability. If the FMR1 origin is positioned to stably replicate the CGG repeats as we propose, changing the distance or orientation of the CGG repeats relative to the FMR1 origin should disrupt that stability. By moving the origin to an ectopic site, such mutations could be engineered.

A systematic deletion of transcription factor binding sites in the FMR1 promoter would indicate which of these factors, if any, contribute to the activity of the FMR1 origin. Thus, one could use the well-characterized FMR1 promoter to further elucidate a potential link between transcription and replication regulation. Integrating replication timing studies into this investigation might uncover the reason for delayed replication of the FMR1 locus upon FMR1 silencing.

The strategy and data presented in chapter 2 should provide a framework for future investigations of potential origins at other TNR disease loci. Although other studies suggest that origins may be found near other disease-relevant TNRs, a more thorough investigation should uncover whether origins are a common feature associated with these loci and the extent of the role that origins may play in stabilizing these and other repeat tracts in the genome.

Finally, the most important questions left unanswered from this investigation are whether a change in origin usage at the FMR1 locus actually occurs during development, and whether unmethylated, expanded CGG repeats undergo contractions because they are replicated from the FMR1 origin. The 22 week fetal fragile X cell line GM07072 had a reduction in FMR1 origin activity, suggesting a developmental change in origin usage. Untransformed fetal cells are available from the Coriell cell repository as young as 8 weeks old. If the reduction in FMR1 origin usage in early developmental cells is also seen in these fetal cells, it would considerably strengthen our hypothesis regarding the involvement of FMR1 origin usage in CGG stability *in vivo*.

Discussion of the Ectopic Integration Strategy and the Genetic Dissection of the DHFR Origin-Beta

Ectopic origin placement: the mammalian version of the yeast ARS assay. Understanding the workings of origins of DNA replication is likely to involve a complex and dynamic set of interactions, possibly including the regulation of replication factor binding, chromatin structure, epigenetic modifications, nuclear localization, association with the nuclear matrix, DNA flexibility, and transcription (4, 27, 37, 40, 61, 115, 123). Our current understanding of regulatory elements at mammalian origins stems from a limited number of often disconnected case examples, which hampers firm conclusions of exactly what is needed for DNA replication to initiate at specific chromosomal locations. Even though specific DNA sequences were shown to be necessary for replication to initiate, the lack of a clear consensus sequence in the origins of higher eukaryotes has made a bio-informatics approach to origin identification problematic (4, 60, 61). The report of a possible loose metazoan consensus sequence (149) is provocative but will require further investigation. For now, in order to test the dynamic interactions required for replication initiation, ectopic systems for origin mutational analysis are critical, just as the use of ectopically-expressed reporter genes have been essential for elucidating the components of gene promoters. The use of stably-integrated origin fragments, either site-specific or random, new ARS-like episomal analyses of origin function (84, 149, 183), and eukaryotic viral origins (202) each have different benefits and caveats. A rigorous comparison of these systems should facilitate the efficient progression of our understanding of origins of DNA replication.

The results and conclusions presented in Chapter 3 demonstrate that the DNA sequence components required for origin activity can be genetically investigated using stable ectopic integration of origin constructs at either non-specific or specific integration sites. If these methods are largely equivalent, as the data in Chapter 3 suggest, the non-specific integration system represents the quickest and technically easiest method for testing the effects of genetic mutations within the origin fragment. However, the non-specific integration system has the disadvantage that the origin copy number and chromatin context in each cell is variable, making some biochemical analyses of ectopic origin fragments potentially problematic. Within a pool of randomly integrated origin fragments, some origins may be active and some may be inactive, but initiation activity is measured as the average initiation activity at all locations. Assaying for DNase or MNase hypersensitive sites, for example, would be easier to interpret at a specific integration site where all origins are isogenic. When assaying such sites in a mixed population of active and inactive origins in different chromosomal contexts, the results may be more difficult to interpret.

The human laminB2 origin and hamster DHFR ori-beta are capable of functioning at multiple ectopic chromosomal locations (7, 8, 141), suggesting that elements necessary for creating the required chromosomal context are contained within the ectopic origin fragment. This may not be the case with other cloned fragments of mammalian origins, and more case examples will be required before the viability of the non-specific ectopic system is accepted as a general test applicable to origins. However, since we directly compared the specific and non-specific systems for analyzing ori-beta constructs for initiation activity and found them to be equivalent, we have bridged a gap between the

investigations of ori-beta and the investigations of the c-myc and beta-globin origins at specific chromosomal locations.

Replacement of origin *cis*-elements with functional elements containing non-homologous sequences. Mammalian origins of DNA replication are comprised of multiple *cis*-acting functional sequence elements (4). As discussed in Chapter 3, some investigations of mammalian origin functional elements have successfully swapped non-homologous sequences that can function independently to promote origin activity (8, 58). The successful substitution of the DNR element with *Xenopus* and murine ribosomal RNA transcriptional elements illustrates the concept that transcription factors do more than just promote transcription (93). Another illustration of this concept is seen during V(D)J recombination, where transcription factor binding, but not transcription, is required for recombination (140, 167). Although these transcription elements were able to substitute for DNR, at this point it is unclear what role they, or DNR, are playing in promoting replication initiation at ori-beta.

Future Directions for the DHFR Origin-Beta

What is DNR doing? Although DNR can be functionally replaced with the 5S RNA gene or the SB2 element (Figure 18), it is still unclear what these three elements have in common. Of course, it is also formally possible that these elements have nothing in common and are promoting ori-beta activity by completely different mechanisms. A clarification of these results could be obtained by replacing DNR with only the TFIIIA or TTF-1 binding sites, to determine if the functional replacement of DNR is due to these

transcription factors alone. The DNR replacement data suggests that the first 132 bp of the DNR element may be binding some type of *trans*-acting protein factor that has a similar function to TFIIIA, TTF-1, and/or Ku70/80. To understand what DNR is doing, it would be very helpful to know what, if any, proteins are binding to DNR.

Possible DNR binding proteins. GA repeats in approximately 250 *Drosophila* genes serve as the binding site for the *trithorax*-like GAGA-factor (GAF) protein, a transcription activator/repressor and chromatin-remodeling factor (101, 102, 192, 209). GAF has been described as "establishing the transcriptional potential of genes without necessarily leading to gene expression" (206). It also plays an important role in proper chromosome segregation during mitosis (17). The functions of GAF appear to overlap with TTF-1, and its role in DNR to create a favorable chromatin environment is plausible. Although a clear homolog for *Drosophila* GAF has not been described, the hamster and mouse protein Pur1 binds GA repeats and is a potent transcriptional activator of the rat insulin promoter in hamster, mouse, rat, and human Hela cells (91). The human transcription factor MAZ is the homolog of Pur1 with 94% sequence identity (178). Interestingly, MAZ (Myc-associated Z-finger protein) binds to the c-myc gene P1 and P2 promoter regions within and flanking the well-characterized c-myc origin (20, 41). Pur1 or MAZ would be a good candidates for potential DNR-binding proteins, and future investigations should begin with these two proteins.

APPENDIX TABLE 3. Real-time PCR primers

Primer name	Sequence (5' to 3')	[Mg^{++}]	Ta*	Location**	Size***
MCM4 (Genbank accession # U63630)					
EX6b-F	TACCTGTGGGTAAGAGATGAGTTG	3.5 mM	60°	10691	209 bp
EX6b-R	CTCTATACATGCAACGACTTGGG			10900	
UPR4-F	GGACATTACAGATGCATTTCTC	3 mM	60°	13119	215 bp
UPR4-R	AAGAGTTCCAAGTTTGTTCCTC			13334	
UPR5-F	CCCCCGACCGCGTTTGTCGTTT	3 mM	64°	13284	198 bp
UPR-R1	GGCCAGTAAGCGCGCCTCTTTGG			13482	
LaminB2 (relative to replication start site (2))					
LB2C1-SX	GTTAACAGTCAGGCGCATGGGCC	3.5 mM	60°	-3931	240 bp
LB2C1-DX	CCATCAGGGTCACCTCTGGTTCC			-3691	
#BE-SX	ACTTTCTGAAGGAGGCTCTG	4 mM	60°	-1623	192 bp
#BE-DX	GGCAATAGCTCACCGTTTAC			-1431	
LO-F	GCGTCACAGCACAACCTGC	3.5 mM	60°	-136	266 bp
LO-R	TCTTTCTTAGACATCCGCTTC			+130	
B48-SX	TAGCTACACTAGCCAGTGACCTTTTCC	3.5 mM	60°	+173	167 bp
B48IIb-DX	GACTGGAAACTTTTTTCTACAAC			+6	
SE17-SX	AATTTCCACTCCACAGCCGT	2 mM	64°	+2532	208 bp
SE17-DX	TAGGAGTGACCCTAGTGACT			+2324	
Fragile X (Genbank accession # L29074)					
FraX-30 SX	CCTGAAAGGATAAATCTTGCCC	3.5 mM	60°	35023	191 bp
FraX-30 DX	CCTATGTTCAGATCTTCAACCC			35214	
FraX-20 SX	AACCCTATTTCTGGCTCTGGAC	3.5 mM	60°	24999	185 bp
FraX-20 DX	CCCTCCAAGTTATCAAAGATTCC			25184	
FraX-16 SX	AGTCAATTATTGGGGTCAACCAC	3 mM	60°	17238	205 bp
FraX-16 DX	GGGCAGAAATCAGAAACTCCAG			17443	
#FraX-14 SX	GGTCTGCACTGATGGAAGAAC	3.5 mM	64°	16070	182 bp
#FraX-14 DX	TAACACTTCTCCACAGCAGTTC			16252	
FraX-12 SX	TTTCAGGAAGACCCTAACATGG	3.5 mM	60°	14658	199 bp
FraX-12 DX	GAAGTTTCATGGCATATATTTAGG			14857	
#FraX-11.5 SX	TGAGCTGGGGATGGGCGAGG	2 mM	64°	14276	182 bp
#FraX-11.5 DX	AGGAGGCGGCCCGGCTGAAG			14458	
FraX-11 SX	TCCAATGGCGCTTTCTACAAGG	3 mM	62°	13922	244 bp
FraX-11 DX	GCTGGTCTCTCATTTCGATAGG			14236	
FraX-1c SX	TCTGTCTTTCGACCCGGCAC	3 mM	64°	13527	211 bp
FraX-1c DX	GGAAGTGAAACCGAAACGGAG			13738	
FraX-1d SX	GCGCGTCTGTCTTTCGACCC	3.5 mM	64°	13529	224 bp
FraX-1d DX	CCCTCCACCGGAAGTGAAACC			13753	

APPENDIX TABLE 3. (cont.)

FraX-1.2 SX	TCTGCAGAAATGGGCGTTCTGG	3 mM	64°	13425	178 bp
FraX-1.2 DX	CTCTCTCTTCAAGAGGCCTGGG			13603	
FraX-1.4 SX	CCTATTCTCGCCTTCCACTCC	2 mM	62°	13247	188 bp
FraX-1.4 DX	ATTTCTGCAGAGGTGCACTCA			13435	
FraX-1.5 SX	TCCCGACTCAATCCATGTCCC	2 mM	62°	13051	188 bp
FraX-1.5 DX	AGGCGAGAATAGGGGTGAAGG			13239	
FraX-2 SX	AGCATCCCGAAGGGAACATGG	2 mM	58°	12747	189 bp
FraX-2 DX	ACTGTATGTGCACCCTGTGCC			12936	
#FraX-2.5 SX	GAATGTGGCCCTAGATCCACC	3.5 mM	60°	12331	205 bp
#FraX-2.5 DX	AGAATAGGCGCTTCCATGATGG			12536	
FraX-3 SX	CTGAAAACCTTAAGGTGCAGGG	3.5 mM	60°	11963	185 bp
FraX-3 DX	CTCCAAGTGTAAGCTGTTGTTC			12148	
#FraX-3.5 SX	CTAGATGCCCGATCAGTAGGG	3.5 mM	60°	11561	223 bp
#FraX-3.5 DX	TGCAATATGTTCGATAGAGATCC			11784	
FraX-5 SX	AACAGTGCTGGAATAACTGGAC	3.5 mM	60°	9998	197 bp
FraX-5 DX	CTTCACTTCATTCCAGTGCATG			10195	
#FraX-6 SX	TCTGCTGACACTGTAATGGTGG	3.5 mM	60°	9598	184 bp
#FraX-6 DX	GTACTTCCTAATCTCAATCTCCCTCCTC			9782	
#FraX-6.5 SX	AGGAAGTCTCACTCATTGCTGG	3.5 mM	60°	9056	189 bp
#FraX-6.5 DX	GGCTCCTGTAAACTTTCTTGTAC			9245	
#FraX-7 SX	GTAACAGATCTGAGCAGACACC	3.5 mM	62°	8870	199 bp
#FraX-7 DX	GAGTGAGACTTCCTGTTACTCC			9069	
FraX-10 SX	ATACATGCAAAGGGCTAGGTCC	3.5 mM	60°	1024	180 bp
FraX-10 DX	GATCTCTCTATGCTTCAGTTTCC			1204	

Hamster DHFR ori-beta (Relative to BamHI site at position 1)

#pp12-SX	GAGCAGCAAGGAACTGAAG	3.5 mM	58°	16	246 bp
#pp12-DX	GTGTATTACTGCACAGTAGCA			262	
pp8-SX	CTCTCTCATAGTTCTCAGGC	3.5 mM	58°	470	200 bp
pp8-DX	GTCCTCGGTATTAGTTCTCC			670	
#pp2a-SX	GTCCCTGCCTCAAAACACAA	3 mM	58°	1070	122 bp
#pp2a-DX	TAGTGCGTCTTTAAGACCTG			1192	
pp2-SX	GTCCCTGCCTCAAAACACAA	3.5 mM	58°	1070	278 bp
pp2-DX	CTGCCTTCATGCTGACATTTGTC			1348	
pp2b-SX	GCACTAGATGCTGAACTTAACAG	3.5 mM	58°	1184	164 bp
pp2b-DX	CTGCCTTCATGCTGACATTTGTC			1348	
pp6-SX	AACTGGCTTCCCAAGAAATT	3.5 mM	58°	1517	168 bp
pp6-DX	AACCTCTGAACTGTAAGCTG			1685	
pp3-SX	GGACACTAAGTCTAGGTACTACA	3.5 mM	58°	3882	258 bp
pp3-DX	GCTGGGATAAGTTGAAATCC			4140	
#ppK12-SX	CAGGACCAATGTGATACAAC	3 mM	58°	5389	185 bp
#ppK12-DX	AGCTTAAGGCTCACTTATGG			5574	

APPENDIX TABLE 3. (cont.)

Human beta-Globin locus (Genbank #GI455025)					
#Globin40.9-SX	GCAAGCAATACAAATAAT	3 mM	58°	~40900	156 bp
#Globin40.9-DX	ACCACAAACACAAACAGG			~40900	
Globin54.8-SX		3.5 mM	60°	~54800	
Globin54.8-DX				~54800	
#Globin62-SX	GAGGTACGGCTGTCATCACTT	3 mM	58°	62005	210 bp
#Globin62-DX	CAGGGCAGTAACGGCAGA			62214	
#Globin62.2-SX	CCTGAGGAGAGGTCTGGCGT	2 mM	60°	62220	207 bp
#Globin62.2-DX	CCTAAGGGTGGGAAAATAGACC			62427	
#Globin62.4-SX	GGTCTATTTTCCCACCCTTAGG	3 mM	60°	62406	209 bp
#Globin62.4-DX	CAGTGCAGCTCACTCAGTGT			62615	
#Globin72-SX	CCAGAATCTACAATGAACTC	3.5 mM	60°	72319	135 bp
#Globin72-DX	TGATGGCTAGTGATGATG			72453	
FRT integration site (113)					
ppHyg-F	TGCTCCGCATTGGTCTTGA	2 mM	60°		73 bp
ppHyg-R	TGCGCCCAAGCTGCAT				
ppNeo1-SX	GAAGGGACTGGCTGCTATTGG	2 mM	60°		201 bp
ppNeo1-DX	TGATCGACAAGACCGGCTTCC				
ppTK-F	AGCAAGAAGCCACGGAAGTC	2 mM	60°		102 bp
ppTK-R	GTTGCGTGGTGGTGGTTTTC				
ppTK3-SX	GCGACGATATCGTCTACGTACCC	3.5 mM	64°		182 bp
ppTK3-DX	CGGTCACGGCATAAGGCATG				
RT-PCR of cDNA					
FMR1 E13-F	ACAAAGGACAGCATCGCTAATG	3.5 mM	60°	42386	193 bp
FMR1 E14-R	CCATTCCTTGACCATCATCAGTC			45046	
#FMR1 E1-4F	CTCCAATGGCGCTTTCTACAAG	?	?	13982	198 bp
#FMR1 E1-4R	AGGCTCTTTTTCATTTGCTCTGG			30133	
#HDHB E5-6F	TGTGAGCACTCCAAACTTTCTAAGC	?	?	exon 5[a]	156 bp
#HDHB E5-6R	GCTGAGATTCTGCTCTATGGTTTGTC			exon 6[a]	
#HDHB E10-F	GAAAGTAACGTTCGACTGTGCAATG	4 mM	60°	exon 10[a]	191 bp
#HDHB E11-R	CTGAAAAGTGTGAATAGTTCTTGC			exon 11[a]	
#ACTB E1-3F	AGCACAGAGCCTCGCCTTTG	3.5 mM	60°	14[b]	202 bp
#ACTB E1-3R	TGACCCATGCCCACCATCAC			216[b]	

*"Ta" refers to the annealing temperature of the primer set.
** The location for each primer set refers to the 5' end of the primer.
***Size refers to the amplicon length of each primer set.
# indicates primers not described elsewhere in this document
[a] human HDHB gene
[b] human beta-actin cDNA sequence; nt corresponds to Genbank accession # BC002409

APPENDIX TABLE 3. (cont.)

Notes on primers:

- Each reaction is done as a "touchdown" reaction consisting of 50 total cycles. The first 5 cycles uses an annealing temperature of T_a+4°, the next 5 cycles uses an annealing temperature of T_a+2°, and the final 40 cycles are done at the annealing temperature listed for each primer. The exception is those reactions with a T_a of 64: these use T_a+2° for 5 cycles, T_a+1° for 5 cycles, and the T_a for 40 cycles.
- The cycling parameters were as follows: 10 minutes at 95°; 5 cycles at 95° for 15 sec, T_a+4° for 5 sec, and 72° for 15 sec; 5 cycles at 95° for 15 sec, T_a+2° for 5 sec, and 72° for 15 sec; 40 cycles at 95° for 15 sec, T_a for 5 sec, and 72° for 15 sec. Some CG-rich amplicons used a melting temperature of 96° or 97° (these are listed below).
- The incubation times are very short because of the Roche LightCycler used. All reactions are done in glass capillary tubes in an open-air chamber, which allows very fast heating and cooling. If the reactions are done in another format all conditions will need to be adapted accordingly.

CG-rich amplicons:

Primer set	Melting temperature
1c	96°
1d	96°
1.2	96° for the first 10 min incubation, then 97° for all others
11	96°

APPENDIX TABLE 4. Real-time PCR efficiencies of target sequences in various DNA samples from HCT116 cells

	genomic HCT116 (cut) [a]			gen. HCT116 (uncut)			nascent HCT116		
Primer set	slope[b]	error[c]	eff.[d]	slope[b]	error[c]	eff.[d]	slope[b]	error[c]	eff.[d]
EX6b	-3.266	0.0526	2.02	-3.163	0.0249	2.07	-3.305	0.0699	2.01
UPR4	-3.438	0.0581	1.95	-3.765	0.0870	1.84	-3.453	0.0795	1.95
LB2C1	-3.626	0.1090	1.89	-3.360	0.0633	1.98	-4.019	0.1730	1.77
B48IIb	-3.621	0.0586	1.89	-3.374	0.0643	1.98	-3.690	0.0718	1.87
10	-3.239	0.0379	2.04	-3.235	0.0627	2.04	-3.754	0.127	1.85
3	-3.315	0.0489	2.00	-3.347	0.0518	1.99	-3.436	0.0416	1.95
1.5	-3.574	0.0854	1.90	-3.618	0.1060	1.89	-3.118	0.0738	2.09
1.4	-3.470	0.1040	1.94	-3.680	0.1940	1.87	-3.509	0.1080	1.93
1.2	-3.114	0.0715	2.08	n.d.[e]			-4.039	0.1500	1.77
1c	-3.289	0.0924	2.01	-5.621	0.0354	1.51	-3.913	0.0685	1.80
1d	-3.118	0.0468	2.09	n.d. [e]			-4.362	0.1230	1.70
11	-3.822	0.0427	1.83	-4.216	0.039	1.73	-4.377	0.0559	1.69
20	-3.269	0.0372	2.02	n.d. [e]			-3.425	0.1030	1.96

[a] The MCM4 and FMR1 primer sets were used to amplify genomic DNA cut with NarI and XhoI. The LaminB2 primer sets used genomic DNA cut with NarI only, since XhoI cuts within the B48IIb amplicon.
[b] Slope of the line when the crossing point of serial dilutions of the given sample are plotted against the LOG of the sample starting concentration
[c] Error (mean squared error) is calculated by the Roche LightCycler software using the equation [Error = $\Sigma(\Delta x*2)/n$], where Δx = the vertical distance between data point and regression line and n = number of data points.
[d] eff. (Efficiency) is calculated from the slope with the equation (efficiency = $e^{-1/slope}$). An efficiency of 1 refers to no amplification, and an efficiency of 2 refers to perfect amplification.
[e] no data. Samples were not tested.

APPENDIX TABLE 5. Nascent DNA quantitations.
*raw values correspond to the copy number equivalent in 1 pg of genomic HCT116 DNA [cut with XhoI/NarI] (so a "3108" nascent DNA raw value has the same copy number as 3108 pg genomic HCT116)
*raw values are the average of 2 duplicate real-time PCR reactions
*Note: there is approximately one diploid genome per 6pg in HCT116

500-1000 bp HCT116 nascent DNA

	prep#1		prep#2		prep#3		prep#4			
primer	raw	norm.	raw	norm.	raw	norm.	raw	norm.	Aver.	SD
EX6b	3108	1.0	2646	1.0	2604	1.0	2016	1.0	1.00	0.00
UPR4	34092	11.0	26430	10.0	26772	10.3	20160	10.0	10.31	0.46
UPR5	29898	9.6	25662	9.7	31392	12.1	20166	10.0	10.34	1.15
LB2C1	6546	1.0	4920	1.0	6876	1.0	4704	1.0	1.00	0.00
LO	49854	7.6	48018	9.8	46758	6.8	32202	6.8	7.76	1.39
B48IIb	57840	8.8	50760	10.3	46542	6.8	33552	7.1	8.26	1.64
10	2112	0.8	1896	0.9	1260	0.9	1062	1.0	0.88	0.06
3	2586	1.0	2226	1.0	1626	1.1	1134	1.0	1.05	0.06
2	7068	2.7	4548	2.1	4284	3.0	3738	3.4	2.80	0.52
1.5	10332	4.0	7452	3.5	8868	6.2	7536	6.8	5.11	1.63
1.4	17166	6.6	12810	6.0	18150	12.7	11820	10.7	8.97	3.22
1.2	3018	1.2	2580	1.2	3756	2.6	2328	2.1	1.77	0.71
1c	3372	1.3	2238	1.0	3048	2.1	2850	2.6	1.76	0.71
1d	1410	0.5	684	0.3	1398	1.0	1398	1.3	0.77	0.42
11	1044	0.4	1032	0.5	1200	0.8	1122	1.0	0.68	0.29
12	4632	1.8	3108	1.5	3510	2.4	2286	2.1	1.94	0.42
20	3138	1.2	2280	1.1	1416	1.0	1128	1.0	1.07	0.09

*all values are for 2uL 1:4 dilution of original nascent preps

1000-2000 bp HCT116 nascent DNA

	prep#1		prep#2		prep#3		prep#4		Aver.	SD
primer	raw	norm.	raw	norm.	raw	norm.	raw	norm.		
EX6b	1650	1.0	1950	1.0	1782	1.0	1596	1.0	1.00	0.00
UPR4	8922	5.4	16218	8.3	12390	7.0	13008	8.2	7.21	1.34
UPR 5	12714	7.7	21288	10.9	16914	9.5	17658	11.1	9.79	1.56
LB2C1	2004	1.0	2136	1.0	1560	1.0	1926	1.0	1.00	0.00
LO	n.d.		n.d.		n.d.		n.d.			
B48IIb	10572	5.3	10332	4.8	6972	4.5	10380	5.4	4.99	0.42
10	870	0.6	1488	0.7	948	0.7	732	0.7	0.69	0.03
3	1668	1.2	2568	1.2	1374	1.0	1260	1.2	1.17	0.10
2	3966	2.9	6684	3.2	4650	3.5	2922	2.8	3.09	0.30
1.5	4500	3.3	7428	3.5	5448	4.1	5118	4.9	3.95	0.71
1.4	8772	6.4	15318	7.3	9066	6.7	8934	8.6	7.26	0.93
1.2	1026	0.8	2052	1.0	2208	1.6	2160	2.1	1.36	0.60
1c	1362	1.0	3588	1.7	3636	2.7	2964	2.8	2.06	0.87
1d	1356	1.0	1206	0.6	804	0.6	1176	1.1	0.82	0.28
11	432	0.3	936	0.4	1164	0.9	1218	1.2	0.70	0.39
12	1656	1.2	1638	0.8	1878	1.4	3168	3.0	1.61	0.99
20	1548	1.1	2232	1.1	1710	1.3	1140	1.1	1.14	0.09

*all values are for 2uL 1:4 dilution of original nascent preps

APPENDIX TABLE 5. (cont.)

500-1000 bp Hela S3 nascent DNA

	prep #1		prep #2		Aver.	SD
primer	raw	norm.	raw	norm.		
EX6b	372	1.0	754	1.0	1.00	0.00
UPR4	1454	3.9	2342	3.1	3.51	0.57
LB2C1	681	1.0	2038	1.0	1.00	0.00
B48IIb	2067	3.0	7618	3.7	3.39	0.50
10	280	0.5	281	0.6	0.56	0.03
3	842	1.6	693	1.4	1.53	0.12
1.5	2380	4.6	3695	7.7	6.14	2.22
1.4	2248	4.3	3662	7.6	5.98	2.35
1c	1293	2.5	3778	7.9	5.19	3.82
11	209	0.4	362	0.8	0.58	0.25
20	439	0.8	463	1.0	0.91	0.09

*all values are for 3uL original nascent preps (out of a total of 100uL)

500-1000 bp PFW nascent DNA

	prep #1		prep #2		Aver.	SD
primer	raw	norm.	raw	norm.		
EX6b	135	1.0	82	1.0	1.00	0.00
UPR4	990	7.3	692	8.4	7.89	0.78
LB2C1	302	1.0	183	1.0	1.00	0.00
B48IIb	1176	3.9	1524	8.3	6.11	3.14
10	156	0.9	86	0.5	0.74	0.29
3	207	1.3	226	1.4	1.33	0.10
1.5	1374	8.3	1231	7.6	8.00	0.49
1.4	2107	12.8	721	4.5	8.64	5.88
1c	470	2.9	754	4.7	3.77	1.29
11	105	0.6	198	1.2	0.93	0.42
20	131	0.8	171	1.1	0.93	0.19

*all values are for 2uL original nascent preps (out of a total of 100uL)

500-1000 bp GM05381 nascent DNA

	prep #1		prep #2		prep #3		prep #4		Aver.	SD
primer	raw	norm.	raw	norm.	raw	norm.	raw	norm.		
EX6b	15.6	1.0	6.4	1.0	36.7	1.0	22.6	1.0	1.00	0.00
UPR4	98	6.3	43	6.7	250	6.8	207	9.2	7.24	1.30
LB2C1	94	1.0	36	1.0	201	1.0	76.0	1.0	1.00	0.00
B48IIb	394	4.2	234	6.5	764	3.8	782	10.3	6.20	2.98
10	25	0.3	11	0.5	23	0.6	75	0.8	0.55	0.22
3	181	2.1	34	1.6	93	2.3	138	1.5	1.87	0.36
1.5	237	2.8	133	6.2	225	5.5	203	2.3	4.18	1.97
1.4	397	4.6	194	9.1	251	6.1	603	6.7	6.63	1.86
1c	245	2.8	126	5.9	0	0.0	34	0.4	2.28	2.73
11	65	0.8	0	0.0	0	0.0	94	1.0	0.45	0.53
20	52	0.6	19	0.9	7.5	0.2	57	0.6	0.58	0.29

*all values are for 2uL original nascent preps (out of a total of 100uL)

APPENDIX TABLE 5. (cont.)

500-1000 bp GM08400 nascent DNA

	prep #1		prep #2		Aver.	SD
primer	raw	norm.	raw	norm.		
EX6b	9.5	1.0	18.8	1.0	1.00	0.00
UPR4	82	8.6	208	11.1	9.85	1.72
LB2C1	120	1.0	232	1.0	1.00	0.00
B48IIb	416	3.5	396	1.7	2.59	1.24
10	37	0.5	57	0.3	0.38	0.13
3	88	1.1	373	1.9	1.49	0.51
1.5	154	2.0	796	4.0	2.97	1.40
1.4	128	1.6	1459	7.3	4.45	3.97
1c	452	5.8	1371	6.8	6.31	0.73
11	73	0.9	76	0.4	0.66	0.39
20	109	1.4	173	0.9	1.13	0.38

*all values are for 3uL original nascent preps (out of a total of 100uL)

500-1000 bp GM05848 nascent DNA

	prep #1		prep #2		Aver.	SD
primer	raw	norm.	raw	norm.		
EX6b	10	1.0	25	1.0	1.00	0.00
UPR4	30	3.0	126	5.0	4.02	1.44
LB2C1	51	1.0	62	1.0	1.00	0.00
B48IIb	114	2.2	244	3.9	3.09	1.20
10	37	1.3	21	0.8	1.04	0.38
3	42	1.5	36	1.3	1.41	0.11
1.5	105	3.7	197	7.3	5.51	2.53
1.4	324	11.5	283	10.5	10.98	0.70
1c	21	0.7	69	2.6	1.65	1.28
11	26	0.9	17	0.6	0.78	0.21
20	5.7	0.2	24	0.9	0.55	0.49

*all values are for 2uL original nascent preps (out of a total of 100uL)

500-1000 bp GM04026 nascent DNA

	prep #1		prep #2		Aver.	SD
primer	raw	norm.	raw	norm.		
EX6b	135	1.0	62	1.0	1.00	0.00
UPR4	275	2.0	248	4.0	3.02	1.39
LB2C1	97	1.0	127	1.0	1.00	0.00
B48IIb	195	2.0	302	2.4	2.19	0.26
10	73	1.0	57	0.5	0.73	0.38
3	105	1.4	218	1.8	1.61	0.24
1.5	516	7.1	206	1.7	4.37	3.81
1.4	921	12.6	548	4.5	8.54	5.77
1c	164	2.2	159	1.3	1.77	0.67
11	107	1.5	73	0.6	1.03	0.62
20	41	0.6	94	0.8	0.66	0.14

*all values are for 2uL original nascent preps (out of a total of 100uL)

APPENDIX TABLE 5. (cont.)

500-1000 bp GM07072 nascent DNA

	prep #1		prep #2		prep #3		Aver.	SD
primer	raw	norm.	raw	norm.	raw	norm.		
EX6b	2562	1.0	228	1.0	124	1.0	1.00	0.00
UPR4	25940	10.1	1732	7.6	2100	16.9	11.55	4.83
LB2C1	5772	1.0	474	1.0	225	1.0	1.00	0.00
B48IIb	25282	4.4	2448	5.2	1958	8.7	6.08	2.30
10	2165	0.6	200	0.5	113	0.7	0.59	0.13
3	4471	1.1	419	1.0	186	1.2	1.12	0.11
1.5	14932	3.8	1876	4.4	541	3.5	3.93	0.47
1.4	10558	2.7	1083	2.6	370	2.4	2.56	0.16
1c	8922	2.3	2061	4.9	78	0.5	2.56	2.20
11	1858	0.5	238	0.6	48	0.3	0.45	0.13
20	5030	1.3	650	1.5	164	1.1	1.30	0.24

*all values are for 3uL original nascent preps (out of a total of 100uL)

APPENDIX TABLE 6. Quantitative RT-PCR of FMR1 mRNA

Cell Line	primer set FMR-13/14	Average	STD dev	% of HCT
HCT #1	6094	5837.33	831.27	100.000
HCT #2	6510			
HCT #3	4908			
HAF #1	6507	6349.67	282.09	108.777
HAF #2	6024			
HAF #3	6518			
GM05381 #1	5250	5292.00	244.72	90.658
GM05381 #2	5071			
GM05381 #3	5555			
GM08400 #1	5463	5517.67	800.40	94.524
GM08400 #2	6344			
GM08400 #3	4746			
GM07072 #1	42	30.67	12.66	0.525
GM07072 #2	33			
GM07072 #3	17			
GM04026 #1	0.56	0.28	0.28	0.005
GM04026 #2	0			
GM04026 #3	0.29			
GM05848 #1	1.4	1.30	0.17	0.022
GM05848 #2	1.4			
GM05848 #3	1.1			

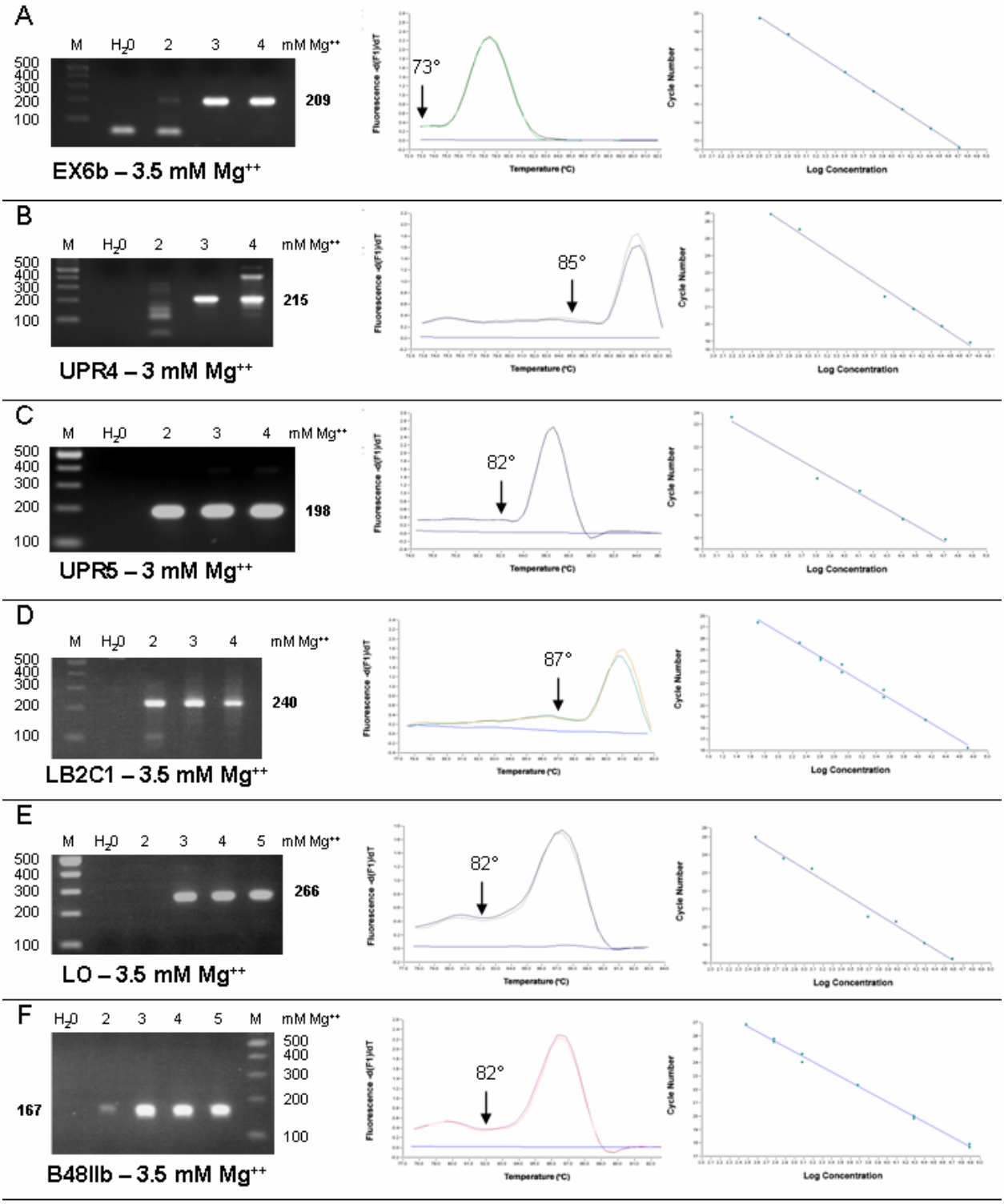
*template = 0.5uL of 20uL cDNA reaction (cDNA reaction used 1ug total RNA as template)
*Standard = serial dilutions of HCT116 #1 PCR product (1 corresponds to ~1 trillionth of the reaction product)
*Each value is the average of two real-time PCR reactions done in parallel

APPENDIX TABLE 7. Relative RNA quantitation by densitometry of Figure 9B

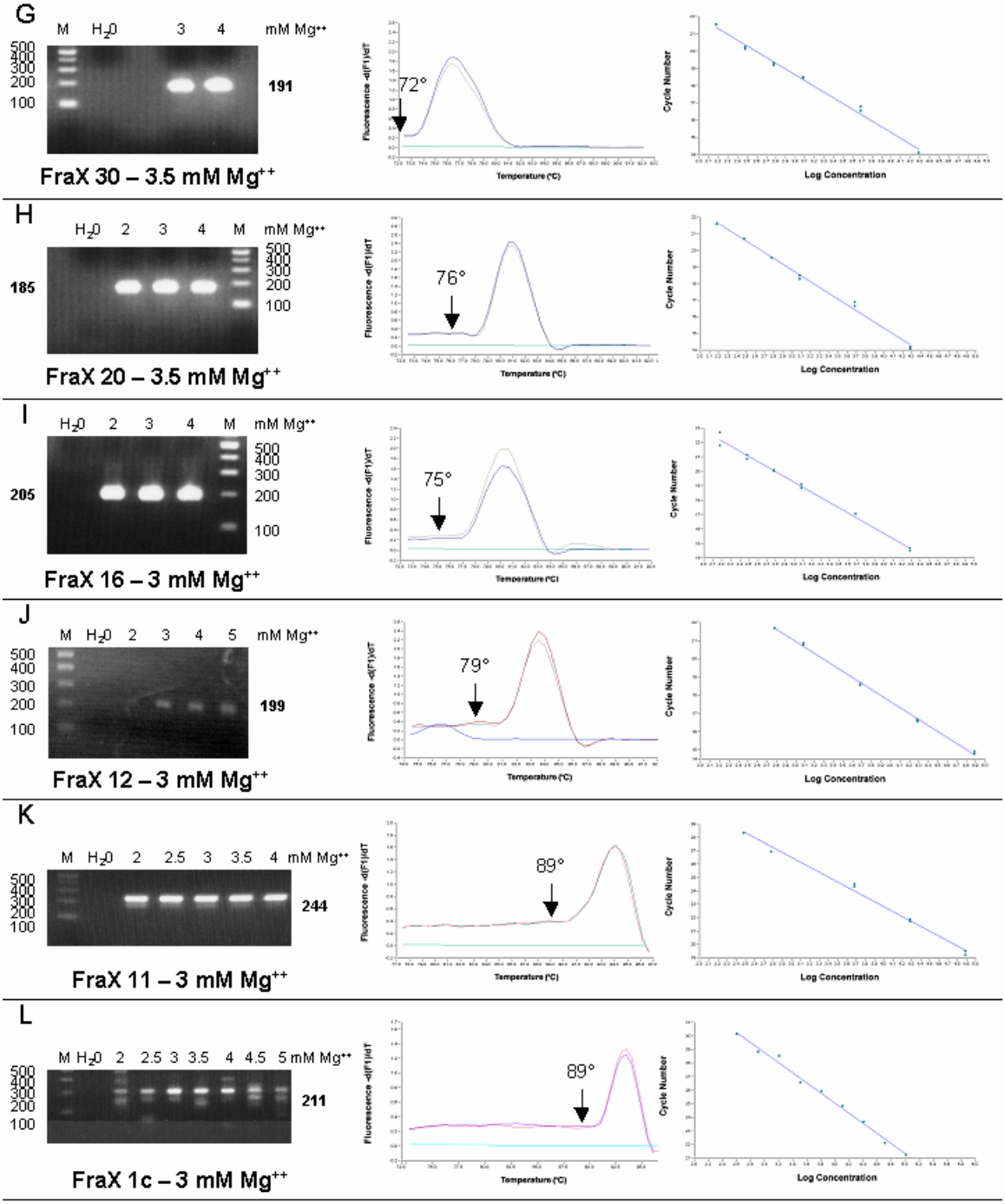
Lane	Sample	volume	minus background	normalized
1	HCT	425624	184013	0.98
2	HAF	439968	198357	1.06
3	GM05381	452190	210579	1.12
4	GM08400	411653	170042	0.91
5	GM07072	424373	182762	0.97
6	GM04026	435037	193426	1.03
7	GM05848	415351	173740	0.93
8	EMPTY	241611	0	
	AVERAGE 1-7:		187560	

*1 ug of total RNA was loaded in each lane
*Densitometry was done using IPlabgel software
*"background"" was determined by measuring volume in an empty lane
*Normalization was achieved by dividing the volume of each lane by the average across all lanes

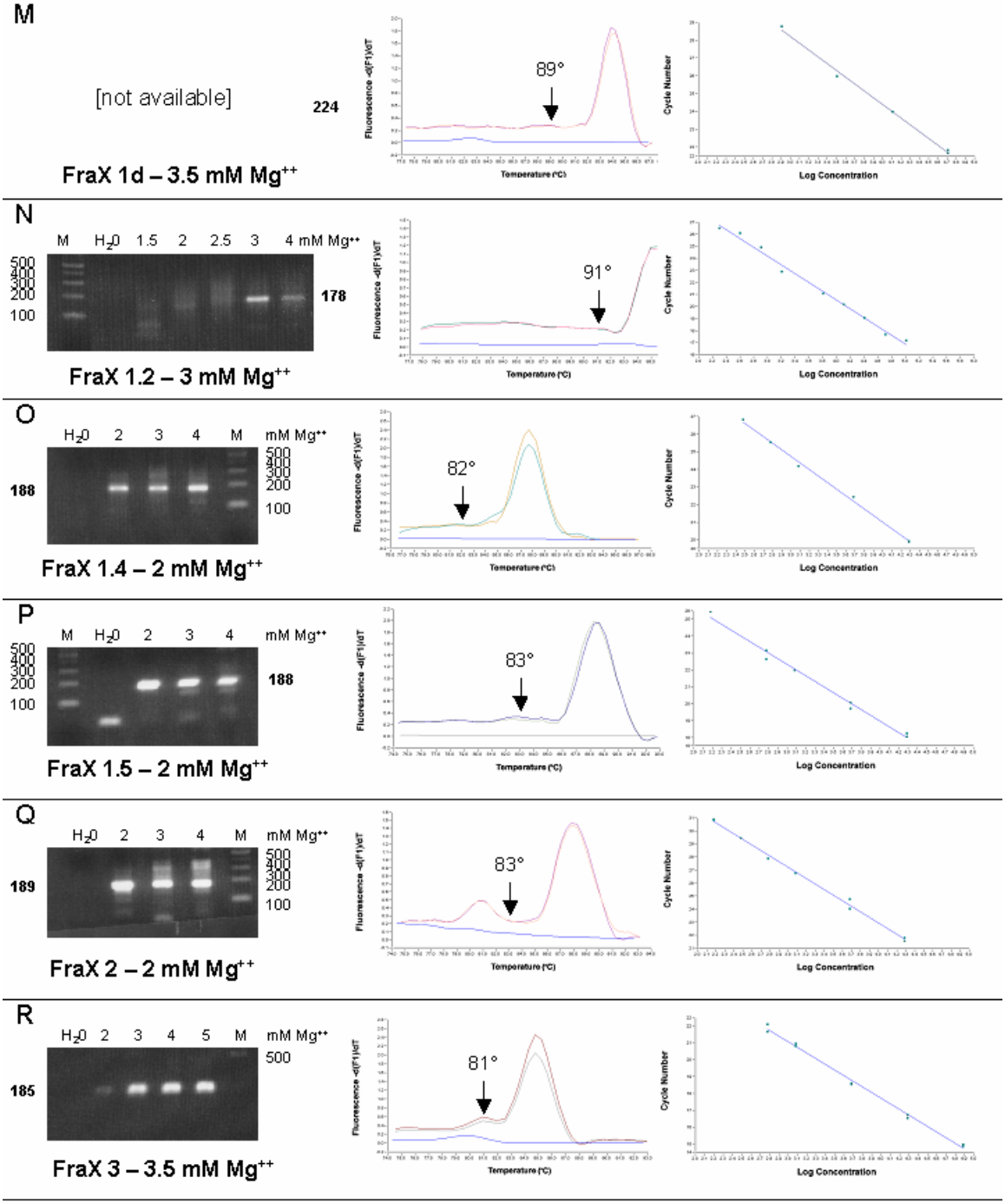
APPENDIX FIGURE 22



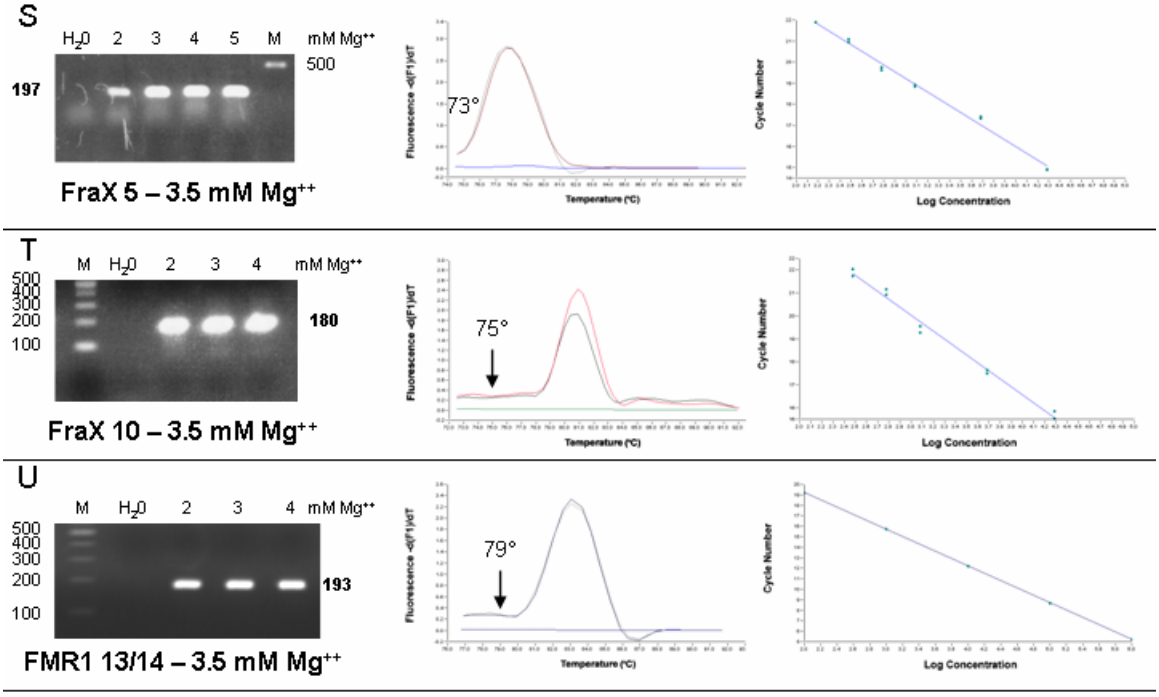
APPENDIX FIGURE 22 (cont.)



APPENDIX FIGURE 22 (cont.)



APPENDIX FIGURE 22 (cont.)



Appendix Figure 22. Characterization of real-time PCR primers. PCR conditions were optimized for each primer set at the MCM4 locus (A-C), Lamin B2 locus (D-F), Fragile X locus (G-T), and FMR1 cDNA (U) to give a single detectable product after 50 cycles. Left panels: Real-time PCR products produced in magnesium titration experiments were electrophoresed on an agarose gel and stained with ethidium bromide (M = marker, numbers to the side of the gel indicate sizes in bp). In each case, the H_2O lane shows products generated when water was used instead of template with 3 or 3.5 mM Mg^{++}.

The middle panel of each row shows a sample melting curve analysis of PCR products amplified using genomic HCT116 DNA template in duplicate, or with water as a negative control. In these experiments, the PCR products after 50 cycles are re-annealed, then the temperature is gradually raised while fluorescence is measured. The derivative of the fluorescence is given to show a peak at the temperature where the product melts. In instances where a small amount of non-target product is observed with a lower melting temperature (as in row Q), the reaction is raised above that temperature before fluorescence is measured whenever possible, so only the target product is detected. Arrows indicate the temperature at which fluorescence is measured for each primer set. In rows D and P, small products are seen on the agarose gel when water is used instead of template, but the melting curve analysis shows that these products melt below 72°C.

The right panel of each row shows a sample standard curve for each primer set using serial dilutions of HCT116 genomic DNA. The tight fit of the points to the line on each curve indicates a broad concentration range and minimal error for the quantitation of starting DNA template.

NUCLEOSOMES HAVE AN ORDERED ARRANGEMENT AROUND THE DNR ELEMENT AT ORI-BETA IN CHINESE HAMSTER OVARY (CHO) CELLS.

Introduction

A microarray-based investigation of nucleosome positioning in *S. cerevisiae* revealed that functionally important genomic sequences such as gene promoters have an ordered arrangement of nucleosomes, but other regions of DNA have essentially random placement of nucleosomes (220). Presumably, this trend occurs because the correct positioning of nucleosomes is necessary for many DNA-binding proteins to access their target sites (108, 195, 196, 213). To investigate the pattern of nucleosome positions around the DNR region, we used an *in vivo* micrococcal nuclease (MNase) sensitivity assay. Purified DNA from MNase-treated chromatin is digested with a specific restriction enzyme to create a uniform end to fragments at the region of interest. Following Southern blotting, a probe is hybridized to the sequence immediately flanking the restriction site, thus indirectly end-labeling all of the MNase fragments from the region of interest that share a common end at the restriction site. Fragment lengths indicate the nearest site of MNase cleavage on each molecule of DNA. Discrete bands appearing on the Southern blot indicate groups of DNA lengths corresponding to areas on chromatin hypersensitive to MNase. Since MNase cuts preferentially in the ~50 bp linker region between 146 bp nucleosome core particles, this method has been used to determine the approximate positions of nucleosomes on chromatin *in vivo* (207). However, tightly associated protein complexes can also create a nucleosome-like gap in MNase accessibility (194).

Materials and Methods

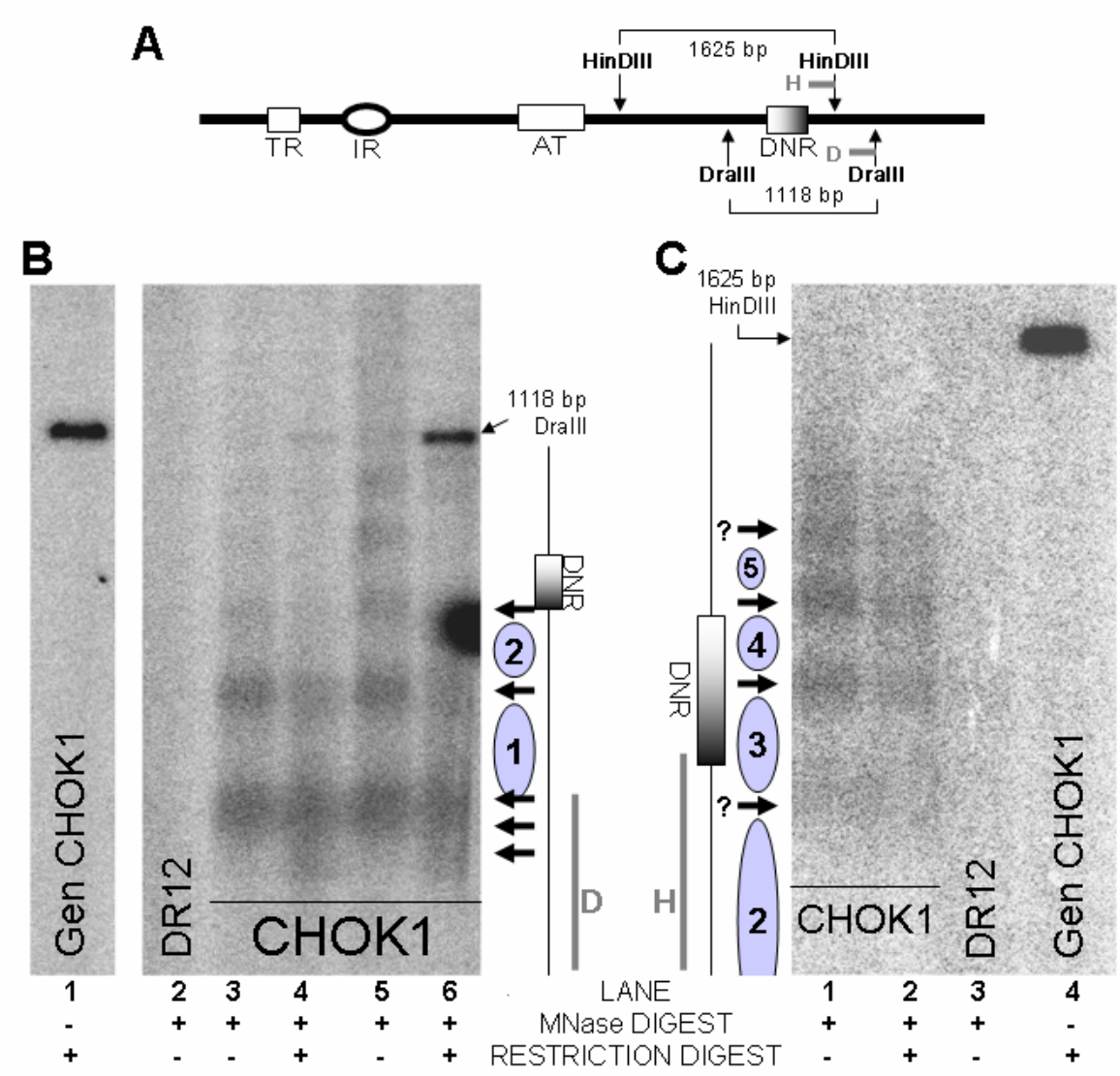
MNase-hypersensitive sites were detected *in vivo* on chromatin essentially as previously described (26). Briefly, purified nuclei containing intact chromatin from CHO-K1 cells was subjected to limited digestion with MNase (Sigma), then digested overnight at 37°C with Proteinase K and purified with repeated phenol:chloroform extractions, RNase A digestion, and ethanol precipitation. The purified MNase-treasted DNA was then digested with either DraIII or HinDIII, to create uniform ends in the population of MNase fragments around DNR. Forty µg of MNase-treated DNA, with or without restriction enzyme digestion, was electrophoresed through a 1.4% agarose gel and blotted to a Highbond N+ membrane (Amersham #RPN203B) for Southern blot analysis. Southern blot hybridization and washes were performed as described above for the non-specific integration system using Church buffer. The probe D for DraIII-digested DNA was a 210 bp fragment located immediately upstream of a DraIII site at nucleotide 5028 (relative to the BamHI site upstream of ori-beta). Probe D was amplified by PCR using the primer sequences 5'-TCTTTTAAGATACTTGGCCCC-3' and 5'-GTGCCACATCCCTGATGATC-3'. Probe H for HinDIII-digested DNA was a 208 bp fragment located immediately upstream of a HinDIII site at nucleotide 4636 (relative to the BamHI site). Probe H was amplified by PCR using the primer sequences 5'-ATCAGACTGGTCCCATATCC-3' and 5'-TTAGAAATTCATCATGAAAAATCTG-3'. By using probes located immediately next to restriction enzyme cut sites, MNase fragments at the region of interest are indirectly end-labeled at the restriction site when visualized after Southern blot hybridization.

Results and Discussion

The region next to a DraIII site downstream of DNR was investigated using the MNase-sensitivity assay. Probe D hybridized only to a 1118 bp band in non-MNase-treated, DraIII-digested genomic CHO-K1 DNA, indicating the expected DraIII fragment (Figure 23A and 23B, lane 1). Probe D did not show any appreciable hybridization to MNase-treated DR12 hamster DNA, which lacks the ori-beta sequence, indicating that probe D bound specifically to the ori-beta target sequence (Figure 23B, lane 2). A ladder of mono-, di-, tri-, and poly-nucleosome-sized DNA fragments in MNase-treated CHO-K1 DNA suggests that nucleosomes around the probe D sequence have a specific arrangement (Figure 23B, lanes 3 and 5). Hypersensitive sites observed in MNase-treated, DraIII-digested CHO-K1 DNA suggests an ordered arrangement of nucleosomes specifically between the *DraIII* site and DNR (Figure 23B, lanes 4 and 6). These results suggest an ordered arrangement of nucleosomes immediately downstream of the DNR element at the endogenous ori-beta in CHO-K1 cells.

To analyze the pattern of nucleosome positions across the DNR sequence, the MNase-hypersensitivity assay was repeated using a HinDIII site downstream of DNR. Probe H hybridized to only a 1625 bp sequence corresponding the expected size of the HinDIII fragment in HinDIII-digested genomic CHO-K1 DNA (Figure 23A and 23C, lane 4), but it did not hybridize appreciably to any sequences in DR12 MNase-treated DNA (Figure 23C, lane 3), indicating that probe H hybridized specifically to its target sequence in ori-beta. A ladder of hypersensitive sites corresponding to the sizes nucleosome particles was observed in MNase-treated CHO-K1 DNA, suggesting that the

nucleosomes have a specific pattern around the HinDIII site (Figure 23C, lane 1). These hypersensitive sites were visible after the MNase-treated DNA was digested with HinDIII, suggesting that nucleosomes have an ordered arrangement across DNR (Figure 23C, lane 2). An ordered arrangement of nucleosomes around DNR suggests that a specific arrangement may be needed for the binding of *trans*-acting factors and/or that the peculiar sequence of DNR is able to position nucleosomes.



Appendix Figure 23. A regular pattern of MNase hypersensitive sites is detected at and around the DNR element. (A) Diagram of the DHFR ori-beta with pertinent HinDIII- and DraIII-digestion fragments labeled. The location of the HinDIII fragment-specific (H) and DraIII fragment-specific (D) probes are drawn as gray lines. (B, C) Southern blots of MNase-digested chromatin from CHOK-1 and DR12 cells, with or without restriction enzyme digestion. Probe D (panel B) or probe H (panel C) were hybridized to the blots. The locations where sequences of the DNR element (shaded box) and probe hybridization sites (gray line) would migrate on the gel are drawn next to the blots. MNase sites (bold arrows) and nucleosome-sized spaces (gray ovals) are drawn for clarity. Below each blot, + shows the use of MNase or restriction digestion, as indicated. The sizes corresponding to the 1118 bp DraIII fragment (B) or 1625 HinDIII fragment (C) are shown.

REFERENCES CITED

1. **Abdurashidova, G., M. B. Danailov, A. Ochem, G. Triolo, V. Djeliova, S. Radulescu, A. Vindigni, S. Riva, and A. Falaschi.** 2003. Localization of proteins bound to a replication origin of human DNA along the cell cycle. Embo J **22:**4294-303.

2. **Abdurashidova, G., M. Deganuto, R. Klima, S. Riva, G. Biamonti, M. Giacca, and A. Falaschi.** 2000. Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science **287:**2023-6.

3. **Aggarwal, B. D., and B. R. Calvi.** 2004. Chromatin regulates origin activity in Drosophila follicle cells. Nature **430:**372-6.

4. **Aladjem, M. I., and E. Fanning.** 2004. The replicon revisited: an old model learns new tricks in metazoan chromosomes. EMBO Rep **5:**686-91.

5. **Aladjem, M. I., L. W. Rodewald, J. L. Kolman, and G. M. Wahl.** 1998. Genetic dissection of a mammalian replicator in the human beta-globin locus. Science **281:**1005-9.

6. **Alexiadis, V., P. D. Varga-Weisz, E. Bonte, P. B. Becker, and C. Gruss.** 1998. In vitro chromatin remodelling by chromatin accessibility complex (CHRAC) at the SV40 origin of DNA replication. Embo J **17:**3428-38.

7. **Altman, A. L., and E. Fanning.** 2001. The Chinese hamster dihydrofolate reductase replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequence elements for activity. Mol Cell Biol **21:**1098-110.

8. **Altman, A. L., and E. Fanning.** 2004. Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. Mol Cell Biol **24:**4138-50.

9. **Anachkova, B., and J. L. Hamlin.** 1989. Replication in the amplified dihydrofolate reductase domain in CHO cells may initiate at two distinct sites, one of which is a repetitive sequence element. Mol Cell Biol **9:**532-40.

10. **Anderson, S., and M. L. DePamphilis.** 1979. Metabolism of Okazaki fragments during simian virus 40 DNA replication. J Biol Chem **254:**11495-504.

11. **Anglana, M., F. Apiou, A. Bensimon, and M. Debatisse.** 2003. Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. Cell **114:**385-94.

12. **Antequera, F.** 2004. Genomic specification and epigenetic regulation of eukaryotic DNA replication origins. Embo J **23:**4365-70.

13. **Balakumaran, B. S., C. H. Freudenreich, and V. A. Zakian.** 2000. CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in Saccharomyces cerevisiae. Hum Mol Genet **9:**93-100.

14. **Bell, S. P.** 2002. The origin recognition complex: from simple origins to complex functions. Genes Dev **16:**659-72.

15. **Bell, S. P., and A. Dutta.** 2002. DNA replication in eukaryotic cells. Annu Rev Biochem **71:**333-74.

16. **Berberich, S., A. Trivedi, D. C. Daniel, E. M. Johnson, and M. Leffak.** 1995. In vitro replication of plasmids containing human c-myc DNA. J Mol Biol **245:**92-109.

17. **Bhat, K. M., G. Farkas, F. Karch, H. Gyurkovics, J. Gausz, and P. Schedl.** 1996. The GAGA factor is required in the early Drosophila embryo not only for transcriptional regulation but also for nuclear division. Development **122:**1113-24.

18. **Bianchi, A., R. D. Wells, N. H. Heintz, and M. S. Caddle.** 1990. Sequences near the origin of replication of the DHFR locus of Chinese hamster ovary cells adopt left-handed Z-DNA and triplex structures. J Biol Chem **265:**21789-96.

19. **Bielinsky, A. K., and S. A. Gerbi.** 1998. Discrete start sites for DNA synthesis in the yeast ARS1 origin. Science **279:**95-8.

20. **Bossone, S. A., C. Asselin, A. J. Patel, and K. B. Marcu.** 1992. MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination. Proc Natl Acad Sci U S A **89:**7452-6.

21. **Brattain, M. G., W. D. Fine, F. M. Khaled, J. Thompson, and D. E. Brattain.** 1981. Heterogeneity of malignant cells from a human colonic carcinoma. Cancer Res **41:**1751-6.

22. **Brinton, B. T., M. S. Caddle, and N. H. Heintz.** 1991. Position and orientation-dependent effects of a eukaryotic Z-triplex DNA motif on episomal DNA replication in COS-7 cells. J Biol Chem **266:**5153-61.

23. **Burhans, W. C., J. E. Selegue, and N. H. Heintz.** 1986. Isolation of the origin of replication associated with the amplified Chinese hamster dihydrofolate reductase domain. Proc Natl Acad Sci U S A **83:**7790-4.

24. **Burhans, W. C., L. T. Vassilev, M. S. Caddle, N. H. Heintz, and M. L. DePamphilis.** 1990. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell **62:**955-65.

25. **Caddle, M. S., R. H. Lussier, and N. H. Heintz.** 1990. Intramolecular DNA triplexes, bent DNA and DNA unwinding elements in the initiation region of an amplified dihydrofolate reductase replicon. J Mol Biol **211:**19-33.

26. **Carey, M., and S. T. Smale.** 2000. Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

27. **Chakalova, L., E. Debrand, J. A. Mitchell, C. S. Osborne, and P. Fraser.** 2005. Replication and transcription: shaping the landscape of the genome. Nat Rev Genet **6:**669-77.

28. **Chastain, P. D., 2nd, S. M. Cohen, B. P. Brylawski, M. Cordeiro-Stone, and D. G. Kaufman.** 2006. A Late Origin of DNA Replication in the Trinucleotide Repeat Region of the Human FMR2 Gene. Cell Cycle **5:**869-72.

29. **Cleary, J. D., K. Nichol, Y. H. Wang, and C. E. Pearson.** 2002. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat Genet **31:**37-46.

30. **Cleary, J. D., and C. E. Pearson.** 2003. The contribution of cis-elements to disease-associated repeat instability: clinical and experimental evidence. Cytogenet Genome Res **100:**25-55.

31. **Cleary, J. D., and C. E. Pearson.** 2005. Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. Trends Genet **21:**272-80.

32. **Clyne, R. K., and T. J. Kelly.** 1997. Identification of autonomously replicating sequence (ARS) elements in eukaryotic cells. Methods **13:**221-33.

33. **Cohen, S. M., B. P. Brylawski, M. Cordeiro-Stone, and D. G. Kaufman.** 2002. Mapping of an origin of DNA replication near the transcriptional promoter of the human HPRT gene. J Cell Biochem **85:**346-56.

34. **Cohen, S. M., B. P. Brylawski, M. Cordeiro-Stone, and D. G. Kaufman.** 2003. Same origins of DNA replication function on the active and inactive human X chromosomes. J Cell Biochem **88:**923-31.

35. **Cvetic, C., and J. C. Walter.** 2005. Eukaryotic origins of DNA replication: could you please be more specific? Semin Cell Dev Biol **16:**343-53.

36. **Daniel, A., L. Ekblom, and S. Phillips.** 1984. Fragile X expression suppressed in either FUdR or methotrexate treated fibroblasts by pretreatment with 5-azacytidine. Am J Med Genet **17:**255-7.

37. **Debatisse, M., F. Toledo, and M. Anglana.** 2004. Replication initiation in mammalian cells: changing preferences. Cell Cycle **3:**19-21.

38. **Delgado, S., M. Gomez, A. Bird, and F. Antequera.** 1998. Initiation of DNA replication at CpG islands in mammalian chromosomes. Embo J **17:**2426-35.

39. **DePamphilis, M. L.** 1999. Replication origins in metazoan chromosomes: fact or fiction? Bioessays **21:**5-16.

40. **DePamphilis, M. L.** 2000. Review: nuclear structure and DNA replication. J Struct Biol **129:**186-97.

41. **DesJardins, E., and N. Hay.** 1993. Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-myc promoters. Mol Cell Biol **13:**5710-24.

42. **Dhar, V., A. I. Skoultchi, and C. L. Schildkraut.** 1989. Activation and repression of a beta-globin gene in cell hybrids is accompanied by a shift in its temporal replication. Mol Cell Biol **9:**3524-32.

43. **Diffley, J. F., and B. Stillman.** 1988. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. Proc Natl Acad Sci U S A **85:**2120-4.

44. **Dijkwel, P. A., and J. L. Hamlin.** 1995. The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand start sites. Mol Cell Biol **15:**3023-31.

45. **Ermakova, O. V., L. H. Nguyen, R. D. Little, C. Chevillard, R. Riblet, N. Ashouian, B. K. Birshtein, and C. L. Schildkraut.** 1999. Evidence that a single replication fork proceeds from early to late replicating domains in the IgH locus in a non-B cell line. Mol Cell **3:**321-30.

46. **Espinas, M. L., E. Jimenez-Garcia, A. Martinez-Balbas, and F. Azorin.** 1996. Formation of triple-stranded DNA at d(GA.TC)n sequences prevents nucleosome assembly and is hindered by nucleosomes. J Biol Chem **271:**31807-12.

47. **Feng, W., D. Collingwood, M. E. Boeck, L. A. Fox, G. M. Alvino, W. L. Fangman, M. K. Raghuraman, and B. J. Brewer.** 2006. Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. Nat Cell Biol **8:**148-55.

48. **Fernandez-Lopez, L., E. Pineiro, R. Marcos, A. Velazquez, and J. Surralles.** 2004. Induction of instability of normal length trinucleotide repeats within human disease genes. J Med Genet **41:**e3.

49. **Fisher, D., and M. Mechali.** 2003. Vertebrate HoxB gene expression requires DNA replication. Embo J **22:**3737-48.

50. **Freudenreich, C. H., S. M. Kantrow, and V. A. Zakian.** 1998. Expansion and length-dependent fragility of CTG repeats in yeast. Science **279:**853-6.

51. **Fu, H., L. Wang, C. M. Lin, S. Singhania, E. E. Bouhassira, and M. I. Aladjem.** 2006. Preventing gene silencing with human replicators. Nat Biotechnol **24:**572-6.

52. **Fuster, C., R. Miro, C. Templado, L. Barrios, and J. Egozcue.** 1989. Expression of folate-sensitive fragile sites in lymphocyte chromosomes. Hum Genet **81:**243-6.

53. **Garber, K., K. T. Smith, D. Reines, and S. T. Warren.** 2006. Transcription, translation and fragile X syndrome. Curr Opin Genet Dev **16:**270-5.

54. **Gatchel, J. R., and H. Y. Zoghbi.** 2005. Diseases of unstable repeat expansion: mechanisms and common principles. Nat Rev Genet **6:**743-55.

55. **Gecz, J., A. K. Gedeon, G. R. Sutherland, and J. C. Mulley.** 1996. Identification of the gene FMR2, associated with FRAXE mental retardation. Nat Genet **13:**105-8.

56. **Genc, B., H. Muller-Hartmann, M. Zeschnigk, H. Deissler, B. Schmitz, F. Majewski, A. von Gontard, and W. Doerfler.** 2000. Methylation mosaicism of 5'-(CGG)(n)-3' repeats in fragile X, premutation and normal individuals. Nucleic Acids Res **28:**2141-52.

57. **Gerber, J. K., E. Gogel, C. Berger, M. Wallisch, F. Muller, I. Grummt, and F. Grummt.** 1997. Termination of mammalian rDNA replication: polar arrest of replication fork movement by transcription termination factor TTF-I. Cell **90:**559-67.

58. **Ghosh, M., G. Liu, G. Randall, J. Bevington, and M. Leffak.** 2004. Transcription factor binding and induced transcription alter chromosomal c-myc replicator activity. Mol Cell Biol **24:**10193-207.

59. **Giacca, M., L. Zentilin, P. Norio, S. Diviacco, D. Dimitrova, G. Contreas, G. Biamonti, G. Perini, F. Weighardt, S. Riva, and et al.** 1994. Fine mapping of a replication origin of human DNA. Proc Natl Acad Sci U S A **91:**7119-23.

60. **Gilbert, D. M.** 2004. In search of the holy replicator. Nat Rev Mol Cell Biol **5:**848-55.

61. **Gilbert, D. M.** 2001. Making sense of eukaryotic DNA replication origins. Science **294:**96-100.

62. **Glover, T. W., M. F. Arlt, A. M. Casper, and S. G. Durkin.** 2005. Mechanisms of common fragile site instability. Hum Mol Genet **14 Spec No. 2:**R197-205.

63. **Gomez, M., and N. Brockdorff.** 2004. Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage. Proc Natl Acad Sci U S A **101:**6923-8.

64. **Griffiths, M. J., and M. C. Strachan.** 1991. A single lymphocyte culture for fragile X induction and prometaphase chromosome analysis. J Med Genet **28:**837-9.

65. **Grummt, I., U. Maier, A. Ohrlein, N. Hassouna, and J. P. Bachellerie.** 1985. Transcription of mouse rDNA terminates downstream of the 3' end of 28S RNA and involves interaction of factors with repeated sequences in the 3' spacer. Cell **43:**801-10.

66. **Gu, Y., Y. Shen, R. A. Gibbs, and D. L. Nelson.** 1996. Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. Nat Genet **13:**109-13.

67. **Hagerman, R. J., C. E. Hull, J. F. Safanda, I. Carpenter, L. W. Staley, R. A. O'Connor, C. Seydel, M. M. Mazzocco, K. Snow, S. N. Thibodeau, and et al.** 1994. High functioning fragile X males: demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression. Am J Med Genet **51:**298-308.

68. **Hagerman, R. J., P. McBogg, and P. J. Hagerman.** 1983. The fragile X syndrome: history, diagnosis, and treatment. J Dev Behav Pediatr **4:**122-30.

69. **Hannan, R. D., A. Cavanaugh, W. M. Hempel, T. Moss, and L. Rothblum.** 1999. Identification of a mammalian RNA polymerase I holoenzyme containing components of the DNA repair/replication system. Nucleic Acids Res **27:**3720-7.

70. **Hansen, R. S., T. K. Canfield, A. D. Fjeld, and S. M. Gartler.** 1996. Role of late replication timing in the silencing of X-linked genes. Hum Mol Genet **5:**1345-53.

71. **Hansen, R. S., T. K. Canfield, A. D. Fjeld, S. Mumm, C. D. Laird, and S. M. Gartler.** 1997. A variable domain of delayed replication in FRAXA fragile X

chromosomes: X inactivation-like spread of late replication. Proc Natl Acad Sci U S A **94:**4587-92.

72. **Hansen, R. S., T. K. Canfield, M. M. Lamb, S. M. Gartler, and C. D. Laird.** 1993. Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell **73:**1403-9.

73. **Harikrishnan, K. N., M. Z. Chow, E. K. Baker, S. Pal, S. Bassal, D. Brasacchio, L. Wang, J. M. Craig, P. L. Jones, S. Sif, and A. El-Osta.** 2005. Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nat Genet **37:**254-64.

74. **Harvey, K. J., and J. Newport.** 2003. CpG methylation of DNA restricts prereplication complex assembly in Xenopus egg extracts. Mol Cell Biol **23:**6769-79.

75. **Hay, R. T., and M. L. DePamphilis.** 1982. Initiation of SV40 DNA replication in vivo: location and structure of 5' ends of DNA synthesized in the ori region. Cell **28:**767-79.

76. **Heintz, N. H., and J. L. Hamlin.** 1982. An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc Natl Acad Sci U S A **79:**4083-7.

77. **Hirst, M. C., and P. J. White.** 1998. Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure. Nucleic Acids Res **26:**2353-8.

78. **Howard-Peebles, P. N.** 1980. Fragile sites in human chromosomes II: demonstration of the fragile site Xq27 in carriers of X-linked mental retardation. Am J Med Genet **7:**497-501.

79. **Howard-Peebles, P. N., and J. M. Friedman.** 1985. Unaffected carrier males in families with fragile X syndrome. Am J Hum Genet **37:**956-64.

80. **Hyrien, O.** 2000. Mechanisms and consequences of replication fork arrest. Biochimie **82:**5-17.

81. **Hyrien, O., C. Maric, and M. Mechali.** 1995. Transition in specification of embryonic metazoan DNA replication origins. Science **270:**994-7.

82. **Ireland, M. J., S. S. Reinke, and D. M. Livingston.** 2000. The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast. Genetics **155:**1657-65.

83. **Jacob, F., and S. Brenner.** 1963. [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon.]. C R Hebd Seances Acad Sci **256:**298-300.

84. **Jenke, A. C., I. M. Stehle, F. Herrmann, T. Eisenberger, A. Baiker, J. Bode, F. O. Fackelmayer, and H. J. Lipps.** 2004. Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome. Proc Natl Acad Sci U S A **101:**11322-7.

85. **Jeong, S., and A. Stein.** 1994. Micrococcal nuclease digestion of nuclei reveals extended nucleosome ladders having anomalous DNA lengths for chromatin assembled on non-replicating plasmids in transfected cells. Nucleic Acids Res **22:**370-5.

86. **Jin, Y., T. A. Yie, and A. M. Carothers.** 1995. Non-random deletions at the dihydrofolate reductase locus of Chinese hamster ovary cells induced by alpha-particles simulating radon. Carcinogenesis **16:**1981-91.

87. **Kaguni, J. M.** 2006. DnaA: Controlling the Initiation of Bacterial DNA Replication and More. Annu Rev Microbiol.

88. **Keller, C., O. Hyrien, R. Knippers, and T. Krude.** 2002. Site-specific and temporally controlled initiation of DNA replication in a human cell-free system. Nucleic Acids Res **30:**2114-23.

89. **Keller, C., E. M. Ladenburger, M. Kremer, and R. Knippers.** 2002. The origin recognition complex marks a replication origin in the human TOP1 gene promoter. J Biol Chem **277:**31430-40.

90. **Kemp, M. G., M. Ghosh, G. Liu, and M. Leffak.** 2005. The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells. Nucleic Acids Res **33:**325-36.

91. **Kennedy, G. C., and W. J. Rutter.** 1992. Pur-1, a zinc-finger protein that binds to purine-rich sequences, transactivates an insulin promoter in heterologous cells. Proc Natl Acad Sci U S A **89:**11498-502.

92. **Kobayashi, T., T. Rein, and M. L. DePamphilis.** 1998. Identification of primary initiation sites for DNA replication in the hamster dihydrofolate reductase gene initiation zone. Mol Cell Biol **18:**3266-77.

93. **Kohzaki, H., and Y. Murakami.** 2005. Transcription factors and DNA replication origin selection. Bioessays **27:**1107-16.

94. **Kumari, D., and K. Usdin.** 2001. Interaction of the transcription factors USF1, USF2, and alpha -Pal/Nrf-1 with the FMR1 promoter. Implications for Fragile X mental retardation syndrome. J Biol Chem **276:**4357-64.

95. **Kurose, A., T. Tanaka, X. Huang, F. Traganos, and Z. Darzynkiewicz.** 2006. Synchronization in the cell cycle by inhibitors of DNA replication induces histone H2AX phosphorylation: an indication of DNA damage. Cell Prolif **39:**231-40.

96. **Ladenburger, E. M., C. Keller, and R. Knippers.** 2002. Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication. Mol Cell Biol **22:**1036-48.

97. **Lahue, R. S., and D. L. Slater.** 2003. DNA repair and trinucleotide repeat instability. Front Biosci **8:**s653-65.

98. **Langst, G., P. B. Becker, and I. Grummt.** 1998. TTF-I determines the chromatin architecture of the active rDNA promoter. Embo J **17:**3135-45.

99. **Langst, G., T. A. Blank, P. B. Becker, and I. Grummt.** 1997. RNA polymerase I transcription on nucleosomal templates: the transcription termination factor TTF-I induces chromatin remodeling and relieves transcriptional repression. Embo J **16:**760-8.

100. **Lassar, A. B., P. L. Martin, and R. G. Roeder.** 1983. Transcription of class III genes: formation of preinitiation complexes. Science **222:**740-8.

101. **Lehmann, M.** 2004. Anything else but GAGA: a nonhistone protein complex reshapes chromatin structure. Trends Genet **20:**15-22.

102. **Leibovitch, B. A., Q. Lu, L. R. Benjamin, Y. Liu, D. S. Gilmour, and S. C. Elgin.** 2002. GAGA factor and the TFIID complex collaborate in generating an open chromatin structure at the Drosophila melanogaster hsp26 promoter. Mol Cell Biol **22:**6148-57.

103. **Lemaitre, J. M., E. Danis, P. Pasero, Y. Vassetzky, and M. Mechali.** 2005. Mitotic remodeling of the replicon and chromosome structure. Cell **123:**787-801.

104. **Lenzmeier, B. A., and C. H. Freudenreich.** 2003. Trinucleotide repeat instability: a hairpin curve at the crossroads of replication, recombination, and repair. Cytogenet Genome Res **100:**7-24.

105. **Leu, T. H., and J. L. Hamlin.** 1989. High-resolution mapping of replication fork movement through the amplified dihydrofolate reductase domain in CHO cells by in-gel renaturation analysis. Mol Cell Biol **9:**523-31.

106. **Levenson, V., and J. L. Hamlin.** 1993. A general protocol for evaluating the specific effects of DNA replication inhibitors. Nucleic Acids Res **21:**3997-4004.

107. **Li, C. J., J. A. Bogan, D. A. Natale, and M. L. DePamphilis.** 2000. Selective activation of pre-replication complexes in vitro at specific sites in mammalian nuclei. J Cell Sci **113 (Pt 5):**887-98.

108. **Lieb, J. D., and N. D. Clarke.** 2005. Control of transcription through intragenic patterns of nucleosome composition. Cell **123:**1187-90.

109. **Lieber, M. R.** 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. Bioessays **19:**233-40.

110. **Lim, J. H., A. B. Booker, T. Luo, T. Williams, Y. Furuta, O. Lagutin, G. Oliver, T. D. Sargent, and J. R. Fallon.** 2005. AP-2alpha selectively regulates fragile X mental retardation-1 gene transcription during embryonic development. Hum Mol Genet **14:**2027-34.

111. **Lin, S., and D. Kowalski.** 1997. Functional equivalency and diversity of cis-acting elements among yeast replication origins. Mol Cell Biol **17:**5473-84.

112. **Lipford, J. R., and S. P. Bell.** 2001. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. Mol Cell **7:**21-30.

113. **Liu, G., M. Malott, and M. Leffak.** 2003. Multiple functional elements comprise a Mammalian chromosomal replicator. Mol Cell Biol **23:**1832-42.

114. **Lopez-Estrano, C., J. B. Schvartzman, D. B. Krimer, and P. Hernandez.** 1998. Co-localization of polar replication fork barriers and rRNA transcription terminators in mouse rDNA. J Mol Biol **277:**249-56.

115. **MacAlpine, D. M., and S. P. Bell.** 2005. A genomic view of eukaryotic DNA replication. Chromosome Res **13:**309-26.

116. **MacAlpine, D. M., H. K. Rodriguez, and S. P. Bell.** 2004. Coordination of replication and transcription along a Drosophila chromosome. Genes Dev **18:**3094-105.

117. **Machida, Y. J., J. L. Hamlin, and A. Dutta.** 2005. Right place, right time, and only once: replication initiation in metazoans. Cell **123:**13-24.

118. **Malott, M., and M. Leffak.** 1999. Activity of the c-myc replicator at an ectopic chromosomal location. Mol Cell Biol **19:**5685-95.

119. **Malter, H. E., J. C. Iber, R. Willemsen, E. de Graaff, J. C. Tarleton, J. Leisti, S. T. Warren, and B. A. Oostra.** 1997. Characterization of the full fragile X syndrome mutation in fetal gametes. Nat Genet **15:**165-9.

120. **Marahrens, Y., and B. Stillman.** 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science **255:**817-23.

121. **Mariappan, S. V., P. Catasti, L. A. Silks, 3rd, E. M. Bradbury, and G. Gupta.** 1999. The high-resolution structure of the triplex formed by the GAA/TTC triplet repeat associated with Friedreich's ataxia. J Mol Biol **285:**2035-52.

122. **McMurray, C. T.** 1999. DNA secondary structure: a common and causative factor for expansion in human disease. Proc Natl Acad Sci U S A **96:**1823-5.

123. **Meister, P., A. Taddei, and S. M. Gasser.** 2006. In and out of the replication factory. Cell **125:**1233-5.

124. **Mendez, J., and B. Stillman.** 2000. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol Cell Biol **20:**8602-12.

125. **Mesner, L. D., E. L. Crawford, and J. L. Hamlin.** 2006. Isolating apparently pure libraries of replication origins from complex genomes. Mol Cell **21:**719-26.

126. **Mesner, L. D., and J. L. Hamlin.** 2005. Specific signals at the 3' end of the DHFR gene define one boundary of the downstream origin of replication. Genes Dev **19:**1053-66.

127. **Messer, W.** 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev **26:**355-74.

128. **Miret, J. J., L. Pessoa-Brandao, and R. S. Lahue.** 1998. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A **95:**12438-43.

129. **Mirkin, S. M.** 2006. DNA structures, repeat expansions and human hereditary disorders. Curr Opin Struct Biol **16:**351-8.

130. **Mirkin, S. M., and E. V. Smirnova.** 2002. Positioned to expand. Nat Genet **31:**5-6.

131. **Mitas, M.** 1997. Trinucleotide repeats associated with human disease. Nucleic Acids Res **25:**2245-54.

132. **Moore, H., P. W. Greenwell, C. P. Liu, N. Arnheim, and T. D. Petes.** 1999. Triplet repeats form secondary structures that escape DNA repair in yeast. Proc Natl Acad Sci U S A **96:**1504-9.

133. **Mostoslavsky, R., N. Singh, T. Tenzen, M. Goldmit, C. Gabay, S. Elizur, P. Qi, B. E. Reubinoff, A. Chess, H. Cedar, and Y. Bergman.** 2001. Asynchronous replication and allelic exclusion in the immune system. Nature **414:**221-5.

134. **Moutou, C., M. C. Vincent, V. Biancalana, and J. L. Mandel.** 1997. Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic. Hum Mol Genet **6:**971-9.

135. **Murano, I., A. Kuwano, and T. Kajii.** 1989. Fibroblast-specific common fragile sites induced by aphidicolin. Hum Genet **83:**45-8.

136. **Nenguke, T., M. I. Aladjem, J. F. Gusella, N. S. Wexler, and N. Arnheim.** 2003. Candidate DNA replication initiation regions at human trinucleotide repeat disease loci. Hum Mol Genet **12:**1021-8.

137. **Nichol Edamura, K., M. R. Leonard, and C. E. Pearson.** 2005. Role of replication and CpG methylation in fragile X syndrome CGG deletions in primate cells. Am J Hum Genet **76:**302-11.

138. **Nichol, K., and C. E. Pearson.** 2002. CpG methylation modifies the genetic stability of cloned repeat sequences. Genome Res **12:**1246-56.

139. **O'Donnell, W. T., and S. T. Warren.** 2002. A decade of molecular studies of fragile X syndrome. Annu Rev Neurosci **25:**315-38.

140. **Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz.** 2006. Regulation of TCRbeta gene assembly by a promoter/enhancer holocomplex. Immunity **24:**381-91.

141. **Paixao, S., I. N. Colaluca, M. Cubells, F. A. Peverali, A. Destro, S. Giadrossi, M. Giacca, A. Falaschi, S. Riva, and G. Biamonti.** 2004. Modular structure of the human lamin B2 replicator. Mol Cell Biol **24:**2958-67.

142. **Panigrahi, G. B., J. D. Cleary, and C. E. Pearson.** 2002. In vitro (CTG)*(CAG) expansions and deletions by human cell extracts. J Biol Chem **277:**13926-34.

143. **Pearson, C. E., K. Nichol Edamura, and J. D. Cleary.** 2005. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet **6:**729-42.

144. **Pelizon, C., S. Diviacco, A. Falaschi, and M. Giacca.** 1996. High-resolution mapping of the origin of DNA replication in the hamster dihydrofolate reductase gene domain by competitive PCR. Mol Cell Biol **16:**5358-64.

145. **Pelletier, R., M. M. Krasilnikova, G. M. Samadashwily, R. Lahue, and S. M. Mirkin.** 2003. Replication and expansion of trinucleotide repeats in yeast. Mol Cell Biol **23:**1349-57.

146. **Pietrobono, R., M. G. Pomponi, E. Tabolacci, B. Oostra, P. Chiurazzi, and G. Neri.** 2002. Quantitative analysis of DNA demethylation and transcriptional reactivation of the FMR1 gene in fragile X cells treated with 5-azadeoxycytidine. Nucleic Acids Res **30:**3278-85.

147. **Pietrobono, R., E. Tabolacci, F. Zalfa, I. Zito, A. Terracciano, U. Moscato, C. Bagni, B. Oostra, P. Chiurazzi, and G. Neri.** 2005. Molecular dissection of the events leading to inactivation of the FMR1 gene. Hum Mol Genet **14:**267-77.

148. **Pollard, L. M., R. Sharma, M. Gomez, S. Shah, M. B. Delatycki, L. Pianese, A. Monticelli, B. J. Keats, and S. I. Bidichandani.** 2004. Replication-mediated instability of the GAA triplet repeat mutation in Friedreich ataxia. Nucleic Acids Res **32:**5962-71.

149. **Price, G. B., M. Allarakhia, N. Cossons, T. Nielsen, M. Diaz-Perez, P. Friedlander, L. Tao, and M. Zannis-Hadjopoulos.** 2003. Identification of a cis-element that determines autonomous DNA replication in eukaryotic cells. J Biol Chem **278:**19649-59.

150. **Puck, T. T., and P. I. Marcus.** 1955. A rapid method for viable cell titration and clone production with Hela cells in tissue culture: the use of X-irradiated cells to supply conditioning factors. Proc Natl Acad Sci U S A **41:**432-7.

151. **Putter, V., and F. Grummt.** 2002. Transcription termination factor TTF-I exhibits contrahelicase activity during DNA replication. EMBO Rep **3:**147-52.

152. **Raghuraman, M. K., E. A. Winzeler, D. Collingwood, S. Hunt, L. Wodicka, A. Conway, D. J. Lockhart, R. W. Davis, B. J. Brewer, and W. L. Fangman.** 2001. Replication dynamics of the yeast genome. Science **294:**115-21.

153. **Rao, B. S., H. Manor, and R. G. Martin.** 1988. Pausing in simian virus 40 DNA replication by a sequence containing (dG-dA)27.(dT-dC)27. Nucleic Acids Res **16:**8077-94.

154. **Refsland, E. W., and D. M. Livingston.** 2005. Interactions among DNA ligase I, the flap endonuclease and proliferating cell nuclear antigen in the expansion and contraction of CAG repeat tracts in yeast. Genetics **171:**923-34.

155. **Rein, T., T. Kobayashi, M. Malott, M. Leffak, and M. L. DePamphilis.** 1999. DNA methylation at mammalian replication origins. J Biol Chem **274:**25792-800.

156. **Rosche, W. A., T. Q. Trinh, and R. R. Sinden.** 1995. Differential DNA secondary structure-mediated deletion mutation in the leading and lagging strands. J Bacteriol **177:**4385-91.

157. **Rossi, M. L., V. Purohit, P. D. Brandt, and R. A. Bambara.** 2006. Lagging strand replication proteins in genome stability and DNA repair. Chem Rev **106:**453-73.

158. **Saha, S., Y. Shan, L. D. Mesner, and J. L. Hamlin.** 2004. The promoter of the Chinese hamster ovary dihydrofolate reductase gene regulates the activity of the local origin and helps define its boundaries. Genes Dev **18:**397-410.

159. **Samadashwily, G. M., G. Raca, and S. M. Mirkin.** 1997. Trinucleotide repeats affect DNA replication in vivo. Nat Genet **17:**298-304.

160. **Sasaki, T., S. Ramanathan, Y. Okuno, C. Kumagai, S. S. Shaikh, and D. M. Gilbert.** 2006. The Chinese hamster dihydrofolate reductase replication origin decision point follows activation of transcription and suppresses initiation of replication within transcription units. Mol Cell Biol **26:**1051-62.

161. **Sasaki, T., T. Sawado, M. Yamaguchi, and T. Shinomiya.** 1999. Specification of regions of DNA replication initiation during embryogenesis in the 65-kilobase DNApolalpha-dE2F locus of Drosophila melanogaster. Mol Cell Biol **19:**547-55.

162. **Schaarschmidt, D., E. M. Ladenburger, C. Keller, and R. Knippers.** 2002. Human Mcm proteins at a replication origin during the G1 to S phase transition. Nucleic Acids Res **30:**4176-85.

163. **Schepers, A., M. Ritzi, K. Bousset, E. Kremmer, J. L. Yates, J. Harwood, J. F. Diffley, and W. Hammerschmidt.** 2001. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. Embo J **20:**4588-602.

164. **Schwaiger, M., and D. Schubeler.** 2006. A question of timing: emerging links between transcription and replication. Curr Opin Genet Dev **16:**177-83.

165. **Schweitzer, J. K., and D. M. Livingston.** 1999. The effect of DNA replication mutations on CAG tract stability in yeast. Genetics **152:**953-63.

166. **Schweitzer, J. K., and D. M. Livingston.** 1998. Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. Hum Mol Genet **7:**69-74.

167. **Sen, R., and E. Oltz.** 2006. Genetic and epigenetic regulation of IgH gene assembly. Curr Opin Immunol **18:**237-42.

168. **Shimizu, M., R. Gellibolian, B. A. Oostra, and R. D. Wells.** 1996. Cloning, characterization and properties of plasmids containing CGG triplet repeats from the FMR-1 gene. J Mol Biol **258:**614-26.

169. **Sibani, S., G. B. Price, and M. Zannis-Hadjopoulos.** 2005. Decreased origin usage and initiation of DNA replication in haploinsufficient HCT116 Ku80+/- cells. J Cell Sci **118:**3247-61.

170. **Sikes, M. L., A. Meade, R. Tripathi, M. S. Krangel, and E. M. Oltz.** 2002. Regulation of V(D)J recombination: a dominant role for promoter positioning in gene segment accessibility. Proc Natl Acad Sci U S A **99:**12309-14.

171. **Simon, I., T. Tenzen, R. Mostoslavsky, E. Fibach, L. Lande, E. Milot, J. Gribnau, F. Grosveld, P. Fraser, and H. Cedar.** 2001. Developmental regulation of DNA replication timing at the human beta globin locus. Embo J **20:**6150-7.

172. **Simpson, R. T.** 1990. Nucleosome positioning can affect the function of a cis-acting DNA element in vivo. Nature **343:**387-9.

173. **Slater, M. L.** 1973. Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle. J Bacteriol **113:**263-70.

174. **Smeets, H. J., A. P. Smits, C. E. Verheij, J. P. Theelen, R. Willemsen, I. van de Burgt, A. T. Hoogeveen, J. C. Oosterwijk, and B. A. Oostra.** 1995. Normal phenotype in two brothers with a full FMR1 mutation. Hum Mol Genet **4:**2103-8.

175. **Smith, K. T., B. Coffee, and D. Reines.** 2004. Occupancy and synergistic activation of the FMR1 promoter by Nrf-1 and Sp1 in vivo. Hum Mol Genet **13:**1611-21.

176. **Smith, K. T., R. D. Nicholls, and D. Reines.** 2006. The gene encoding the fragile X RNA-binding protein is controlled by nuclear respiratory factor 2 and the CREB family of transcription factors. Nucleic Acids Res **34:**1205-15.

177. **Snapka, R. M., and P. A. Permana.** 1993. SV40 DNA replication intermediates: analysis of drugs which target mammalian DNA replication. Bioessays **15:**121-7.

178. **Song, J., H. Murakami, H. Tsutsui, H. Ugai, C. Geltinger, T. Murata, M. Matsumura, K. Itakura, I. Kanazawa, K. Sun, and K. K. Yokoyama.** 1999. Structural organization and expression of the mouse gene for Pur-1, a highly conserved homolog of the human MAZ gene. Eur J Biochem **259:**676-83.

179. **Stillman, B.** 2001. DNA replication. Genomic views of genome duplication. Science **294:**2301-4.

180. **Stinchcomb, D. T., K. Struhl, and R. W. Davis.** 1979. Isolation and characterisation of a yeast chromosomal replicator. Nature **282:**39-43.

181. **Subramanian, P. S., D. L. Nelson, and A. C. Chinault.** 1996. Large domains of apparent delayed replication timing associated with triplet repeat expansion at FRAXA and FRAXE. Am J Hum Genet **59:**407-16.

182. **Sutherland, G. R.** 1988. The role of nucleotides in human fragile site expression. Mutat Res **200:**207-13.

183. **Takeda, D. Y., Y. Shibata, J. D. Parvin, and A. Dutta.** 2005. Recruitment of ORC or CDC6 to DNA is sufficient to create an artificial origin of replication in mammalian cells. Genes Dev **19:**2827-36.

184. **Tao, L., Z. Dong, M. Leffak, M. Zannis-Hadjopoulos, and G. Price.** 2000. Major DNA replication initiation sites in the c-myc locus in human cells. J Cell Biochem **78:**442-57.

185. **Tassone, F., R. J. Hagerman, L. W. Gane, and A. K. Taylor.** 1999. Strong similarities of the FMR1 mutation in multiple tissues: postmortem studies of a male with a full mutation and a male carrier of a premutation. Am J Med Genet **84:**240-4.

186. **Terracciano, A., P. Chiurazzi, and G. Neri.** 2005. Fragile X syndrome. Am J Med Genet C Semin Med Genet **137:**32-7.

187. **Theis, J. F., and C. S. Newlon.** 2001. Two compound replication origins in Saccharomyces cerevisiae contain redundant origin recognition complex binding sites. Mol Cell Biol **21:**2790-801.

188. **Thiriet, C., and J. J. Hayes.** 1998. Functionally relevant histone-DNA interactions extend beyond the classically defined nucleosome core region. J Biol Chem **273:**21352-8.

189. **Todorovic, V., A. Falaschi, and M. Giacca.** 1999. Replication origins of mammalian chromosomes: the happy few. Front Biosci **4:**D859-68.

190. **Todorovic, V., S. Giadrossi, C. Pelizon, R. Mendoza-Maldonado, H. Masai, and M. Giacca.** 2005. Human origins of DNA replication selected from a library of nascent DNA. Mol Cell **19:**567-75.

191. **Van Esch, H.** 2006. The Fragile X premutation: new insights and clinical consequences. Eur J Med Genet **49:**1-8.

192. **van Steensel, B., J. Delrow, and H. J. Bussemaker.** 2003. Genomewide analysis of Drosophila GAGA factor target genes reveals context-dependent DNA binding. Proc Natl Acad Sci U S A **100:**2580-5.

193. **Veaute, X., and R. P. Fuchs.** 1993. Greater susceptibility to mutations in lagging strand of DNA replication in Escherichia coli than in leading strand. Science **261:**598-600.

194. **Verdin, E., P. Paras, Jr., and C. Van Lint.** 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. Embo J **12:**3249-59.

195. **Vermaak, D., and A. P. Wolffe.** 1998. Chromatin and chromosomal controls in development. Dev Genet **22:**1-6.

196. **Verschure, P. J., A. E. Visser, and M. G. Rots.** 2006. Step out of the groove: epigenetic gene control systems and engineered transcription factors. Adv Genet **56:**163-204.

197. **Vogel, W., W. Schempp, and I. Sigwarth.** 1978. Comparison of thymidine, fluorodeoxyuridine, hydroxyurea, and methotrexate blocking at the G1/S phase transition of the cell cycle, studied by replication patterns. Hum Genet **45:**193-8.

198. **Vogelauer, M., L. Rubbi, I. Lucas, B. J. Brewer, and M. Grunstein.** 2002. Histone acetylation regulates the time of replication origin firing. Mol Cell **10:**1223-33.

199. **Vujcic, M., C. A. Miller, and D. Kowalski.** 1999. Activation of silent replication origins at autonomously replicating sequence elements near the HML locus in budding yeast. Mol Cell Biol **19:**6098-109.

200. **Wallisch, M., E. Kunkel, K. Hoehn, and F. Grummt.** 2002. Ku antigen supports termination of mammalian rDNA replication by transcription termination factor TTF-I. Biol Chem **383:**765-71.

201. **Walter, J., and J. W. Newport.** 1997. Regulation of replicon size in Xenopus egg extracts. Science **275:**993-5.

202. **Wang, J., S. E. Lindner, E. R. Leight, and B. Sugden.** 2006. Essential elements of a licensed, mammalian plasmid origin of DNA synthesis. Mol Cell Biol **26:**1124-34.

203. **Wang, L., C. M. Lin, S. Brooks, D. Cimbora, M. Groudine, and M. I. Aladjem.** 2004. The human beta-globin replication initiation region consists of two modular independent replicators. Mol Cell Biol **24:**3373-86.

204. **Wang, L., C. M. Lin, J. O. Lopreiato, and M. I. Aladjem.** 2006. Cooperative sequence modules determine replication initiation sites at the human {beta}-globin locus. Hum Mol Genet **15:**2613-22.

205. **Wang, S., P. A. Dijkwel, and J. L. Hamlin.** 1998. Lagging-strand, early-labelling, and two-dimensional gel assays suggest multiple potential initiation sites in the Chinese hamster dihydrofolate reductase origin. Mol Cell Biol **18:**39-50.

206. **Wang, Y. V., H. Tang, and D. S. Gilmour.** 2005. Identification in vivo of different rate-limiting steps associated with transcriptional activators in the presence and absence of a GAGA element. Mol Cell Biol **25:**3543-52.

207. **Weinmann, A. S., S. E. Plevy, and S. T. Smale.** 1999. Rapid and selective remodeling of a positioned nucleosome during the induction of IL-12 p40 transcription. Immunity **11:**665-75.

208. **White, P. J., R. H. Borts, and M. C. Hirst.** 1999. Stability of the human fragile X (CGG)(n) triplet repeat array in Saccharomyces cerevisiae deficient in aspects of DNA metabolism. Mol Cell Biol **19:**5675-84.

209. **Wilkins, R. C., and J. T. Lis.** 1997. Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation. Nucleic Acids Res **25:**3963-8.

210. **Wilmes, G. M., and S. P. Bell.** 2002. The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. Proc Natl Acad Sci U S A **99:**101-6.

211. **Wohrle, D., U. Salat, D. Glaser, J. Mucke, M. Meisel-Stosiek, D. Schindler, W. Vogel, and P. Steinbach.** 1998. Unusual mutations in high functioning fragile X males: apparent instability of expanded unmethylated CGG repeats. J Med Genet **35:**103-11.

212. **Wohrle, D., U. Salat, H. Hameister, W. Vogel, and P. Steinbach.** 2001. Demethylation, reactivation, and destabilization of human fragile X full-mutation alleles in mouse embryocarcinoma cells. Am J Hum Genet **69:**504-15.

213. **Wolffe, A. P., and H. Kurumizaka.** 1998. The nucleosome: a powerful regulator of transcription. Prog Nucleic Acid Res Mol Biol **61:**379-422.

214. **Wyrick, J. J., J. G. Aparicio, T. Chen, J. D. Barnett, E. G. Jennings, R. A. Young, S. P. Bell, and O. M. Aparicio.** 2001. Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science **294:**2357-60.

215. **Yan, P., X. Mao, L. Wang, X. Zha, and C. Lu.** 2002. HBV C promoter Sp1 binding sequence functionally substitutes for the yeast ARS1 ABF1 binding site. DNA Cell Biol **21:**737-42.

216. **Yang, Z., and J. J. Hayes.** 2003. Xenopus transcription factor IIIA and the 5S nucleosome: development of a useful in vitro system. Biochem Cell Biol **81:**177-84.

217. **Yang, Z., R. Lau, J. L. Marcadier, D. Chitayat, and C. E. Pearson.** 2003. Replication inhibitors modulate instability of an expanded trinucleotide repeat at the myotonic dystrophy type 1 disease locus in human cells. Am J Hum Genet **73:**1092-105.

218. **Yeshaya, J., R. Shalgi, M. Shohat, and L. Avivi.** 1999. FISH-detected delay in replication timing of mutated FMR1 alleles on both active and inactive X-chromosomes. Hum Genet **105:**86-97.

219. **Yu, A., M. D. Barron, R. M. Romero, M. Christy, B. Gold, J. Dai, D. M. Gray, I. S. Haworth, and M. Mitas.** 1997. At physiological pH, d(CCG)15 forms a hairpin containing protonated cytosines and a distorted helix. Biochemistry **36:**3687-99.

220. **Yuan, G. C., Y. J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando.** 2005. Genome-scale identification of nucleosome positions in S. cerevisiae. Science **309:**626-30.

221. **Zhou, J., C. Chau, Z. Deng, W. Stedman, and P. M. Lieberman.** 2005. Epigenetic control of replication origins. Cell Cycle **4:**889-92.