MECHANISMS OF COMMON FRAGILE SITE INSTABILITY AND CANCER

by

Ryan L. Ragland

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Human Genetics)
in The University of Michigan
2009

Doctoral Committee:

Professor Thomas W. Glover, Chair
Professor Sally A. Camper
Associate Professor Mats E. Ljungman
Assistant Professor David O. Ferguson
Assistant Professor Joann Sekiguchi

For my family and friends.

Acknowledgments

First and foremost I would like to acknowledge my family and friends without whom I would not be the person or scientist that I am today. This is exceptionally true for my wife Mary, whose love and support made this entire process possible. I would like to acknowledge my committee and its chair and my mentor Dr. Thomas Glover for all their support and understanding throughout these years. Also I would like to acknowledge Dr. Martin Arlt who, whether he likes it or not, is personally responsible for my development as a scientist. I would like to acknowledge the University of Michigan comprehensive cancer center in general and Dr. Eric Fearon in specific for their unwavering financial and scientific support of my work. Finally, special thanks to all of the members of the Glover lab past and present who have made my time here memorable and enjoyable and whose expertise was indispensable.

Table of Contents

II. Stably transfected common fragile site sequences exhibit instability at ectopic sites

III. Mice with hypomorphic Atr expression exhibit increased APH-induced DNA damage and checkpoint abnormalities.

IV. Discussion and Future Directions

List of Figures

List of Tables

Chapter I

Introduction

Background

Every year over 10 million people are diagnosed with cancer, a disease that accounts for approximately twelve percent of all deaths worldwide (http://www.who.int/cancer/ en/). Significantly, although there are over 200 different types of cancer, the fundamental basis for all cancers is the accumulation of genetic abnormalities. There are currently an estimated 25,000 genes in the human genome. However, only an estimated 300 of these genes are known to play a role in the development of cancer, suggesting that although cancer is extremely complex, the pool of genes involved is limited.

It is generally accepted that one cell must gain at least six major characteristics before it is considered to have tumorigenic potential (Hanahan and Weinberg, 2000). Those characteristics are: the ability to evade apoptosis; limitless growth potential; insensitivity to anti-growth signals; self-sustained growth signals; sustained angiogenesis; and the ability to invade local tissues and to metastasize. The accumulation of any one of these characteristics may be caused by the mutation of one or several different genes. Given that the normal rate for accumulation of mutation in the human genome is approximately 175 mutations per diploid genome per generation, and given the number

of genes that must be mutated before a cell can become transformed, it is clear that rate of cancer formation cannot be accounted for by the normal rate of accumulation of mutations (Nachman and Crowell, 2000).

Alteration of genetic factors, such as genes involved in the sensing and repair of DNA damage, can increase the rate of mutation accumulation and have been broadly categorized as "enabling characteristics" of tumorigenesis. An increase in mutation rate can be due to a variety of additional factors including long term or intense exposure to environmental carcinogens such as ultraviolet radiation (UV), exposure to infectious agents such as the human papilloma virus, or a genetic predisposition to tumorigenesis caused by inheritance of altered genes. Regardless of the origin, the ultimate outcome of an increase in any of these factors is to amplify the rate of accumulation of DNA alterations thus leading to an increase in the rate of tumorigenesis.

Not all loci in the human genome are equivalent in their susceptibility or response to mutagenesis. There are some sites, referred to as fragile sites, that are significantly more susceptible to certain types of DNA damage. Fragile sites are defined as chromosomal loci that are particularly sensitive to replication stress, forming cytogenetically visible gaps and breaks on metaphase chromosomes (Figure 1-1). These sites are broadly classified into two main categories, rare and common, based on their population frequency, pattern of inheritance, and method of induction (Durkin and Glover, 2007). The first site defined as "fragile" was discovered in 1970 by Loveiren et. al, who described a heritable recurrent chromosome break in the long arm of chromosome 16 (Magenis et al., 1970). Following this discovery, a number of families

with X-linked mental retardation and associated rare fragile sites were described (Harvey et al., 1977).

Rare fragile sites

Rare fragile sites are seen in a small proportion of individuals (<5%) and are inherited in a Mendelian manner. Some, such as FRAXA, are associated with human genetic disorders and their study has led to the identification of nucleotide-repeat expansion as a frequent mutational mechanism in humans. There are currently 31 known rare fragile sites. Nine of these have been cloned; FRAXA (Xq27.3), FRAXE (Xq28), FRAXF (Xq28), FRA10A (10q23.3), FRA10B (10q25.2), FRA11B (11q23.3), FRA12A (12q13.1), FRA16A (16q13.11), and FRA16B (16q22.1) (Reviewed in (Sutherland, 2003)).

Rare fragile sites are subdivided into groups based on their mode of induction. Most are folate sensitive, induced by inhibition of folic acid metabolism and leading to perturbation of DNA synthesis. Other rare fragile sites are induced by compounds such as distamycin-A, which binds to the AT-rich DNA sequences present at these sites. The expression of all rare fragile sites studied at the molecular level is dependent on an expansion of repeat DNA sequences. Most, such as the folate sensitive fragile site FRAXA (Xq27.3) in the *FMR-1* gene, are caused by an expansion of CGG trinucleotide repeats. As an example, normal individuals have anywhere from 1-50 CGG repeats. In the disease state, the number of repeat units increases dramatically to 200-2000 repeats. Trinucleotide repeats are inherited in a Mendelian manner but also follow a phenomenon known as anticipation, whereby an increase in the repeat unit number is associated with

an accompanying increase in the severity of symptoms and an earlier age of onset of disease in succeeding generations. Other rare fragile sites are associated with an expanded AT-rich minisatellite repeat sequence. The rare fragile sites FRA16B (16q22.1) and FRA10B (10q25.2) are examples of expansion of the existing AT-rich minisatellite repeats from just a few copies to 33 and 42 base pairs, respectively, which causes fragility.

While rare fragile site repeats exist in normal unaffected individuals. it is only the expanded forms of these sites that are unstable, prone to DNA damage, and associated with disease. Intermediate forms of rare fragile site alleles also exist. These intermediate alleles contain repeat sizes that are between the normal and disease associated repeat length. Intermediate alleles are not normally associated with disease and are not as prone to damage as the more expanded disease associated alleles.

Common fragile sites

Common fragile sites (CFSs) comprise the largest class of fragile sites. These sites are found in all individuals and are considered to be a normal component of the human genome. The first CFSs were discovered when cells were grown under conditions of folate stress to diagnose the instability of the FRAXA locus. During this clinical experiment, it was observed that cells from both normal individuals and individuals with fragile X syndrome, displayed recurrent chromosome breaks not associated with the fragile X locus. Follow up work published in 1984 showed that when cells were cultured in the presence of the DNA polymerase inhibitor aphidicolin (APH), a number of sites in

the genome were specifically and reproducibly induced to form gaps and breaks on metaphase chromosomes (Glover et al., 1984).

Unlike rare fragile sites, the instability of CFSs is not caused by expansion of trinucleotide repeats. CFSs form gaps and breaks on metaphase chromosomes when DNA synthesis is partially inhibited by folate stress or using chemical compounds such as APH, bromodeoxyuridine (BrdU), or 5-azacytidine (Sutherland et al., 1985). In addition to cytogenetically identifiable gaps and breaks, CFSs exhibit a number of other characteristics of unstable DNA in cultured cells, including: gross chromosome deletions (Wang et al., 1993); smaller submicroscopic deletions (Durkin et al., 2008); translocations (Glover and Stein, 1988); intrachromosomal gene amplifications (Coquelle et al., 1997); and sister chromatid exchanges (Glover and Stein, 1987). In cells pretreated with APH, CFSs are preferred sites for the integration of transfected and viral DNA sequences (Rassool et al., 1991; Ferber et al., 2003; Thorland et al., 2003).

Currently, there are over 70 different CFSs with one or more found in most human chromosomes (Figure 1-2). However, the exact number of CFSs is not well defined. This number is not well defined because increased replication stress induces more loci in the genome to express recurrent breaks and be defined as CFSs. Just 20 of the more than 70 CFSs account for more than 80% of all gaps and breaks seen in aphidicolin-treated lymphocytes (Glover et al., 1984). The most highly expressed of these sites are FRA3B (3p14.2) and FRA16D (16q23), which accounts for 25-34% of all chromosome gaps and breaks found after low dose APH treatment in human lymphocytes (Glover et al., 1984).

The boundaries of CFSs are not well defined. A number of groups have used FISH experiments with multiple tiled DNA probes, demarcating whether a single probe had most breaks crossing, distal, or proximal, to the FISH signal, in an attempt to refine the region that is most prone to forming gaps and breaks (Wilke et al., 1994; Arlt et al., 2002). Unlike rare fragile sites, CFS span several hundred kilobases to over a megabase in size and gaps and breaks can occur anywhere within these large regions. Despite their size, a number of these sites are fully contained within large genes such as the tumor suppressor genes fragile histidine triad (*FHIT***)** at 3p14 and WW domain-containing oxidoreductase (*WWOX)* at 16q23. Both the sequence and the fragility of these CFSs is conserved during mammalian evolution. Orthologs of human CFSs have been found in the syntenic regions of a number of other mammalian species, including: other primates; cat; dog; pig; horse; cow; Indian mole rat; deer mouse; and laboratory mouse strongly suggesting an underlying function for these sites in the normal functioning of the cell (Soulie and De Grouchy, 1981; Elder and Robinson, 1989; Smeets and van de Klundert, 1990; Stone et al., 1991; McAllister and Greenbaum, 1997; Glover et al., 1998; Ruiz-Herrera et al., 2004). Furthermore, replication slow zones exist in yeast and may function as counterparts to mammalian CFSs (Roeder and Fink, 1980; Dunham et al., 2002; Lemoine et al., 2005).

Mechanisms of CFS expression

Although there are a large number of loci defined as CFSs, we do not currently know what makes these loci so very different from non-fragile loci. It has been shown repeatedly that APH, an inhibitor of replicative DNA polymerases, causes preferential

damage at CFSs suggesting that these sites are more sensitive to this form of replication stress than the rest of the genome (Glover et al., 1984). During normal replication, the helicase and polymerase complexes are closely associated with each other and move at similar rates.

The helicase complex is responsible for unwinding and separating the DNA strands ahead of the rest of the replication machinery. This is necessary for accurate replication of DNA by the polymerase complex. However, when replication forks encounter sequences that are particularly difficult to replicate through or are inhibited in some other manner, such as by the addition of low dose APH, separation of the helicase and polymerase complexes can occur (Byun et al., 2005). The separation of these two complexes increases the amount of single stranded unreplicated DNA that exists between them. In most cases, an increase in the amount of single strand DNA between these complexes does not cause the cell distress. However, in the case of CFSs It is hypothesized that these sites contain many more sequences that are capable of forming secondary structures than is average for the rest of the genome (Boldog et al., 1997; Mishmar et al., 1998; Arlt et al., 2002; Mirkin and Mirkin, 2007; Zhang and Freudenreich, 2007). Therefore, an increase in the size of the single strand DNA region between the polymerase and helicase complexes at a CFS could lead to the preferential formation of secondary structures. An increase in the formation of secondary structures could serve to inhibit and slow replication leading to the preferential formation of gaps and breaks on metaphase chromosomes and CFSs (Figure 1-3).

Additional support for this hypothesis is found when cells are treated with both the polymerase inhibitor APH and the topoisomerase I inhibitor camptothecin (CPT). It

has been shown that APH is capable of uncoupling the replicative polymerases from the helicase complex and inducing gaps and breaks at CFSs (Byun et al., 2005). In a recent series of unpublished experiments, Arlt et al. have demonstrated that APH-induced CFS damage is prevented when cells are treated with low doses CPT (Arlt et al. unpublished data). In addition to a reduction of CFS expression, it was shown that there was also a reduction in overall levels of phosphorylated RPA and foci formation. These data suggest that in APH treated cells the further addition of CPT brings about a reduction in the size of the single strand region between the polymerase and helicase complexes. This reduction of single strand DNA between these complexes could serve to reduce the possibility of secondary structure formation, and given the proposed hypothesis, could account for the reduction in CFS expression. Extrapolation of this hypothesis suggests that anything capable of stalling replication forks or causing an increase in the separation of the polymerase and helicase complexes through some other means, could potentially contribute to the fragility of CFSs. Currently, the two best candidates that are unique in some way to CFSs are the association of late replication timing with CFSs and the genetic sequence of CFSs.

Replication timing of CFSs

Strong association of late replication with many CFSs has been demonstrated repeatedly (Le Beau et al., 1998; Wang et al., 1999; Hellman et al., 2000; Pelliccia et al., 2008). In addition, treatment of cells with APH perturbs the replication timing of these loci. CFSs are normally very late replicating, perhaps some of the last sites in the genome to replicate, completing replication in late S and even into G2 phase of the cell

cycle. Because these sites are very late replicating, they are especially prone to further replication delay as would be brought on by the addition of APH. Given these data, it is possible that CFSs may be entering mitosis incompletely replicated, or perhaps, not fully condensed leading to the formation of the gaps and breaks that we observe on metaphase chromosomes.

This assertion is supported by the fact that as the dose of APH increases, and replication timing is further delayed, more loci become "fragile" and complete chromosome fragmentation is observed. This finding suggest that any site in the genome can become "fragile" if its replication is completed late enough. Rare fragile sites have also been found to be associated with late replication timing and alleles with expanded repeats are later replicating than their normal counterparts (Hansen et al., 1993; Hansen et al., 1997; Handt et al., 2000; Zlotorynski et al., 2003). Finally, Wang et al. found that there was an allele specific relationship between late replication and overall fragility at the FRA3B locus (Wang et al., 1999). While it is clear that a number of CFSs are associated with late replication, and that there is a correlation between APH induced replication delay and an increase in gaps and breaks, it has not been shown experimentally that late replication is either necessary or sufficient for the increased instability of CFSs.

Sequence analysis of CFSs

The genomic instability associated with rare fragile sites is caused by the expansion of a single tri- or di- nucleotide sequence (reviewed in (Sutherland, 2003)). This implies that if sequence is involved in CFS fragility there may be a similar sequence

motif present in these sites. However, repeated in-depth examination of the sequence of CFSs by several different groups has not uncovered any single sequence that could account for the fragility of CFSs. While similar in appearance to rare fragile sites at the cytogenetic level, it is currently believed that the expression of CFSs is not caused by the expansion of di- or tri- nucleotide or other simple repeat sequences. Gaps and breaks can occur anywhere in the large (0.5 to >3Mb) CFS regions. This suggests that sequences contributing to the instability of these sites must be present throughout the entire CFS region.

All cloned common fragile sites are AT-rich and contain numerous repeat elements such as LINEs, SINEs, and other long repeat elements. In addition to these sequence features, several groups have proposed AT di-nucleotide repeates and flexibility peaks as possible sequence motifs that may contribute to the instability of CFSs. Flexibility peaks are defined as sequences that have a high local variation in DNA twist angle as measured using the TwistFlex program (Mishmar et al., 1998). Many CFSs contain a high number of flexibility peaks as compared to control sequences. Based on this finding, it has been suggested that the formation of abnormal DNA structures at flexibility peaks during replication could be a causal factor in the instability of CFSs (Boldog et al., 1997; Mishmar et al., 1998; Arlt et al., 2002; Mirkin and Mirkin, 2007). In support of this hypothesis, it has recently been shown that in yeast, an AT-rich sequence containing variable stretches of perfect AT repeats taken from a flexibility peak found in the human CFS FRA16D, can cause fork stalling and chromosome breakage (Zhang and Freudenreich, 2007). While this particular motif is principal in the literature, these findings do not rule out the contribution of additional sequence motifs that are

capable of forming strong secondary structures. In support of this assertion, Rozier et al. found that high repetitive element content may contribute to the instability of CFSs (Rozier et al., 2004).

Cellular regulation of CFSs

Throughout life, cells are exposed to a wide variety of chemical, biological, and physical events that can result in alterations of the DNA sequence. These alterations can occur in a wide variety of forms such as: direct DNA damage; single strand DNA breaks; double strand DNA breaks; bulky adducts; and can also occur due to faulty repair of existing DNA damage leading to insertions, deletions, translocations and other phenomenon. Because the correct sensing and repair of damaged DNA is fundamentally important to the viability of the individual cell, and the organism as a whole, cells have evolved many different protein pathways to sense DNA damage, arrest the cell cycle, and repair existing damage accurately. In extreme cases, when the cellular machinery is not sufficient to repair damage, these pathways can trigger apoptosis in an effort to prevent the replication of cells with unstable genomes. This decisive mechanism is designed to preserve the whole organism at the expense of single cells. These cellular responses can be activated at various stages of the cell cycle and are collectively termed DNA damage checkpoints. If these checkpoints are defective and damaged DNA is allowed to replicate, it can lead to an accumulation of genomic alterations that are the root cause of the malignant transformation of a normal cell.

CFSs are particularly sensitive to stalled replication forks. Stalled replication forks do not cause DNA damage intrinsically but can activate cell cycle checkpoints and

permit the formation of complex secondary structures, single strand DNA breaks, and if not correctly sensed or repaired, can become double strand DNA breaks. Stalled replication forks are regulated primarily by the checkpoint protein Ataxia Telengiectasia and Rad3 Related (ATR) (Casper et al., 2002). Because CFSs are particularly prone to forming stalled replication forks, they are also particularly sensitive to loss of functional levels of the proteins that regulate these events, such as ATR (Casper et al., 2002; Alderton et al., 2004; Casper et al., 2004). Specifically, it was found that in human cells in which functional levels of ATR were reduced using either siRNA, dominant negative kinase dead ATR, or cre-lox mediated deletion of ATR, there was a strong, statistically significant increase in both total gaps and breaks per metaphase and gaps and breaks at the common fragile sites FRA3B and FRA16D (Casper et al., 2002). Examination of known downstream targets of ATR suggested several candidate genes that could also regulate the stability of CFSs.

Because it is directly affected by the loss of ATR and because of its importance as a tumor suppressor gene, Breast Cancer Type 1 (BRCA1) was the first gene, post ATR, to be analyzed for its role in regulating the stability of CFSs. Arlt et al. showed that in HCC1937 line a BRCA1-null breast cancer cell line, in mouse embryo fibroblast line with inactivating mutations in both p53 and Brca1, and in HeLa cells depleted for BRCA1 by the addition of a BRCA1 specific siRNA, there was a statistically significant increase in gaps and breaks on metaphase chromosomes both overall and specifically at CFSs (Arlt et al., 2004). In addition, these experiments suggested that there was a specific role for BRCA1 in the G2/M phases of the cell cycle as proteins with G2/M specific deficiency were unable to prevent CFS breaks.

Both Cell Cycle Checkpoint Kinase 1 (CHK1) and Cell Cycle Checkpoint Kinase 2 (CHK2) are downstream phosphorylation targets of ATR (Durkin et al., 2006). Using similar siRNA techniques as described above, human cells were depleted for either one or both of these proteins. It was found that under conditions of replication stress, depletion of CHK1 but not CHK2 caused an increase in gaps and breaks at the CFSs FRA3B and FRA16D (Durkin et al., 2006). Depletion of both CHK1 and CHK2 showed no significant increase in gaps and breaks compared to depletion of CHK1 alone. This finding further defines a role for the ATR pathway and CHK1 in the regulation of stalled replication forks at CFSs, and suggests that the ATM pathway, via CHK2, is not required for this regulation.

The Fanconi Anemia Complementation group D2 protein (FANCD2) is activated by monoubiquitination in normal cells in response to DNA damage. It has been shown that FANCD2 is targeted to nuclear foci that colocalize with BRCA1 (Schreiber, 1979). Using both a patient derived cell line lacking FANCD2 expression and RNAi techniques, Howlett et al. examined the role of the Fanconi anemia pathway in the regulation of CFSs (Howlett et al., 2005). It was found that depletion of FANCD2 in the presence of low dose APH or hydroxyurea resulted in a strong statistically significant increase in total gaps and breaks and gaps and breaks at CFSs (Howlett et al., 2005). Furthermore, the addition of low dose APH also led to a strong activation of the Fanconi Anemia pathway in the form of increased monoubiquitinated FANCD2. These results suggested that FANCD2 is also required for correct maintenance of APH-induced DNA damage.

The structural maintenance of chromosomes (SMC) family of genes includes SMC1 and SMC3 which form a heterodimer involved in the regulation of sister

chromatid cohesion and repair of gaps and breaks (Strunnikov and Jessberger, 1999). Musio et al. (Musio et al., 2005) showed that depletion of either SMC1 or SMC3 via siRNA caused an increase in CFS expression in APH treated normal human fibroblasts. Furthermore, they found that SMC1 is phosphorylated in an ATR-dependent and not an ATM-independent manner (Musio et al., 2005). As with the CHK1 results, this finding suggested that the regulation of damage at CFSs is primarily dependent on the ATR and not the ATM pathway.

Hydroxyurea Sensitive 1 (HUS1) is a cell cycle checkpoint gene that is phosphorylated in response to DNA damage (Kostrub et al., 1998). Complete loss of Hus1 in mice results in lethality in both mouse embryos and cell culture (Zhu and Weiss, 2007). However, conditional loss of Hus1 in a mouse model causes an increase in CFS instability, apoptosis, double strand DNA breaks as monitored by γ–H2ax formation, and a reduction in cell doubling time (Zhu and Weiss, 2007). These results indicate that HUS1 is also important in the regulation of damage at CFSs.

While these data suggest that it is the ATR and not the ATM pathway that is primarily responsible for sensing and signaling damage at CFSs, a recent finding by Ozeri-Galai et. al (Ozeri-Galai et al., 2008) found that in the absence of ATR, the ATM pathway is activated and responsible for the avoidance of damage at CFSs. This finding is not surprising as although both ATR and ATM primarily respond to different forms of DNA damage, there is a considerable amount of overlap of function between these two proteins. It has been shown that ATR is activated in an ATM-dependent manner following IR (Myers and Cortez, 2006) and that ATM is activated in an ATR-dependent manner following UV irradiation (Stiff et al., 2006). Given these data, fully

understanding the relationship between ATR, ATM, and the regulation of damage at CFSs may be more complex than originally believed.

All of these studies help us understand and define the cellular machinery that is monitoring CFSs in specific, and by extrapolation, the whole genome for arrest replication forks. However, although we know a lot about the sensing and downstream signaling that occurs when a replication fork stalls, the question of what pathways are involved in repair of existing damage at CFSs remains.

Repair of damage at CFSs

While the ATR pathway is responsible for the sensing, down stream signaling, and cell cycle stalling after the formation of stalled replication forks, additional pathways must be activated to enact correct repair of lesions at these sites. If left unrepaired, stalled replication forks can lead to both single and, eventually, double strand DNA breaks (Saintigny et al., 2001; Lundin et al., 2002; Ozeri-Galai et al., 2008). There are currently two cellular pathways that are known to be responsible for repairing the majority of double strand DNA breaks. These pathways are the homologous recombination pathway (HR) and the non-homologous end joining pathway (NHEJ) (Jackson, 2002).

The HR pathway repairs double strand DNA breaks by resecting the available DNA ends localized to the break leaving 3'single strand DNA overhangs. These single strand regions then invade homologous regions of the sister chromatid and replicate through the damaged regions in a RAD51 dependent manner. This type of repair is considered to be conservative and error free as there is a sister chromatid present for accurate copying of information across damaged regions. Because of the necessity for one chromatid to invade

and copy from its homologous counterpart, it has been hypothesized that sister chromatid exchanges (SCEs) are a product of sites repaired successfully by the HR pathway (Wilson and Thompson, 2007).

On the other hand, NHEJ mediated repair of DNA damage does not require the presence of a homologous sister chromatid. This pathway repairs a double strand DNA break by religating the broken ends of the DNA together. Repairs made via this pathway are more error prone due to the resection necessary to create the blunt ends used in ligation and because no undamaged template from which to accurately copy the damaged regions is used. The NHEJ pathway requires the function of several different proteins including DNA-PKcs and Ligase IV to enact repair (Critchlow and Jackson, 1998; Lundin et al., 2002).

In 2005, Schwartz et al., demonstrated that under conditions of APH-induced replication stress, siRNA-mediated knock down of RAD51, DNA-PKcs, or Ligase IV caused a significant increase in overall gaps and breaks on metaphase chromosomes and at the CFSs FRA3B and FRA16D (Schwartz et al., 2005). In addition, they found that APH-induced replication stress caused the foci formation of both RAD51 and DNA-PKcs, and that these foci were co-localized with markers of DNA damage such as, γ–H2AX and phosphorylated MDC1. These findings support the assertion that double strand breaks are capable of forming at CFS loci as a result of replication stress and further suggest that repair of CFSs by both the HR and NHEJ pathways is necessary for the stability of these sites. Supporting these findings, it was shown in 1987 that after low dose APH treatment, 70% of all gaps and breaks at the CFS FRA3B also contained an SCE, suggesting a role for HR mediated repair of lesions at CFSs (Glover and Stein, 1987).

CFSs and cancer

Cancer is caused by the accumulation of genes altered in their normal behavior by mechanisms such as DNA damage. CFSs are regions of the genome, frequently associated with genes, which are particularly sensitive to certain forms of DNA damage. Given these facts, the relationship between CFSs and cancer appears unambiguous. In reality, however, the question of whether the instability of CFSs is causal or merely coincidental to cancer is not currently known.

There are three major lines of reasoning suggesting that instability at CFSs may be causal to tumorigenesis. First, CFSs are particularly prone to DNA damage and these sites are commonly found expressed or deleted early in tumorigenesis. Second, CFSs are often co-localized with large genes many of which are known tumor suppressor genes whose loss is frequently found in many different cancers. Third and finally, the regulation of damage at CFSs is primarily associated with the ATR pathway, which involves many genes whose loss or altered expression is strongly associated with tumorigenesis.

While CFSs are stable in untreated somatic cells *in vivo*, these sites are frequently found deleted or rearranged in many different cancer cell lines (Boldog et al., 1994; Wilke et al., 1994; Ohta et al., 1996; Ried et al., 2000; Bednarek et al., 2001; Arlt et al., 2002; Huebner and Croce, 2003; Finnis et al., 2005). In addition, consistent with the finding that CFSs are sensitive to replication stress, deletions at CFS regions also appear very early in tumorigenesis and precede more global genomic deletions and loss of heterozygosity (DiTullio et al., 2002; Gorgoulis et al., 2005). It was recently found that

the same replication stress that causes gaps and breaks on metaphase chromosomes at CFSs also causes large deletions that closely resemble those found in various tumors and cancer cell lines (Durkin et al., 2008). This finding suggests that APH-induced replication stress can cause both the expression of gaps and breaks on metaphase chromosomes and the deletion of large genomic regions that may be an early event in tumorigenesis. The fact that expression of CFSs is a form of DNA damage and DNA damage is an enabling characteristic of tumorigenesis, further supports the idea that the instability of CFSs could be contributory to cancer formation.

Of the CFSs that have been cloned, many are associated with known genes. In the course of discovering and cloning CFSs, several have been localizing using novel sites of HPV integration in cervical cancer cell lines (Ferber et al., 2003; Thorland et al., 2003). These data suggest that regions of high genomic instability, such as are found at CFSs, are particularly prone to viral integrations. In total, over forty of these sites have been localized using such methods and in addition to FRA3B and FRA16D, many of these sites are also associated with genes (Table 1-1). At least two of the genes associated with CFSs have been defined as tumor suppressors, and although currently undefined, many more may also be.

Deletions within FRA3B have been found in solid tumors and have been associated with loss or alteration of the FHIT gene (Glover et al., 1988; Negrini et al., 1996; Arlt et al., 2002). Specifically, alteration or deletion of the FHIT gene has been found in about half of all Barrett's esophagous cases and both gastric and colon carcinomas (Mimori et al., 1999; Fang et al., 2001). Furthermore, in a knock out mouse model, it has been shown that complete loss of FHIT expression causes the strong formation of gastric tumors when mice are treated with the

chemical carcinogen N-nitrosomethylbenzylamine and compared to controls (Dumon et al., 2001). Furthermore, over-expression of FHIT was shown to inhibit gastric tumor formation in nude mice (Siprashvili et al., 1997). Because of these data, it is proposed that FHIT acts as a tumor suppressor gene and that loss of this gene via deletions at the CFS FRA3B could be an initiating factor in gastric and possibly additional tumors.

FRA16D is centered on exons 6, 7, and 8 of WWOX which has been implicated as having a role in both apoptosis and tumor suppression (Bednarek et al., 2001). Deletions of WWOX have been found in a number of gastric adenocarcinomas, multiple myelomas, ovarian, breast, hepatocellular, and prostate carcinomas (Mangelsdorf et al., 2000; Yang and Zhang, 2008). Similar to FHIT, expression of WWOX in breast, ovarian, and lung tumor cell lines normally lacking WWOX expression, resulted in an inhibition of tumor growth *in vivo* (Bednarek et al., 2001; Ludes-Meyers et al., 2003). Additionally, although complete loss of WWOX is lethal, mice with hypomorphic expression of this gene, via a gene trap mouse model, showed an increased incidence of B-cell lymphomas (Ludes-Meyers et al., 2007).

While it is currently unknown how many of the genes associated with CFSs are involved in cancer, there is some indication that more than the two CFSs described above are associated with tumor suppressor genes. The primary example of this is the CFS FRA6E, which spans the PARK2 gene (Denison et al., 2003a; Denison et al., 2003b). While it is not formally established that PARK2 is a tumor suppressor gene, previous work has shown that over-expression of this gene suppressed tumor growth (Denison et al., 2003a; Denison et al., 2003b). Furthermore, PARK2 has been shown to be inactivated in several different tumor types (Cesari et al., 2003; Wang et al., 2004). In addition to PARK2, the gene RORA found

at the CFS FRA15A is also believed to have a tumor suppressive function as it has also been shown to be inactivated in several different tumor types (Zhu et al., 2006).

It has been repeatedly shown that CFSs are frequent sites of large deletions both in tumor cell lines and in treated hybrid cell lines (Glover et al., 1988; Negrini et al., 1996; Mangelsdorf et al., 2000; Arlt et al., 2002; Corbin et al., 2002; Durkin et al., 2008; Yang and Zhang, 2008). These and many similar data suggest that CFSs are frequently unstable in cancers and can lead to the deletion or alteration of the genes associated with them. Lower functional expression or loss of tumor suppressor proteins, such as FHIT or WWOX, may increase susceptibility for development of certain cancers. Supporting the hypothesis that the selective loss of genes associated with CFSs can be causal to tumorigenesis, a recent paper by McAvoy et al., found that there was no relationship between the overall fragility of a given CFS and the frequency that the associated gene is deleted in various cancers (McAvoy et al., 2007). If deletions causing the loss of genes at CFSs were neutral to tumorigenesis, one would expect that the more fragile CFSs would be more often found deleted regardless of the type of cancer examined. Because this is not the case, it appears that there is a selection for the loss of particular CFS and their associated genes in specific cancer types. It is likely that specific cancers, through a process of elimination or out competition similar to evolution on a smaller scale, select for CFS deletions that will most increase their tumorigenic potential. These data may suggest a more overarching role for deletions at CFSs in early tumor formation. Although why there appears to be an association of tumor suppressor genes with CFSs in the first place is currently unknown.

Finally, in addition to CFSs being regions of the genome that are prone to replication stress-induced DNA damage and being regions of the genome that are associated with tumor suppressor genes found to be deleted early in tumorigenesis, CFSs are protected by by a cellular pathway that contains many members well known to have tumor suppressor function. As discussed above, a number of genes involved in the intra-S and G2/M checkpoints have been found to be important in maintaining the stability of CFSs, including *ATR*, *BRCA1*, *CHK1*, *FANCD2*, *HUS1*, and *SMC1* (Casper et al., 2002; Arlt et al., 2004; Howlett et al., 2005; Musio et al., 2005; Durkin et al., 2006; Zhu and Weiss, 2007). More recently, two groups have reported that the ATR and ATM checkpoint pathways are activated very early during tumorigenesis in response to oncogene activity leading to replication stress (DiTullio et al., 2002; Gorgoulis et al., 2005). In response to replication stress ATR phosphorylates CHK1, BRCA1, FANCD2, and possibly P53 as well (Lakin et al., 1999; Liu et al., 2000; Tibbetts et al., 2000; Zhao and Piwnica-Worms, 2001; Smith and La Thangue, 2005). These data suggest that loss of any of these well known tumor suppressors could result in an increase in the instability of CFSs that could additionally contribute to tumorigenesis via an inherent increase in genomic instability or deletion of associated genes.

Ataxia telangiectasia and Rad3 related

Because it allows for a significantly increased rate of mutagenesis, the alteration or loss of genes involved in sensing or repairing DNA damage is considered an "enabling characteristic" of tumorigenesis (Hanahan and Weinberg, 2000). There are a large number of genes that are responsible for sensing and responding to DNA damage.

However, three proteins in particular are central to these roles. These proteins are p53, ATM and ATR. The connection between p53, ATM, and tumorigenesis is well studied and compelling. It has been shown that Li-Fraumeni Syndrome is caused by mutational loss of p53 and is associated with a high rate of a wide variety of cancers (MIM# 191170). In addition, mice lacking the p53 gene develop tumors at a high rate by 6 months of age (Donehower et al., 1992).

Mutations in ATM are the cause of Ataxia Telangiectasia which is also associated with a high rate of tumor formation. Patients with this disease are hypersensitive to UV irradiation (MIM# 607585). In addition, knock out mouse mice lacking Atm have been shown to develop thymic lymphomas by 4 months of age (Barlow et al., 1996). Despite having similar functional significance, the connection between ATR and tumorigenesis has been elusive and difficult to study. The primary reason for this difficulty is that complete loss of ATR is not compatible with life and heterozygous expression results in limited abnormalities.

Given the inherent difficulties in the scientific examination of ATR, almost all of the current knowledge of the function of ATR comes from *in vitro* studies. The ATR kinase is a member of the phosphoinositide 3-kinase-related kinase family of proteins, which includes the ATM kinase. These kinases mediate critical signal transduction roles in the cellular DNA damage response. All three kinases are essential for the maintenance of genomic stability; however, unlike Atm, homozygous mutations in Atr result in early murine embryonic lethality (Brown and Baltimore, 2000).

ATR is indispensable for normal cellular function, and responds primarily to replication stress. The activation of ATR by replication stress early in tumorigenesis

strongly suggests its critical importance in tumor formation (DiTullio et al., 2002; Gorgoulis et al., 2005). It has been shown that ATR is primarily responsible for signaling the intra-S and the G2/M checkpoints in response to replication stress caused by UV irradiation, DNA alkylation, and replication inhibitors such as aphidicolin, hypoxia and folate stress (Arlt et al., 2003; Shechter et al., 2004). Furthermore, it has been shown that ATR phosphorylates a number of proteins important in tumorigenesis as discussed above. It is likely that because of these important roles in normal cellular function that loss of ATR is neither tolerated at the whole organism level nor at the cellular level. This lethality may provide an explanation as to why homozygous inactivating mutations of ATR in tumors are rare. However, heterozygous loss or mutation of ATR is not lethal and has been found in a number of sporadic microsatellite instability-positive stomach, colon, and endometrial tumors (Menoyo et al., 2001; Lewis et al., 2005).

While homozygous mutations in *ATR* are lethal, heterozygous loss or mutation of *ATR* is not lethal. As above, *ATR* mutations have been found in a number of sporadic MMR-deficient stomach, colon, and endometrial tumors (Menoyo et al., 2001; Lewis et al., 2005) and in nasal natural killer T-cell lymphomas (Liu et al., 2006) suggesting that *ATR* functions as a tumor suppressor. Mismatch repair-deficient cancer cell lines containing heterozygous mutations of ATR are hypersensitive to UV, ionizing radiation, cisplatin and topoisomerase inhibitors indicative of a defective response to DNA damage (Lewis et al., 2005). Two groups have studied cancer incidence in heterozygous mice with a knock-out allele. In a small number of animals studied, Fang et al (Fang et al., 2004) did not find a significant increased incidence of early-life tumors in Atr+/- heterozygotes, but did see an ~3 fold increase when on a mismatch repair (MMR) deficient background. Brown and Baltimore

(Brown and Baltimore, 2000) reported a small (3-4 fold) increase in late-life tumors in Atr+/- heterozygotes compared to controls.

Although ATR and ATM are similar in function ATR is early embryonic lethal whereas ATM is not. So what is the essential functional difference between these two genes that makes the loss of one lethal and not the other? ATR responds primarily to RPA coated single strand DNA whereas ATM is activated by unprotected double strand DNA breaks. Single strand RPA coated DNA is found in cells under normal conditions whereas unprotected double strand breaks are an abnormal structure indicative of damage. Therefore, ATR may be present at unperturbed replication forks and regulating cell cycle timing. Alternatively, replication under normal conditions may by itself produce stalled replication forks. These stalled forks could occur as replication proceeds over difficult to replicate regions of the genome, such as CFSs. ATR may be required for the accurate signaling and repair of these sites. This hypothesis is supported by the finding of increased gaps and breaks on untreated cells with ATR knockdown. Whatever the case may be it is clear that ATR is essential to the survival of the cell.

As previously discussed, our laboratory has studied the role of the ATR kinase in the DNA replication stress response using a dominant negative kinase-dead ATR mutant, ATR-specific siRNA, and transient cre-lox mediated ATR depletion (Casper et al., 2002). These experiments demonstrated that transient loss of ATR results in a marked increase in chromosomal instability, particularly at common fragile sites, underscoring ATR's critical role in the maintenance of genomic stability during DNA replication. Current mouse models of Atr mutation are limited in their ability to dissect the biological consequences of Atr deficiency in that a 50% reduction of Atr appears to have only a

partial effect on genome stability and cancer, whereas complete loss of Atr, as would occur with the Knudson two-hit model of tumor suppressor genes, is likely lethal to cancer cells. An appealing alternative approach to this question comes from studies of the human genetic recessive disorder Seckel syndrome.

Seckel syndrome

Seckel syndrome is a genetically heterogeneous group of diseases with three identified subgroups, Seckel Syndrome 1, 2, and 3 (SCKL1, SCKL2, and SCKL3). It is a rare autosomal recessive disorder that is primarily characterized by severe microcephaly, more than five standard deviations below the population mean, and both growth and mental retardation. Clinically, Seckel syndrome presents with a characteristic 'bird-headed' facial appearance. This disease was originally identified in 1960 as bird-headed dwarfism. In 2000, Goodship et al. (Goodship et al., 2000) studied five individuals from two consanguineous Pakistani families. Using a genome screen and homozygosity mapping, they successfully assigned the SCKL1 locus to a region of chromosome 3 (3q22.1-q24) (Goodship et al., 2000). Further detailed examination of candidate genes in this region by O'Driscoll et al. (O'Driscoll et al., 2003) revealed that a mutation in ATR was the cause of SCKL1 in these individuals. Other loci for Seckel syndrome have been mapped to human chromosomes 18p11-q11 (SCKL2) (Borglum et al., 2001) and 14q (SCKL3) (Kilinc et al., 2003) but individual genes responsible for these forms of Seckel have yet to be identified.

SCKL1 is caused by a homozygous point mutation of ATR located in exon 9 (2101A->G). This mutation results in the expression of high levels of an alternate ATR

splice variant excluding sequence from exon 9 introducing a stop codon in exon 10. This stop codon leads to the premature termination of the transcription and loss of expression of the predicted carboxy-truncated ATR protein (O'Driscoll et al., 2003). While this alternate ATR splice variant is expressed at low levels in normal cells, patients with homozygous SCKL1 mutations predominantly express the mis-spliced variant leading to an estimated 90% reduction in normal ATR protein levels.

Heterozygous individuals have an increase in the mis-spliced form of ATR but do not have any described phenotype. The lack of a described phenotype in heterozygous individuals is most likely due to gene dosage effects. In concordance with loss of ATR, our lab has shown that lymphoblasts cultured from individuals with SCKL1 show increased APH-induced breakage (Casper et al., 2004). This was found to be true both for total gaps and breaks per metaphase and also at the specific CFSs FRA3B and FRA16D. Because only two presumably related Pakistani families with a total of five affected young children have been identified with the SCKl1 mutation, it is not yet known if there is an increased risk for tumor formation in humans (Goodship et al., 2000). However, the existing mouse models discussed previously suggest that Atr does have a role in preventing tumorigenesis.

Dissertation summary

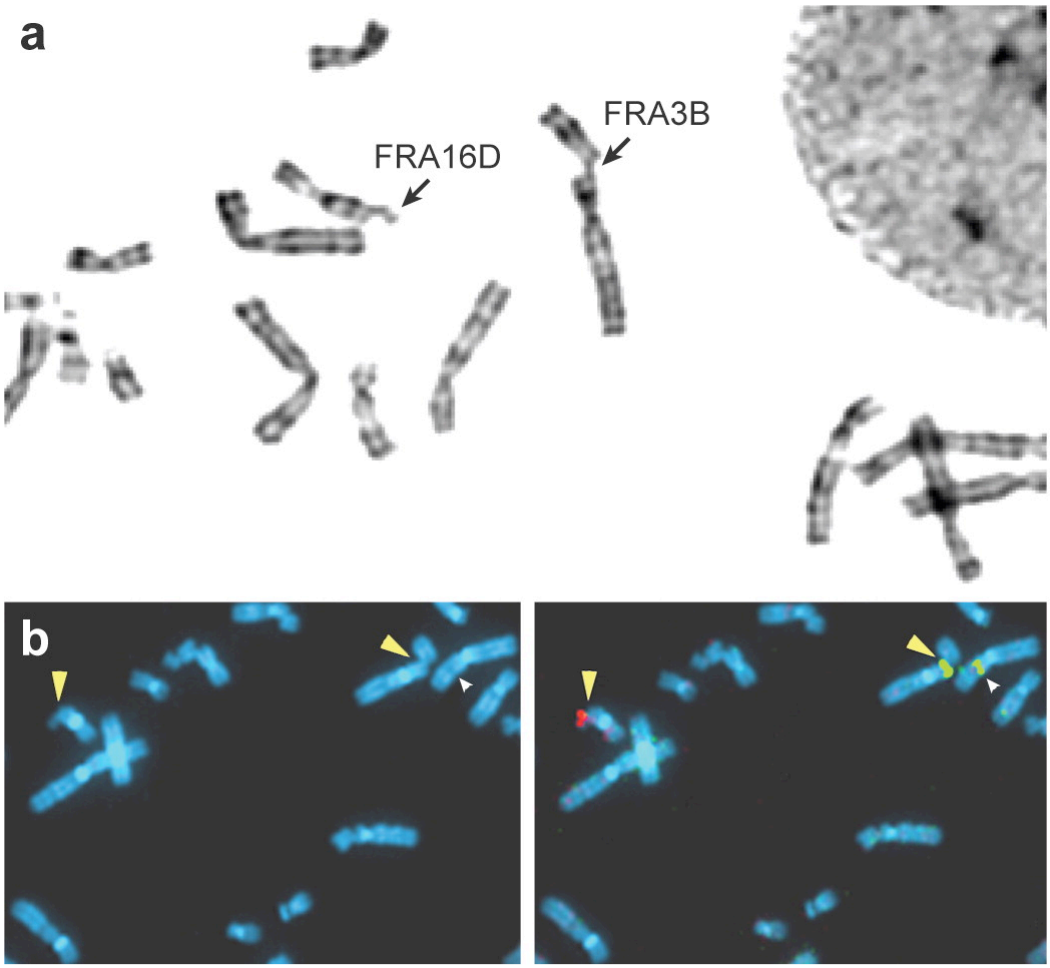
This thesis is focused on two of the major aspects of CFS biology; what makes CFSs "fragile" and how is lack of proper regulation of these sites related to tumorigenesis? Chapter II examines the fundamental genetic aspects that differentiate CFSs from other non-fragile loci in the genome in their response to APH-mediated

replication stress. Specifically, I examined the hypothesis that the increased instability associated with CFSs is due, in part, to genetic sequences located at these loci. This was done by transfecting HCT116 (colon cancer) cells with BACs containing either sequence from the center of the most frequently broken CFS, FRA3B, or control sequence from non-fragile loci and by examining the novel ectopic integrations for their response to low dose APH treatment.

It was found that BACs containing FRA3B sequences retained their instability even when integrated at non-fragile loci and were broken three to seven times as often as the control BAC integrations. This result suggests that the sequence of CFS is sufficient to recapitulate CFS-like instability at novel ectopic loci. Furthermore, it was found that these novel ectopic integrations of FRA3B BACs were not as late replicating as the endogenous FRA3B. This finding suggests that while late replication may be contributory to the instability of the endogenous FRA3B locus, it is not necessary for the instability found at our ectopically integrated FRA3B BACs.

Chapter III is focused on the creation and characterization of a mouse model with hypomorphic expression of the Atr gene. Because of the lethality of complete loss of Atr the examination of the relationship between replication stress, CFS expression, the Atr pathway, and tumorigenesis has not been thoroughly examined. We know that loss of Atr increases the frequency of gaps and breaks at CFSs and that an increase in the expression of CFSs can lead to deletions of associated tumor suppressor genes. However, a definitive study of this relationship has not been performed. Thus, the creation of this novel mouse model will allow us to examine these relationships in more depth than has previously been reported.

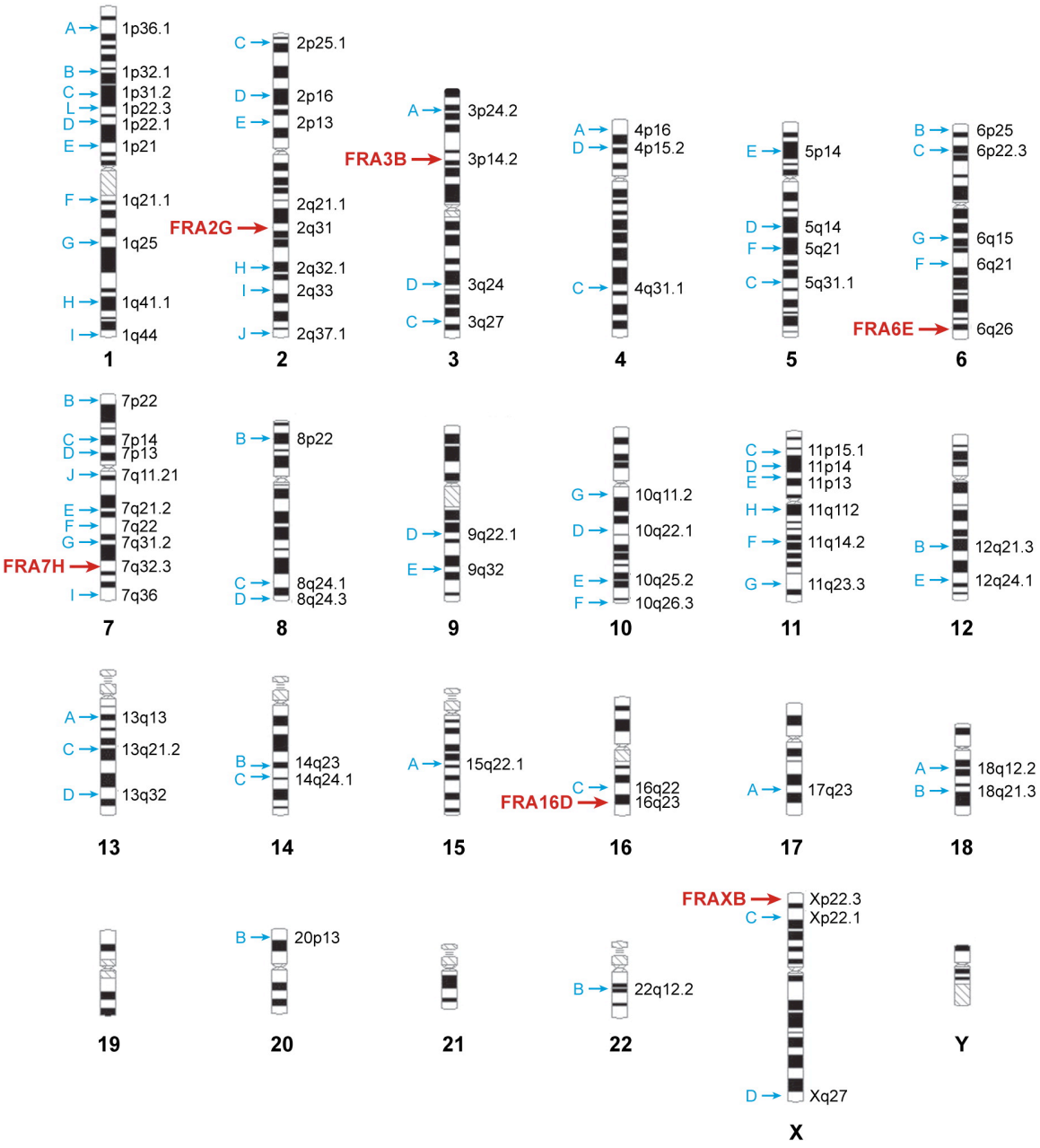
Figure 1-1. Examples of the CFSs FRA3B and FRA16D



AR Durkin SG, Glover TW. 2007.
Annu. Rev. Genet. 41:169–92

Two examples of partial metaphase spreads containing representative G-banded and FISH examples of the CFSs FRA3B and FRA16D. (**A**) A partial metaphase that is G-banded to identify individual chromosomes. Specific fragile sites are as labeled. (**B**) A partial metaphase containing FISH probes for FRA3B (green) and FRA16D (red). The panel on the left does not include the green and red florescence.

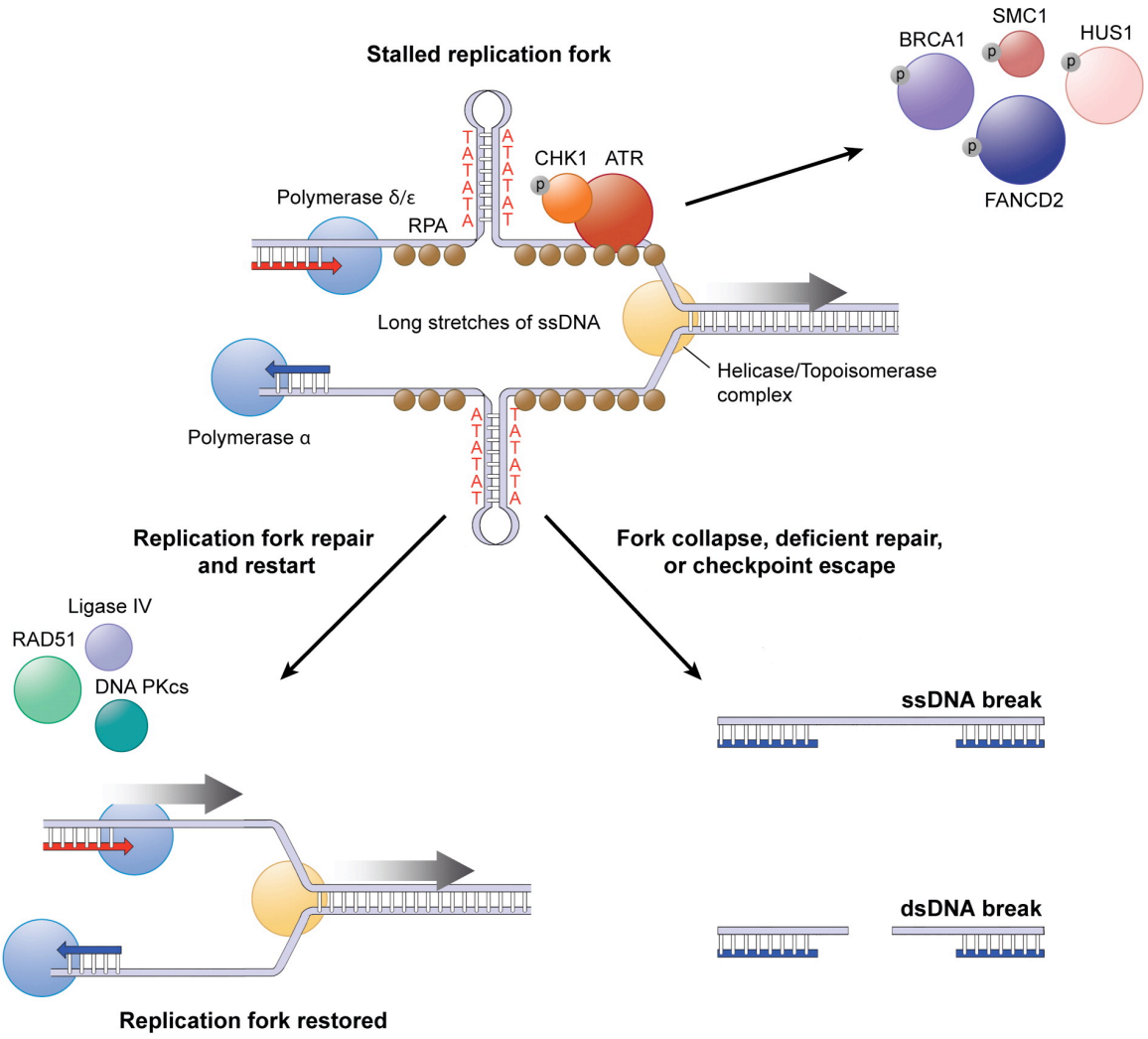
Figure 1-2. Genomic location of all currently identified CFSs.



Durkin SG, Glover TW. 2007.
Annu. Rev. Genet. 41:169–92

Ideogram of human chromosomes indicating the genomic location of known CFSs. Sites in red are the most commonly expressed while sites labeled in blue are less frequently expressed.

Figure 1-3. Model for CFS instability.



Durkin SG, Glover TW. 2007. Annu. Rev. Genet. 41:169–92

Proposed model of CFS instability whereby the helicase and polymerase complexes are separated by APH-induced replication stress. This separation increases the single strand DNA between these complexes thus allowing for the formation of secondary structures and perhaps associated with flexibility peaks or perfect AT repeats. Once formed these structures can recruit RPA to the sites of single strand DNA thus activating the ATR pathway. Correct repair via RAD51, Ligase IV, and/or DNAPKcs leads to replication fork restart whereas incorrect repair or sensing can lead to single and double strand breaks.

Table 1-1. Genes associated with CFSs and their locations

Human CFS	Chromosome Location	Associated Genes
FRA1B	1p32.3	DAB1
FRA2G	2q31	IGRP, RDHL, LRP
FRA2F	2q22.1	LRP1B
FRA3B	3p14.2	FHIT, PTPRG
FRA4B	4q12	PDGFRA/FIPL
FRA4F	4q22.3	GRID2, ARHGAP15
FRA6C	6q22.3	ATXN1/SCA1
FRA6E	6q26	PARK2, MAP3K, LPA
FRA6F	6q21	REV3L, DIF13, FKHRL
FRA7G	7q31.2	CAV1, CAV2, TESTIN, MET
FRA7H	7q32.3	Unidentified gene
FRA7I	7q35	PIP, CNTNAP2
FRA7K	7q31.1	IMMP2L
FRA8C	8q24.1	MYC
FRA9E	9q32-33.1	PAPPA, ROD1 KLF4
FRA10D	10q21.3	CTNNA3
FRA11F	11q14.1	DLG2
FRA13A	13q13.3	NBEA
FRA15A	15q22.2	RORA
FRA16D	16q23.2	WWOX
FRA22B	22q12.3	LARGE
FRAXB	Xp22.3	STS
FRAXC	Xp21.1	DMD, IL1RAPL1

References

Alderton GK, Joenje H, Varon R, Borglum AD, Jeggo PA, O'Driscoll M. 2004. Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway. Hum Mol Genet 13:3127-3138.

Arlt MF, Casper AM, Glover TW. 2003. Common fragile sites. Cytogenet Genome Res 100:92-100.

Arlt MF, Miller DE, Beer DG, Glover TW. 2002. Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells. Genes Chromosomes Cancer 33:82-92.

Arlt MF, Xu B, Durkin SG, Casper AM, Kastan MB, Glover TW. 2004. BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function. Mol Cell Biol 24:6701-6709.

Barlow C, Hirotsune S, Paylor R, Liyanage M, Eckhaus M, Collins F, Shiloh Y, Crawley JN, Ried T, Tagle D, Wynshaw-Boris A. 1996. Atm-deficient mice: a paradigm of ataxia telangiectasia. Cell 86:159-171.

Bednarek AK, Keck-Waggoner CL, Daniel RL, Laflin KJ, Bergsagel PL, Kiguchi K, Brenner AJ, Aldaz CM. 2001. WWOX, the FRA16D gene, behaves as a suppressor of tumor growth. Cancer Res 61:8068-8073.

Boldog F, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H, Drabkin HA. 1997. Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193-203.

Boldog FL, Waggoner B, Glover TW, Chumakov I, Le Paslier D, Cohen D, Gemmill RM, Drabkin HA. 1994. Integrated YAC contig containing the 3p14.2 hereditary renal carcinoma 3;8 translocation breakpoint and the fragile site FRA3B. Genes Chromosomes Cancer 11:216-221.

Borglum AD, Balslev T, Haagerup A, Birkebaek N, Binderup H, Kruse TA, Hertz JM. 2001. A new locus for Seckel syndrome on chromosome 18p11.31-q11.2. Eur J Hum Genet 9:753-757.

Brown EJ, Baltimore D. 2000. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev 14:397-402.

Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA. 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev 19:1040-1052.

Casper AM, Durkin SG, Arlt MF, Glover TW. 2004. Chromosomal instability at common fragile sites in Seckel syndrome. Am J Hum Genet 75:654-660.

Casper AM, Nghiem P, Arlt MF, Glover TW. 2002. ATR regulates fragile site stability. Cell 111:779-789.

Cesari R, Martin ES, Calin GA, Pentimalli F, Bichi R, McAdams H, Trapasso F, Drusco A, Shimizu M, Masciullo V, D'Andrilli G, Scambia G, Picchio MC, Alder H, Godwin AK, Croce CM. 2003. Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27. Proc Natl Acad Sci U S A 100:5956-5961.

Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M. 1997. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell 89:215-225.

Corbin S, Neilly ME, Espinosa R, 3rd, Davis EM, McKeithan TW, Le Beau MM. 2002. Identification of unstable sequences within the common fragile site at 3p14.2: implications for the mechanism of deletions within fragile histidine triad gene/common fragile site at 3p14.2 in tumors. Cancer Res 62:3477-3484.

Critchlow SE, Jackson SP. 1998. DNA end-joining: from yeast to man. Trends Biochem Sci 23:394-398.

Denison SR, Callahan G, Becker NA, Phillips LA, Smith DI. 2003a. Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. Genes Chromosomes Cancer 38:40-52.

Denison SR, Wang F, Becker NA, Schule B, Kock N, Phillips LA, Klein C, Smith DI. 2003b. Alterations in the common fragile site gene Parkin in ovarian and other cancers. Oncogene 22:8370-8378.

DiTullio RA, Jr., Mochan TA, Venere M, Bartkova J, Sehested M, Bartek J, Halazonetis TD. 2002. 53BP1 functions in an ATM-dependent checkpoint pathway that is constitutively activated in human cancer. Nat Cell Biol 4:998-1002.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA, Jr., Butel JS, Bradley A. 1992. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature 356:215-221.

Dumon KR, Ishii H, Fong LY, Zanesi N, Fidanza V, Mancini R, Vecchione A, Baffa R, Trapasso F, During MJ, Huebner K, Croce CM. 2001. FHIT gene therapy

prevents tumor development in Fhit-deficient mice. Proc Natl Acad Sci U S A 98:3346-3351.

Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. 2002. Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 99:16144-16149.

Durkin SG, Arlt MF, Howlett NG, Glover TW. 2006. Depletion of CHK1, but not CHK2, induces chromosomal instability and breaks at common fragile sites. Oncogene 25:4381-4388.

Durkin SG, Glover TW. 2007. Chromosome Fragile Sites. Annu Rev Genet.

Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW. 2008. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. Proc Natl Acad Sci U S A 105:246-251.

Elder FF, Robinson TJ. 1989. Rodent common fragile sites: are they conserved? Evidence from mouse and rat. Chromosoma 97:459-464.

Fang JM, Arlt MF, Burgess AC, Dagenais SL, Beer DG, Glover TW. 2001. Translocation breakpoints in FHIT and FRA3B in both homologs of chromosome 3 in an esophageal adenocarcinoma. Genes Chromosomes Cancer 30:292-298.

Fang Y, Tsao CC, Goodman BK, Furumai R, Tirado CA, Abraham RT, Wang XF. 2004. ATR functions as a gene dosage-dependent tumor suppressor on a mismatch repair-deficient background. Embo J 23:3164-3174.

Ferber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, McGovern R, Gostout BS, Cheung TH, Chung TK, Fu WY, Smith DI. 2003. Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma. Oncogene 22:7233-7242.

Finnis M, Dayan S, Hobson L, Chenevix-Trench G, Friend K, Ried K, Venter D, Woollatt E, Baker E, Richards RI. 2005. Common chromosomal fragile site FRA16D mutation in cancer cells. Hum Mol Genet 14:1341-1349.

Glover TW, Berger C, Coyle J, Echo B. 1984. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:136-142.

Glover TW, Coyle-Morris JF, Li FP, Brown RS, Berger CS, Gemmill RM, Hecht F. 1988. Translocation t(3;8)(p14.2;q24.1) in renal cell carcinoma affects expression of the common fragile site at 3p14(FRA3B) in lymphocytes. Cancer Genet Cytogenet 31:69-73.

Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, Wilke CM, Dierick HA, Beer DG. 1998. The murine Fhit gene is highly similar to its human orthologue and maps to a common fragile site region. Cancer Res 58:3409-3414.

Glover TW, Stein CK. 1987. Induction of sister chromatid exchanges at common fragile sites. Am J Hum Genet 41:882-890.

Glover TW, Stein CK. 1988. Chromosome breakage and recombination at fragile sites. Am J Hum Genet 43:265-273.

Goodship J, Gill H, Carter J, Jackson A, Splitt M, Wright M. 2000. Autozygosity mapping of a seckel syndrome locus to chromosome 3q22. 1-q24. Am J Hum Genet 67:498-503.

Gorgoulis VG, Vassiliou LV, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, Venere M, Ditullio RA, Jr., Kastrinakis NG, Levy B, Kletsas D, Yoneta A, Herlyn M, Kittas C, Halazonetis TD. 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature 434:907-913.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. Cell 100:57-70.

Handt O, Baker E, Dayan S, Gartler SM, Woollatt E, Richards RI, Hansen RS. 2000. Analysis of replication timing at the FRA10B and FRA16B fragile site loci. Chromosome Res 8:677-688.

Hansen RS, Canfield TK, Fjeld AD, Mumm S, Laird CD, Gartler SM. 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 94:4587-4592.

Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD. 1993. Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell 73:1403-1409.

Harvey J, Judge C, Wiener S. 1977. Familial X-linked mental retardation with an X chromosome abnormality. J Med Genet 14:46-50.

Hellman A, Rahat A, Scherer SW, Darvasi A, Tsui LC, Kerem B. 2000. Replication delay along FRA7H, a common fragile site on human chromosome 7, leads to chromosomal instability. Mol Cell Biol 20:4420-4427.

Howlett NG, Taniguchi T, Durkin SG, D'Andrea AD, Glover TW. 2005. The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. Hum Mol Genet 14:693-701.

Huebner K, Croce CM. 2003. Cancer and the FRA3B/FHIT fragile locus: it's a HIT. Br J Cancer 88:1501-1506.

Jackson SP. 2002. Sensing and repairing DNA double-strand breaks. Carcinogenesis 23:687-696.

Kilinc MO, Ninis VN, Ugur SA, Tuysuz B, Seven M, Balci S, Goodship J, Tolun A. 2003. Is the novel SCKL3 at 14q23 the predominant Seckel locus? Eur J Hum Genet 11:851-857.

Kostrub CF, Knudsen K, Subramani S, Enoch T. 1998. Hus1p, a conserved fission yeast checkpoint protein, interacts with Rad1p and is phosphorylated in response to DNA damage. Embo J 17:2055-2066.

Lakin ND, Hann BC, Jackson SP. 1999. The ataxia-telangiectasia related protein ATR mediates DNA-dependent phosphorylation of p53. Oncogene 18:3989-3995.

Le Beau MM, Rassool FV, Neilly ME, Espinosa R, 3rd, Glover TW, Smith DI, McKeithan TW. 1998. Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. Hum Mol Genet 7:755-761.

Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD. 2005. Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. Cell 120:587-598.

Lewis KA, Mullany S, Thomas B, Chien J, Loewen R, Shridhar V, Cliby WA. 2005. Heterozygous ATR mutations in mismatch repair-deficient cancer cells have functional significance. Cancer Res 65:7091-7095.

Liu A, Takakuwa T, Luo WJ, Fujita S, Aozasa K. 2006. Alterations in ATR in nasal NK/T-cell lymphoma and chronic active Epstein-Barr virus infection. Cancer Sci 97:605-610.

Liu Q, Guntuku S, Cui XS, Matsuoka S, Cortez D, Tamai K, Luo G, Carattini-Rivera S, DeMayo F, Bradley A, Donehower LA, Elledge SJ. 2000. Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev 14:1448-1459.

Ludes-Meyers JH, Bednarek AK, Popescu NC, Bedford M, Aldaz CM. 2003. WWOX, the common chromosomal fragile site, FRA16D, cancer gene. Cytogenet Genome Res 100:101-110.

Ludes-Meyers JH, Kil H, Nunez MI, Conti CJ, Parker-Thornburg J, Bedford MT, Aldaz CM. 2007. WWOX hypomorphic mice display a higher incidence of B-cell

lymphomas and develop testicular atrophy. Genes Chromosomes Cancer 46:1129-1136.

Lundin C, Erixon K, Arnaudeau C, Schultz N, Jenssen D, Meuth M, Helleday T. 2002. Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. Mol Cell Biol 22:5869-5878.

Magenis RE, Hecht F, Lovrien EW. 1970. Heritable fragile site on chromosome 16: probable localization of haptoglobin locus in man. Science 170:85-87.

Mangelsdorf M, Ried K, Woollatt E, Dayan S, Eyre H, Finnis M, Hobson L, Nancarrow J, Venter D, Baker E, Richards RI. 2000. Chromosomal fragile site FRA16D and DNA instability in cancer. Cancer Res 60:1683-1689.

McAllister BF, Greenbaum IF. 1997. How common are common fragile sites: variation of aphidicolin-induced chromosomal fragile sites in a population of the deer mouse (Peromyscus maniculatus). Hum Genet 100:182-188.

McAvoy S, Ganapathiraju SC, Ducharme-Smith AL, Pritchett JR, Kosari F, Perez DS, Zhu Y, James CD, Smith DI. 2007. Non-random inactivation of large common fragile site genes in different cancers. Cytogenet Genome Res 118:260-269.

Menoyo A, Alazzouzi H, Espin E, Armengol M, Yamamoto H, Schwartz S, Jr. 2001. Somatic mutations in the DNA damage-response genes ATR and CHK1 in sporadic stomach tumors with microsatellite instability. Cancer Res 61:7727-7730.

Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, Huebner K, Croce CM. 1999. Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. Proc Natl Acad Sci U S A 96:7456-7461.

Mirkin EV, Mirkin SM. 2007. Replication fork stalling at natural impediments. Microbiol Mol Biol Rev 71:13-35.

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B. 1998. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci U S A 95:8141-8146.

Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A, Ried T, Vezzoni P. 2005. SMC1 involvement in fragile site expression. Hum Mol Genet 14:525-533.

Myers JS, Cortez D. 2006. Rapid activation of ATR by ionizing radiation requires ATM and Mre11. J Biol Chem 281:9346-9350.

Nachman MW, Crowell SL. 2000. Estimate of the mutation rate per nucleotide in humans. Genetics 156:297-304.

Negrini M, Monaco C, Vorechovsky I, Ohta M, Druck T, Baffa R, Huebner K, Croce CM. 1996. The FHIT gene at 3p14.2 is abnormal in breast carcinomas. Cancer Res 56:3173-3179.

O'Driscoll M, Ruiz-Perez VL, Woods CG, Jeggo PA, Goodship JA. 2003. A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet 33:497-501.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K. 1996. The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84:587-597.

Ozeri-Galai E, Schwartz M, Rahat A, Kerem B. 2008. Interplay between ATM and ATR in the regulation of common fragile site stability. Oncogene 27:2109-2117.

Pelliccia F, Bosco N, Curatolo A, Rocchi A. 2008. Replication timing of two human common fragile sites: FRA1H and FRA2G. Cytogenet Genome Res 121:196-200.

Rassool FV, McKeithan TW, Neilly ME, van Melle E, Espinosa R, 3rd, Le Beau MM. 1991. Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. Proc Natl Acad Sci U S A 88:6657-6661.

Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S, Nancarrow JK, Woollatt E, Kremmidiotis G, Gardner A, Venter D, Baker E, Richards RI. 2000. Common chromosomal fragile site FRA16D sequence: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells. Hum Mol Genet 9:1651-1663.

Roeder GS, Fink GR. 1980. DNA rearrangements associated with a transposable element in yeast. Cell 21:239-249.

Rozier L, El-Achkar E, Apiou F, Debatisse M. 2004. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. Oncogene 23:6872-6880.

Ruiz-Herrera A, Garcia F, Fronicke L, Ponsa M, Egozcue J, Caldes MG, Stanyon R. 2004. Conservation of aphidicolin-induced fragile sites in Papionini (Primates) species and humans. Chromosome Res 12:683-690.

Saintigny Y, Delacote F, Vares G, Petitot F, Lambert S, Averbeck D, Lopez BS. 2001. Characterization of homologous recombination induced by replication inhibition in mammalian cells. Embo J 20:3861-3870.

Schreiber W. 1979. Jimson seed intoxication: recognition and therapy. Mil Med 144:329-336.

Schwartz M, Zlotorynski E, Goldberg M, Ozeri E, Rahat A, le Sage C, Chen BP, Chen DJ, Agami R, Kerem B. 2005. Homologous recombination and nonhomologous end-joining repair pathways regulate fragile site stability. Genes Dev 19:2715-2726.

Shechter D, Costanzo V, Gautier J. 2004. Regulation of DNA replication by ATR: signaling in response to DNA intermediates. DNA Repair (Amst) 3:901-908.

Siprashvili Z, Sozzi G, Barnes LD, McCue P, Robinson AK, Eryomin V, Sard L, Tagliabue E, Greco A, Fusetti L, Schwartz G, Pierotti MA, Croce CM, Huebner K. 1997. Replacement of Fhit in cancer cells suppresses tumorigenicity. Proc Natl Acad Sci U S A 94:13771-13776.

Smeets DF, van de Klundert FA. 1990. Common fragile sites in man and three closely related primate species. Cytogenet Cell Genet 53:8-14.

Smith L, La Thangue NB. 2005. Signalling DNA damage by regulating p53 co-factor activity. Cell Cycle 4:30-32.

Soulie J, De Grouchy J. 1981. A cytogenetic survey of 110 baboons (Papio cynocephalus). Am J Phys Anthropol 56:107-113.

Stiff T, Walker SA, Cerosaletti K, Goodarzi AA, Petermann E, Concannon P, O'Driscoll M, Jeggo PA. 2006. ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. Embo J 25:5775-5782.

Stone DM, Jacky PB, Hancock DD, Prieur DJ. 1991. Chromosomal fragile site expression in dogs: I. Breed specific differences. Am J Med Genet 40:214-222.

Strunnikov AV, Jessberger R. 1999. Structural maintenance of chromosomes (SMC) proteins: conserved molecular properties for multiple biological functions. Eur J Biochem 263:6-13.

Sutherland GR. 2003. Rare fragile sites. Cytogenet Genome Res 100:77-84.

Sutherland GR, Parslow MI, Baker E. 1985. New classes of common fragile sites induced by 5-azacytidine and bromodeoxyuridine. Hum Genet 69:233-237.

Thorland EC, Myers SL, Gostout BS, Smith DI. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. Oncogene 22:1225-1237.

Tibbetts RS, Cortez D, Brumbaugh KM, Scully R, Livingston D, Elledge SJ, Abraham RT. 2000. Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. Genes Dev 14:2989-3002.

Wang F, Denison S, Lai JP, Philips LA, Montoya D, Kock N, Schule B, Klein C, Shridhar V, Roberts LR, Smith DI. 2004. Parkin gene alterations in hepatocellular carcinoma. Genes Chromosomes Cancer 40:85-96.

Wang L, Darling J, Zhang JS, Huang H, Liu W, Smith DI. 1999. Allele-specific late replication and fragility of the most active common fragile site, FRA3B. Hum Mol Genet 8:431-437.

Wang ND, Testa JR, Smith DI. 1993. Determination of the specificity of aphidicolin-induced breakage of the human 3p14.2 fragile site. Genomics 17:341-347.

Wilke CM, Guo SW, Hall BK, Boldog F, Gemmill RM, Chandrasekharappa SC, Barcroft CL, Drabkin HA, Glover TW. 1994. Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. Genomics 22:319-326.

Wilson DM, 3rd, Thompson LH. 2007. Molecular mechanisms of sister-chromatid exchange. Mutat Res 616:11-23.

Yang J, Zhang W. 2008. WWOX tumor suppressor gene. Histol Histopathol 23:877-882.

Zhang H, Freudenreich CH. 2007. An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in S. cerevisiae. Mol Cell 27:367-379.

Zhao H, Piwnica-Worms H. 2001. ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. Mol Cell Biol 21:4129-4139.

Zhu M, Weiss RS. 2007. Increased common fragile site expression, cell proliferation defects, and apoptosis following conditional inactivation of mouse Hus1 in primary cultured cells. Mol Biol Cell 18:1044-1055.

Zhu Y, McAvoy S, Kuhn R, Smith DI. 2006. RORA, a large common fragile site gene, is involved in cellular stress response. Oncogene 25:2901-2908.

Zlotorynski E, Rahat A, Skaug J, Ben-Porat N, Ozeri E, Hershberg R, Levi A, Scherer SW, Margalit H, Kerem B. 2003. Molecular basis for expression of common and rare fragile sites. Mol Cell Biol 23:7143-7151.

Chapter II

Stably transfected common fragile site sequences exhibit instability at ectopic sites

Abstract

[1]Common fragile sites (CFSs) are loci that are especially prone to forming gaps and breaks on metaphase chromosomes under conditions of replication stress. While much has been learned about the cellular responses to gaps and breaks at CFSs, less is known about what makes these sites inherently unstable. CFS sequences are highly conserved in mammalian evolution and contain a number of sequence motifs that are hypothesized to contribute to their instability. To examine the role of CFS sequences in chromosome breakage, we stably transfected two BACs containing FRA3B sequences and two non-CFS control BACs containing similar sequence content into HCT116 cells and isolated cell clones with BACs integrated at ectopic sites. Integrated BACs were present at just a few to several hundred contiguous copies. Cell clones containing integrated FRA3B BACs showed a significant, three to seven fold increase in aphidicolin-induced gaps and breaks at the integration site as compared to control BACs. Furthermore, many FRA3B integration sites displayed additional chromosome rearrangements associated with CFS instability. Clones were examined for replication

[1] This chapter has previously been published: Ragland RL, Glynn MW, Arlt MF, Glover TW, (2008) Stably transfected common fragile site sequences exhibit instability at ectopic sites. Genes Chromosomes and Cancer. Oct;47(10):860-72.

timing and it was found that the integrated FRA3B sequences were not dependent on late replication for their fragility. This is the first direct evidence in human cells that introduction of CFS sequences into ectopic non-fragile loci is sufficient to recapitulate the instability found at CFSs. These data support the hypothesis that sequences at CFSs are inherently unstable, and are a major factor in the formation of replication stress induced gaps and breaks at CFSs.

Introduction

Common fragile sites (CFSs) are chromosomal loci that are especially prone to forming cytogenetically visible gaps and breaks on metaphase chromosomes under conditions of replication stress such as treatment with low concentrations of the DNA polymerase inhibitor aphidicolin (APH) (reviewed in (Durkin and Glover, 2007)). These sites span several hundred kilobases to over a megabase in size with gaps and breaks occurring throughout the region, and many are located within or span large genes. In addition to cytogenetically identifiable gaps and breaks, CFSs exhibit a number of other characteristics of unstable DNA in cultured cells including gross chromosome deletions (Wang et al., 1993), smaller submicroscopic deletions (Durkin et al., 2008), translocations (Glover and Stein, 1988), intrachromosomal gene amplifications (Coquelle et al., 1997), and sister chromatid exchanges (Glover and Stein, 1987). In cells pretreated with APH, CFSs are also preferred sites for the integration of transfected DNA (Rassool et al., 1991). Despite their instability, CFSs are found in all individuals and are a normal component of the human genome. Furthermore, CFSs are conserved in mammalian evolution(Soulie and De Grouchy, 1981; Elder and Robinson, 1989; Smeets and van de

Klundert, 1990; Stone et al., 1991; McAllister and Greenbaum, 1997; Glover et al., 1998; Ruiz-Herrera et al., 2004) and counterparts exist in yeast (Roeder and Fink, 1980; Dunham et al., 2002; Lemoine et al., 2005).

While CFSs are normally stable in somatic cells in vivo, these sites are frequently rearranged in many cancer cell lines (Arlt et al., 2002; Huebner and Croce, 2003; Finnis et al., 2005). For example, the two most frequently expressed CFSs, FRA3B and FRA16D, are associated with the large tumor suppressor genes *fragile histidine triad (FHIT)* and *WW domain-containing oxidoreductase (WWOX)* respectively (Boldog et al., 1994; Wilke et al., 1994; Ohta et al., 1996; Ried et al., 2000; Bednarek et al., 2001). Heterozygous or homozygous deletion of these tumor suppressors is found early in tumorigenesis and is associated with a number of different cancers (Ohta et al., 1996; Michael et al., 1997; Huebner and Croce, 2003; Bartkova et al., 2005; Gorgoulis et al., 2005). It was recently shown that cells treated with APH form tumor-like deletions at a high frequency at FRA3B supporting the hypothesis that such deletions are a result of replication stress (Durkin et al., 2008).

In recent years, a number of genes involved in the intra-S and G2/M checkpoints have been found to be important in maintaining the stability of CFSs, including *ATR*, *BRCA1*, *CHK1*, *FANCD2*, *HUS1*, and *SMC1* (Casper et al., 2002; Arlt et al., 2004; Howlett et al., 2005; Musio et al., 2005; Durkin et al., 2006; Zhu and Weiss, 2007). In addition, several genes (*RAD51*, *DNA-PKcs*, and *LIG4*) that are involved in both homologous recombination and non-homologous end joining DNA repair pathways have also been reported to be involved in the maintenance of CFSs stability (Schwartz et al., 2005). The importance of the ATR pathway in the maintenance of CFSs implicates

stalled replication forks as having a central role in the instability of these sites. While these studies have provided a basic understanding of the DNA damage response pathways involved in the maintenance of CFS instability, little is understood about what makes CFSs exceptionally susceptible to DNA damage.

There are a number of factors that could contribute to CFS instability. Primary amongst these is sequence. While it is possible that a single sequence motif is responsible for the instability seen at CFSs, none of the cloned CFSs contain expanded di- or trinucleotide repeats like those that are responsible for the instability seen at rare fragile sites (reviewed in (Sutherland, 2003)). In addition, gaps and breaks can occur throughout the large CFS regions, suggesting that if sequence contributes to instability it would most likely result from one or more motifs that are present throughout the entire CFS region. One sequence motif that is in accordance with this hypothesis is the flexibility peak. Flexibility peaks are defined as sequences that have a high local variation in DNA twist angle as measured using the TwistFlex program (Mishmar et al., 1998). Many CFSs contain a high number of flexibility peaks and it has been suggested that the formation of abnormal DNA structures at these sites during replication could be a causal factor in their instability (Boldog et al., 1997; Mishmar et al., 1998; Arlt et al., 2002; Mirkin and Mirkin, 2007). In support of this hypothesis, it has recently been shown that in yeast, an AT-rich sequence containing variable stretches of perfect AT repeats taken from a flexibility peak found in the human FRA16D CFS can cause fork stalling and chromosome breakage (Zhang and Freudenreich, 2007). It has also been suggested that sequence motifs such as high repetitive element content may contribute to the instability of CFSs (Rozier et al., 2004). Finally, because CFSs are some of the latest

sites in the genome to complete replication with some sites remaining unreplicated into the G2 phase of the cell cycle late replication timing may also be causal to the instability of CFSs (Le Beau et al., 1998; Hellman et al., 2000; Palakodeti et al., 2004).

In experiments designed to determine the contribution of sequence to CFS instability, we addressed the question of whether the integration of BAC clones containing sequences from FRA3B are sufficient to recapitulate CFS-like instability at a novel genomic location. We found that cell clones containing either of two adjacent FRA3B BACs integrated at unique non-fragile site loci retained CFS-like instability at the ectopic sites. In addition, we found that the fragility of integrated CFS sequences was not dependent on late replication.

Results

Characterization and retrofitting of CFS and control BACs

Two BAC clones, 431E5 and 530G4, containing genomic inserts within the *FHIT* gene were chosen for these experiments. These two BACs are located in the center of fragility of the most frequently expressed CFS, FRA3B, and overlap by 14.5 kb (Figure 2-1). FRA3B BAC 431E5 is 135.8 kb in length, has 62% AT content, 45% repetitive element content, contains four flexibility peaks, and 11 perfect AT repeats greater than or equal to $(AT)_5$ in size (Table 2-1). The FRA3B BAC 530G4 is 150.7 kb in size, has 62% AT content, 39% repetitive element content, contains eight flexibility peaks, and seven perfect AT repeats greater than or equal to $(AT)_5$ in size.

Two control BACs taken from regions not containing CFSs were chosen to match the sequence content of the FRA3B BACs. Control BAC 250G12 is located at 10q25.1

which does not contain any known genes, is 109.9 kb in size, has 62% AT content, 48% repetitive element content, contains four flexibility peaks, and has 12 perfect AT repeats greater than or equal to $(AT)_5$. Control BAC 412D20 is located in intron 2 of the large gene, FERM domain-containing protein 4A (*FRMD4A*) at 10p13, is 79 kb in size, has 56% AT content, 39% repetitive element content, but contains no flexibility peaks and only two perfect AT repeats greater than or equal to $(AT)_5$ in size. All four BACs were examined for the additional sequence motifs of inverted repeats, palindromes, and tandem repeats all of which are capable of forming strong secondary structures. Both FRA3B and control BACs contain a similar number of tandem repeats and palindromes (Table 2-1). However, control BAC 412D20 contains four times the number of inverted repeats as the experimental BACs, whereas control BAC 250G12 has only half the number of inverted repeats.

All BACs were retrofitted to be resistant to geneticin (G418) using the pRetroES plasmid as described by Kim et al. (Kim et al., 1998) (Figure 2-2). The size, integrity, and sequence content of the genomic inserts for all BACs was confirmed by end sequencing and pulsed field gel electrophoresis (Figure 2-3).

The two FRA3B BACs and two control BACs were stably transfected into HCT116 cells and clones were selected using G418 resistance. Four cell clones containing an integration of the FRA3B BAC 431E5, two cell clones containing the FRA3B BAC 530G4, four cell clones containing the control BAC 412D20, and two cell clones containing the control BAC 250G12 were identified (Figure 2-4). Each clonal cell population contained a single, unique integration site and all integration sites were cytogenetically distinct from any cloned CFSs.

Integrated FRA3B BACs are more fragile than control BACs

To determine if the integrated BACs retained characteristics of CFSs, we scored total gaps and breaks on metaphase chromosomes, gaps and breaks specifically at the sites of integration, and at the endogenous FRA3B loci using FISH with labeled BAC DNA as probe. In the absence of APH, no clone showed a significant number of gaps and breaks at the integration site (p value range 0.118 to 1) (Figure 2-5a) or at the endogenous FRA3B loci (data not shown). Following treatment with 0.4 μM APH, all six cell clones with integrated control BACs showed a low frequency (0-11%) of breakage at the integration site. This level of breakage is not significantly different from untreated cells (p value range 0.49 to 1). However, cell clones containing integrated FRA3B BACs showed a statistically significant three- to nine-fold increase in gaps and breaks at the integration site as compared to untreated clones, and a three- to seven-fold increase in gaps and breaks at the integration site compared to the APH treated control cell clones (p value range 0.023 to < 0.00001) (Figure 2-5b). In addition, cell clones containing integrated FRA3B BACs showed a similar, or slightly higher, level of breakage at the ectopic integration site as compared to the endogenous FRA3B loci following 0.4 μM APH treatment (Figure 2-5c). Because both FRA3B BACs were equally unstable, these results suggest that both BACs contain sequences capable of transferring the instability seen at the FRA3B locus to ectopic sites in the genome.

Integrated FRA3B BACs promote chromosomal rearrangements

In addition to increased gaps and breaks on metaphase chromosomes, eight of the twelve cell clones containing integrated BACs displayed a variety of chromosome rearrangements at the sites of integration (Figure 2-6). Cell clones containing FRA3B BACs had a significantly ($p<0.0001$) higher frequency of these rearrangements as compared to clones containing control BACs, 11.3% (72/638 metaphases) and 4.3% (27/630 metaphases), respectively (Table 2-2). These include dicentric chromosomes, ring chromosomes, laddered amplicons, and massive amplification (Figure 2-6 and Table 2-2). In addition, the FRA3B BAC 530G4 cell clone 24H3 showed a total of 8 metaphases with FISH signals exclusively on fragmented extra-chromosomal elements, and a number of interphase nuclei with micronuclei or chromatin bridge formations (Figure 2-6g,j, and k). The FRA3B BAC 431E5 cell clone 4M contained one metaphase showing a fusion of the sister chromatids at the integration site (Figure 2-6e). The amplifications, laddered FISH signal, sister chromatid fusion, and nuclear chromatin bridges are all consistent with breakage fusion bridge cycle (BFBC) events, which have previously been found to be associated with CFSs (Kuo et al., 1994; Coquelle et al., 1997; Coquelle et al., 2002).

The most striking example of this was found in the FRA3B BAC cell clone 24H3 which demonstrated multiple chromosomal rearrangements indicative of an ongoing repair process occurring at the integration site. In this series of metaphases, a normal signal becomes amplified, most likely through BFBC, and forms multiple breaks that eventually create extra chromosomal elements (Figure 2-6f-i). These extra chromosomal elements are then shuttled into micronuclei for removal from the cell (Figure 2-6j-k).

Three of the six FRA3B BAC cell clones displayed such amplifications with multiple breaks indicating that this process is not limited to the one clone described here.

In total, four of the six cell clones containing FRA3B BAC integrations displayed metaphases with rearrangements at the site of integration, and the majority (47/74 or 63.5%) of these rearrangements were of the types associated with the BFBC. Three of the six control BAC containing cell clones also had metaphases with rearrangements at the integration site. However, compared to the cell clones containing FRA3B BAC integrations, the control cell clones exhibited very few rearrangements of the type associated with BFBC events (3/27 or 11.1%), suggesting that the mechanisms involved in forming the rearrangements in the FRA3B clones are different than those in the control clones.

The BACs integrated in multiple copies

It was apparent from the size and intensity of the FISH signals that all of the cell clones contained multiple copies of the integrated BACs, presumably resulting either from integration of concatamers or amplification following integration. In order to determine the BAC copy number at the integration site and to clarify whether the BACs integrated as whole units or as fragments, we performed quantitative real time PCR (Q-PCR) analysis of genomic DNA from each clone. This analysis was performed using three PCR markers for each clone, one at either end of the genomic insert in the BAC and one in the center.

The copy number of the integrated BACs varied from ~3 copies to over 300 copies (Table 2-3). For most of the cell clones the Q-PCR values for all three markers

were not significantly different indicating that integrated BACs were present as whole units. Although all clones contained sequence corresponding to all three markers, the FRA3B BAC 431E5 cell clone 11M and the control BAC 250G12 cell clone 3 contain different copy numbers of the three markers (Table 2-3). Sequence from marker #1 was overrepresented in both of these clones, indicating that the BACs either did not integrate or did not amplify as a unit. Importantly, all of the FRA3B BAC cell clones have a similar overall copy number as the control BAC cell clones, with the exception of clone 8L, which contained over 300 copies of the FRA3B BAC 431E5.

The frequency of gaps and breaks at integrated FRA3B BACs was not statistically different (p value range 0.49 to 1) between clones with different copy numbers. However, analysis of the BAC copy numbers revealed a trend whereby the fragility of an integrated FRA3B BAC increased with copy number. This trend was seen between the FRA3B BAC 431E5 cell clones 7H, 4M, and 8L and between the FRA3B BAC 530G4 cell clones 24H3 and 1H (Figure 2-5). This suggests that both sequence content and copy number contribute to the fragility of these integrated BACs.

Reduction of BAC copy number reduces but does not eliminate fragility

To address the possibility that the observed instability at the integration sites could be entirely due to the copy number of the integrated BACs and not sequence content, we reduced the copy number of the integrated BACs in five cell clones. To achieve this, we took advantage of the LoxP sites flanking the genomic inserts in the BACs. Cell clones containing multiple copies of the integrated BACs were treated with a retrovirus expressing Cre recombinase to induce deletion of intervening human genomic

sequences. These cells were then sub-cloned and the resulting clones were analyzed using Q-PCR on genomic DNA as before. Using this method, we were able to clone and identify one reduced cell clone derived from both of the control BAC 412D20 cell clone 6H1 and the control BAC 250G12 cell clone 3. Two reduced cell clones derived from the FRA3B BAC 431E5 cell clones 11M and 4M, and one reduced cell clone derived from the FRA3B BAC 530G4 clone 1H were also identified. The BAC copy numbers in the reduced control BAC clones were reduced from 27 and 93 copies to 6 and 2 copies, respectively (Table 2-4). The reduced FRA3B BAC cell clones were reduced from 91 and 168 to 3 and 4 copies respectively, and finally, the reduced clone derived from the FRA3B BAC 530G4 clone 1H was reduced from 106 to 2 copies (Table 2-4). With the exception of PCR marker #3 in the reduced control BAC 412D20 clone 6H1, all three PCR markers used to determine copy number were present in approximately equal amounts in the reduced cell clones suggesting that in the reduced clones the BACs are present as whole units.

We analyzed 75 metaphases from these reduced clones for gaps and breaks at the site of BAC integration by FISH. All data were normalized with respect to total gaps and breaks per metaphase. In the absence of APH, no reduced clone showed any gaps or breaks at the BAC integration site (Figure 2-7a). When treated with 0.4μM APH, all three reduced FRA3B BAC clones showed a somewhat lower frequency of gaps and breaks as compared to the unreduced clones, but still contained a significantly increased frequency of gaps and breaks when compared to the treated reduced control clones (p value range 0.0003 to 0.024) (Figure 2-7b). When treated with 0.4μM APH none of the reduced control cell clones showed any gaps or breaks at the site of integration

suggesting that the small number of breaks seen in the original clones are most likely a result of copy number (Figure 2-7b). In addition, the reduced cell clones containing integrated FRA3B BACs showed a similar level of breakage at the integration site when compared to the endogenous FRA3B loci following 0.4μM APH treatment (Figure 2-7c). These results demonstrate that even at a few copies, the sequence contained in the two FRA3B BACs was able to confer fragile site-like instability to the ectopic site of integration.

Integrated FRA3B BACs replicate earlier than endogenous FRA3B loci

Late replication has been demonstrated for a number of CFSs and is believed to be a causal factor in their fragility (Le Beau et al., 1998; Wang et al., 1999; Focarelli et al., 2006). In order to determine if our integrated FRA3B BACs replicate at the same time as the endogenous FRA3B loci we examined by FISH analysis 108 interphase cells from all three cell clones containing reduced FRA3B BAC insertions. The endogenous FRA3B site was found to have similar replication timing in all three of the reduced clones with doublet signals, indicative of completed replication, observed at 27-35% of the loci scored (Figure 2-8). This frequency of replicated loci is consistent with the previously reported late replication timing of FRA3B (Le Beau et al., 1998). In contrast, the integrated FRA3B BAC sequences showed a significantly higher ($p<0.0001$) percentage of doublet signals across all three cell clones (Figure 2-8). The reduced FRA3B BAC 431E5 cell clones 4M and 11M showed doublet signals at 55% and 59% of their integration sites, respectively. The reduced BAC 530G4 cell clone 1H had doublet signals at 74% of its integration sites, approximately a 2-3 fold increase in the number of

signals that are replicated over the endogenous FRA3B. These data show that the integrated BAC sequences are completing replication much earlier than the endogenous FRA3B loci, yet still show fragile site-like instability.

Discussion

We have shown that in human cells sequences from two adjacent FRA3B BACs can transfer CFS-like instability to new loci in the genome. These data show that sequences taken from different regions of the FRA3B CFS are inherently unstable and that these sequences alone are sufficient to recapitulate much of the instability seen at CFSs. Because all of the FRA3B BAC integration sites were at unique chromosomal loci and retained similar levels of CFS-like instability, the genomic location of the BAC integrations does not appear to be a significant factor in the these experiments.

Although not statistically significant, cell clones containing integrated FRA3B BACs followed a trend whereby the greater the BAC copy number the greater the frequency of metaphases with gaps and breaks at the sites of integration. This was seen when comparing the fragility of the original FRA3B BAC cell clones that contained multiple BAC copies to each other (Figure 2-5b), when original and reduced cell clones were compared to each other (Figure 2-7b), or when the any of the FRA3B BAC cell clones was compared to the endogenous FRA3B loci in the same cell clone (Figures 2-5c and 2-5c).

Gaps and breaks at CFSs have been shown to lead to the BFBC and intrachromosomal gene amplification events in Chinese hamster ovary cells and in cancer cells (Kuo et al., 1994; Coquelle et al., 1997; Coquelle et al., 2002). We found that the

transfer of fragile site sequence is sufficient to recapitulate some of the amplification events that have been seen at endogenous CFS loci in other studies. Clones containing ectopic FRA3B BAC integrations displayed a statistically greater overall number of chromosomal rearrangements and a greater number of rearrangements indicative of BFBC than did the control BAC integrations. These chromosomal rearrangements were not found in the any of the cell clones that were reduced in copy number, suggesting that the copy number of BAC integrations is important in the formation of these rearrangements. However, the statistically significant difference in the number and type of rearrangements between FRA3B and control BAC integrations indicates that sequence content of the FRA3B BACs also contributes to these rearrangements.

The search for a discrete sequence motif that is responsible for CFS fragility has historically been complicated by the large size of the CFS regions. Although we reduced the amount of sequence that we needed to interrogate, from ~1Mb to just over 100 kb, we were not able to precisely identify a sequence motif responsible for CFS fragility. In addition to the AT-rich flexibility peaks (Mishmar et al., 1998), recent work in a yeast model system has shown that short $(AT)_{5-34}$ perfect repeats within CFSs are hot spots for chromosome breakage and that the longer of these repeats $(AT)_{23-34}$ lead to stalled replication in 2D-gel analysis (Zhang and Freudenreich, 2007). It has also been hypothesized that other sequence motifs capable of forming a strong secondary structures or stalling replication could be causal in the formation of gaps and breaks at CFSs (Durkin and Glover, 2007; Mirkin and Mirkin, 2007; Zhang and Freudenreich, 2007). Because both control and FRA3B BACs contain similar AT content, repetitive element content, tandem repeats, and palindrome motifs it is unlikely that the presence of these

features alone are responsible for the differences in instability observed at the ectopic integrations. In addition, inverted repeats are present in control BAC 412D20 at a frequency that is approximately four times that of the FRA3B BACs suggesting that this motif is not sufficient to recapitulate CFS-like instability at an ectopic location. Both AT perfect repeats greater than or equal to $(AT)_5$ and flexibility peaks are present at the same frequency in the control BAC 250G12 as in both of the FRA3B BACs suggesting that presence of these motifs alone is also not responsible for the fragility of these ectopic FRA3B BAC integrations. However, it should be noted that although the frequency of greater than five perfect AT repeats was similar, the longest perfect AT repeat found in the control BACs was an $(AT)_{19}$ repeat, whereas the longest motif in the FRA3B BACs was an $(AT)_{24}$ repeat found in the overlapping region of the two FRA3B BACs. Given that it has been found that AT perfect repeats $(AT)_{21\text{-}34}$ units in size were found to stall replication forks better (Zhang and Freudenreich, 2007) and to form stronger cruciform structures in yeast (Dayn et al., 1991) than do sequences with a lower number of repeats, there may be some threshold effect whereby an $(AT)_{19}$ repeat is not sufficient to contribute to fragility whereas an $(AT)_{24}$ repeat is. An examination of the human genome reveals that there are thousands of perfect $(AT)_{24}$ repeats, many of which are not associated with CFSs. Two of these sites, located at 9q21 and 15q25, were examined using the TwistFlex program. Like the $(AT)_{24}$ repeat found in our FRA3B BACs, these repeats were found to be located in flexibility peaks and surrounded by regions of high AT content (>60%). Because these repeats were found in non-fragile regions of the genome, and were indistinguishable from the $(AT)_{24}$ repeat found in our FRA3B BACs in size or sequence context, it is suggested that although a single $(AT)_{24}$ may be

contributory to the fragility of CFSs, the contribution of sequence to the instability of our BAC integrations, and to CFSs in general, is likely to be more complex.

Although a number of papers have reported that molecularly characterized CFS are highly flexible and AT rich (Boldog et al., 1997; Mishmar et al., 1998; Ried et al., 2000; Shiraishi et al., 2001; Arlt et al., 2002; Limongi et al., 2003; Ferber et al., 2004), a few studies differ in their findings. Helmrich et al. examined sequences from 15 human and 8 mouse CFSs and found no increase in DNA flexibility as compared to controls (Helmrich et al., 2006). Similarly, Tsantoulis et al. describes CFSs as being flexibility peak poor and GC rich (Tsantoulis et al., 2007). Both of these studies included large regions of sequence flanking the CFSs in their analyses, as opposed to a few hundred kilobases of most fragile DNA at the center of molecularly characterized CFSs as in most other studies of CFS sequence. In addition, Tsantoulis et al. included CFSs that have only been mapped at the resolution of a chromosome band in their analysis. These disparities likely account for differences in findings and interpretations.

A few studies have explored the contribution of sequence to CFS instability by examining chromosome breakage at endogenous CFSs that contain large deletions. Hamster human chromosome 3-hybrid cell clones containing large deletions of FRA3B derived during tumorigenesis were found to have no significant reduction in fragility (Corbin et al., 2002). More recently, a significant reduction in the fragility of FRA3B was reported in seven clones with large (several hundred kilobases) APH-induced deletions centered within the fragile site (Durkin et al., 2008). Many of these deletions encompass the sequences contained in our FRA3B BACs. Finally, a tumor cell line containing a deletion of the entire FRAXB locus was found to have completely lost

fragility of FRAXB (Arlt et al., 2002). These findings suggest that many sequence motifs spread throughout the region are likely to contribute to the fragility of CFSs and are in agreement with our results that two adjoining regions of FRA3B were capable of recapitulating CFS-like instability at ectopic sites. Furthermore, these findings imply that whatever features are necessary for CFS fragility are located within the defined CFS sequence boundaries and cannot be eliminated without a complete deletion of the site.

In order to examine other factors that could contribute to CFS fragility, we analyzed the integrated BAC in our reduced clones for replication timing. It was found that ectopic FRA3B BAC integrations were 2-3 times more likely than the endogenous FRA3B sites to be replicated. Despite earlier replication timing, these sites were still unstable and formed gaps and breaks under conditions of replication stress. These findings suggest that ectopic FRA3B BAC integrations do not need to be as late replicating as the endogenous loci to be unstable. Because these findings were determined using a FISH assay with the entire FRA3B BACs as probes, we cannot exclude the possibility that smaller regions within the integrated BACs are late replicating. However these data suggest that late replication is not necessary for the formation of gaps and breaks at our FRA3B ectopic integrations.

Our findings show that sequence alone is a critical factor underlying the instability seen at CFSs, and may help to elucidate further mechanisms involved in CFS instability. While it has long been hypothesized that specific sequences are causal to the fragility of CFSs, this is the first example showing that CFS instability can be transferred to ectopic sites in mammalian cells by more than one region and that this instability can be seen with as little as 300 kb of transferred sequence. These findings are central to

understanding the fundamental mechanisms underlying CFS instability, the role of CFSs in the normal functioning of human cells, and in early breakage events in tumorigenesis.

Materials and methods

Identification and retrofitting of BACs

The Research Genetics (Huntsville, AL.) human genomic BAC library was screened using PCR probes for BAC clones containing sequence from intron 4, exon 5 and intron 5 of *FHIT*. PCR primers were designed using sequences obtained from the UCSC Genome Browser. E. coli containing the BACs were inoculated into 1L LB media containing 25μg/ml chloramphenicol and incubated overnight at 37°C. The pRetroES plasmid was used for retrofitting the BACs to contain neomycin resistance gene following the procedures outlined by Wang et al. ((Wang et al., 2001). Lysis and extraction of the BAC was done following standard CsCl extraction protocols (Wilson, 2001). PFGE analysis was performed on extracted BACs digested with *NotI* in order to determine if correct retrofitting had occurred. In order to determine the exact sequence content of the genomic inserts end sequencing was preformed using T7 and SP6 primers flanking the genomic insert in the BAC.

Sequence analysis

Flexibility peaks ware defined and determined using the TwistFlex program (http://margalit.huji.ac.il/TwistFlex/Home.html), which analyzes DNA flexibility at the twist angle. Repetitive element content (SINEs, LINEs, LTR elements, and DNA elements) was measured and defined using the repeat masker program

(http://www.repeatmasker.org). Additionally the online programs Palindrome (http://bioweb.pasteur.fr/seqanal/interfaces/palindrome.html), Inverted Repeat Finder (Warburton et al., 2004), and Tandem Repeat Finder (Benson, 1999) were used for examination of BAC sequence. Default settings for all values for all programs were used.

Transfection

HCT116 cells were transfected using the DOTAP liposomal transfection reagent (Roche Applied Science). A 10 cm^2 cell culture plate was seeded with $1.5X10^6$ cells and incubated overnight at 37°C. Ten micrograms of supercoiled BAC DNA was subsequently transfected into these cells following the manufacturer's protocols. The medium was replaced after 10 hours of incubation with selective medium containing 500 μg/ml of active G418 (Gibco). Stable, G418 resistant clones were cloned and analyzed for integration of the BAC by real time PCR and fluorescent in situ hybridization.

Cell culture and CFS analysis

HCT116 cells were grown in Dulbecco's modified essential medium (DMEM) containing 10% fetal bovine serum (FBS), L-glutamine, and penicillin/streptomycin at 37°C and 5% CO_2. Cells resistant for G418 were grown as above with medium supplemented with 500 μg/ml G418.

Cells were treated with 0.4 μM aphidicolin for 24 hours prior to metaphase chromosome harvest to induce fragile site expression. Metaphases were harvested by treating the cells with 50 ng/ml colcemid for 45 minutes. Cells were then placed in 0.075 M KCl at 37°C for 18 minutes and fixed in Carnoy fixative (3:1 methanol:glacial acetic

acid) for two hours before replacing the fixative several times. Fixed cells were dropped on slides and aged at room temperature for 48 hours prior to FISH analysis.

Fifty metaphases from each cell clone were scored for total chromosome gaps and breaks as well as gaps and breaks at the sites of FISH probe hybridization. In order to account for differences in treatment from one sample to another, gaps and breaks at the sites of FISH probe hybridization were normalized using total gaps and breaks per cell.

FISH

The BACs used for transfection were used for FISH analysis of the clones containing those BACs. Established FISH protocols were followed (Wilke et al., 1996). Probes were labeled with biotin 14-dATP using the BioNic Translation kit (Invitrogen, Carlsbad, Ca.). Bound probe was detected using fluorescein isothiocynate (FITC) conjugated avidin-DCS (Vector Laboratories, Burlingame, Ca.) followed by fluorescein conjugated anti-avidin immunoglobulin G (IgG). Chromosomes were stained using 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Burlingame, Ca.). Fluorescent signals were detected using a Zeiss Axioscope epifluorescence microscope.

Reduction of copy number

To reduce the copy number of the BACs at the sites of integration, cells were treated with a retrovirus, AdCre1, expressing Cre recombinase (Dr. Frank Graham McMaster University, Hamilton, Ontario, Canada). One well of a six well plate was seeded with $8X10^4$ cells. Cells were allowed to settle overnight and then treated with the retrovirus. Cells were allowed to recover for 48 hours and treated with retrovirus again

as above. After 48 hours of recovery cells were plated out at ~200 cells/10cm^2 for isolation of cell clones.

Quantitative real time PCR

Copy number of transfected BACs was estimated by genomic real time PCR using the SYBR green assay and the iCycler system (BIO RAD). Genomic DNA was extracted by standard alkaline lysis and 50 ng of total genomic DNA was used per PCR reaction. Primers were designed using primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). All primer pairs were optimized using a standard curve of 200 ng, 20 ng, 2 ng, and 0 ng samples of normal HCT116 and/or LD055 genomic DNA. Experimental samples were run in triplicate over two separate experiments for a total of six replicates for each primer pair and eighteen replicates for each clone. A GAPDH primer pair was run for each sample in triplicate over two experiments to determine total DNA input for each sample. All results were analyzed for copy number using the $2^{-\Delta\Delta ct}$ method (Livak and Schmittgen, 2001).

Table 2-1 Location and sequence content of the genomic inserts of the BACs

BAC	431E5	530G4	412D20	250G12
Type	FRA3B	FRA3B	Control	Control
Chr.	3p14.2	3p14.2	10p13	10q25
Size (kb)	135.8	150.7	79	109.9
AT(%)	61.99	62.10	56.07	62.23
Repeat Ele.[a] (%)	44.83	38.96	39.36	48.16
SINE(%)	9.88	9.84	19.57	6.84
LINE(%)	24.42	11.37	7.90	22.95
Flex Peaks (/100kb)	2.95	5.31	0.00	3.64
Inverted Repeats (/100kb)	102.36	104.18	432.91	42.77
Palin.[b] (/100kb)	117.08	69.01	67.09	131.94
Tandem Repeats (/100kb)	15.46	21.90	50.63	26.39
$(AT)_{>5}$ (/100kb)	8.10	4.64	2.53	10.92

[a]Percent of sequence that is repetitive elements
[b]Number of palindromes / 100kb of sequence

Table 2-2 Type and number of abnormal signals seen in clones at the integration site

BAC	Clone	Dicentric Chr.[a]		Ring Chr.[b]		Ring Chr Multi.[c]		Multi. Chr.[d]		Large Amp.[e]		Chr. Ele.[f]		Fused Chr.[g]		Total	
		APH		APH		APH		APH		APH		APH		APH		APH	
		-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Control	2H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
412D20	6H1	2	1	5	4	0	1	0	1	0	0	0	0	0	0	7	7
	4H	2	6	0	0	0	0	0	0	1	1	0	0	0	0	3	7
	3H5	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	2
Control	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
250G12	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRA3B	7H	3	1	0	0	0	0	0	1	2	11	0	0	0	0	5	13
431E5	11M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4M	2	3	0	0	0	0	1	0	0	6	0	0	0	1	3	10
	8L	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	2
FRA3B	24H3	4	7	0	0	0	3	3	3	3	6	4	4	0	0	14	23
530G4	1H	2	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2
Total		16	22	5	4	0	4	4	6	6	25	4	4	0	1	35	66

Number of metaphases containing: [a]FISH signal on a dicentric chromosome.
[b] FISH signal on a ring chromosome.
[c]More than one FISH signal on a ring chromosome.
[d]More than one FISH signal on a single chromosome.
[e] Very large amplifications of FISH signal on a single chromosome.
[f] Extra chromosomal elements marked by FISH signal.
[g] FISH signal at fused sister chromatids.

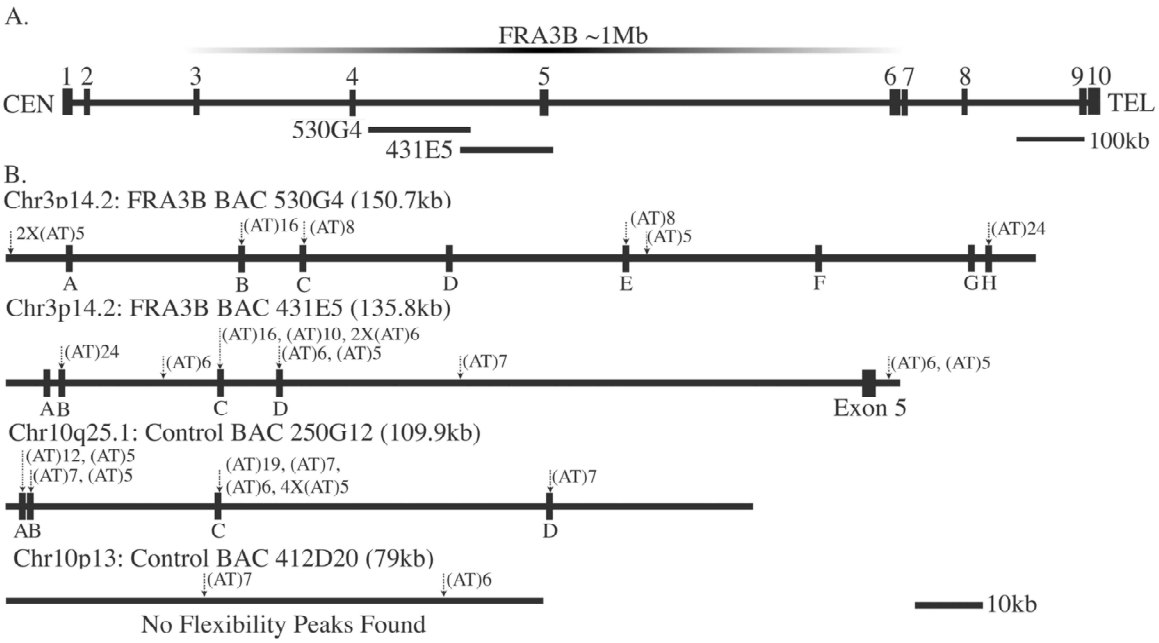
Table 2-3 Copy number analysis for all PCR markers across all BAC clones

BAC	Clone	PCR 1	PCR 2	PCR 3	Avg
Control	2H	12 ± 3	22 ± 6	10± 4	14
412D20	6H1	47±14	14±2	18±4	27
	4H	169±13	120±11	210±24	166
	3H5	188±39	219±56	204±37	204
Control	2	4±2	3±1	2±0.5	3
250G12	3	146±29	90±11	43±4	93
FRA3B	7H	45±11	26±6	45±9	38
431E5	11M	138±38	80±14	54±6	91
	4M	245±107	131±85	128±42	168
	8L	421±240	216±66	312±92	316
FRA3B	24H3	73± 3	44±18	65± 21	61
530G4	1H	90±32	141±40	87±29	106

Table 2-4 Copy number for all PCR markers in the reduced clones

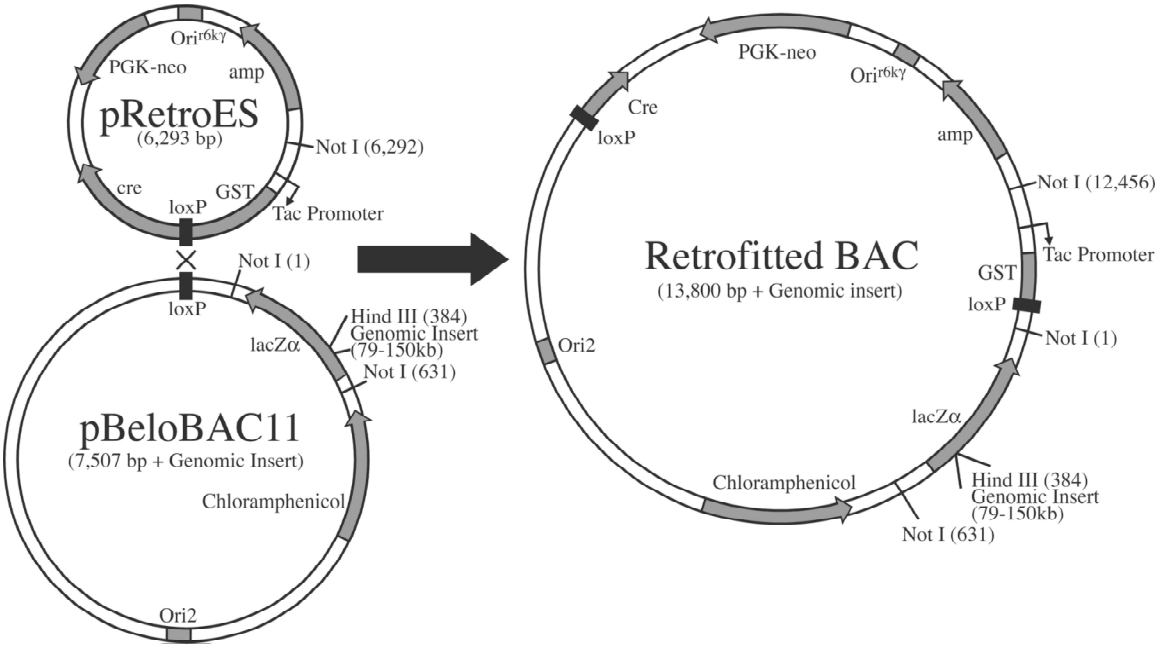
BAC	Clone	PCR 1	PCR 2	PCR 3	Avg
Control 412D20	6H1	3± 2	3±1	14±4	6
Control 250G12	3	1±0.4	2± 1	2±1	2
FRA3B 431E5	11M	3±1	2±0.3	4±0.4	3
	4M	6±3	3±0.9	3± 0.7	4
FRA3B 530G4	1H	2±0.3	2±0.4	2±0.5	2

Figure 2-1 Genomic location, flexibility peak, and (AT) repeat content of BACs



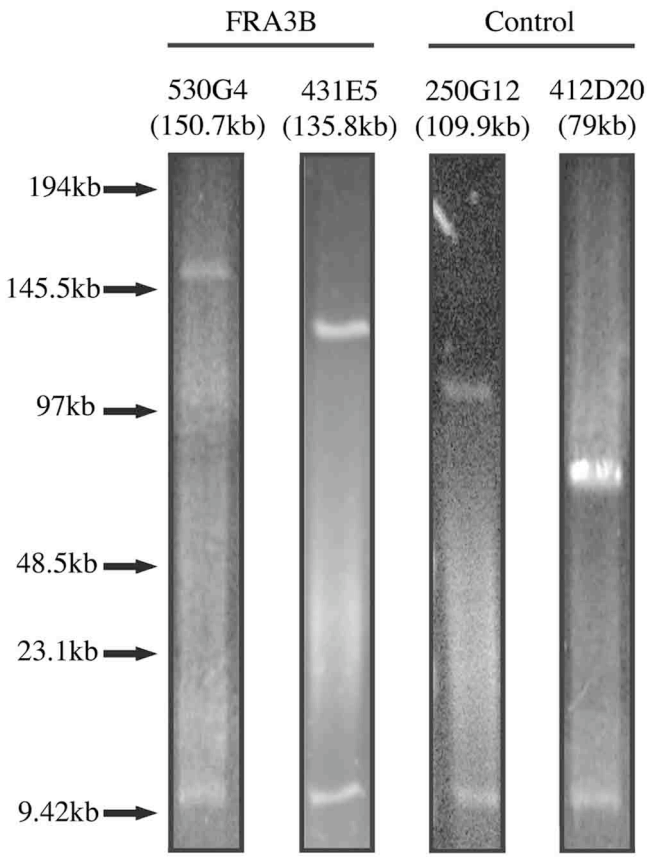
(**A**) Physical map of the FHIT gene and FRA3B locus. The position and size of the two FRA3B BACs (431E5 and 530G4) is indicated. (**B**) Physical map of all four BAC genomic inserts. The size and genomic location of the inserts is indicated. The number and location of flexibility peaks are indicated as letters for each BAC. The location and size of perfect AT repeated greater than $(AT)_4$ is also indicated.

Figure 2-2 Method of BAC retrofitting to contain NEO resistance



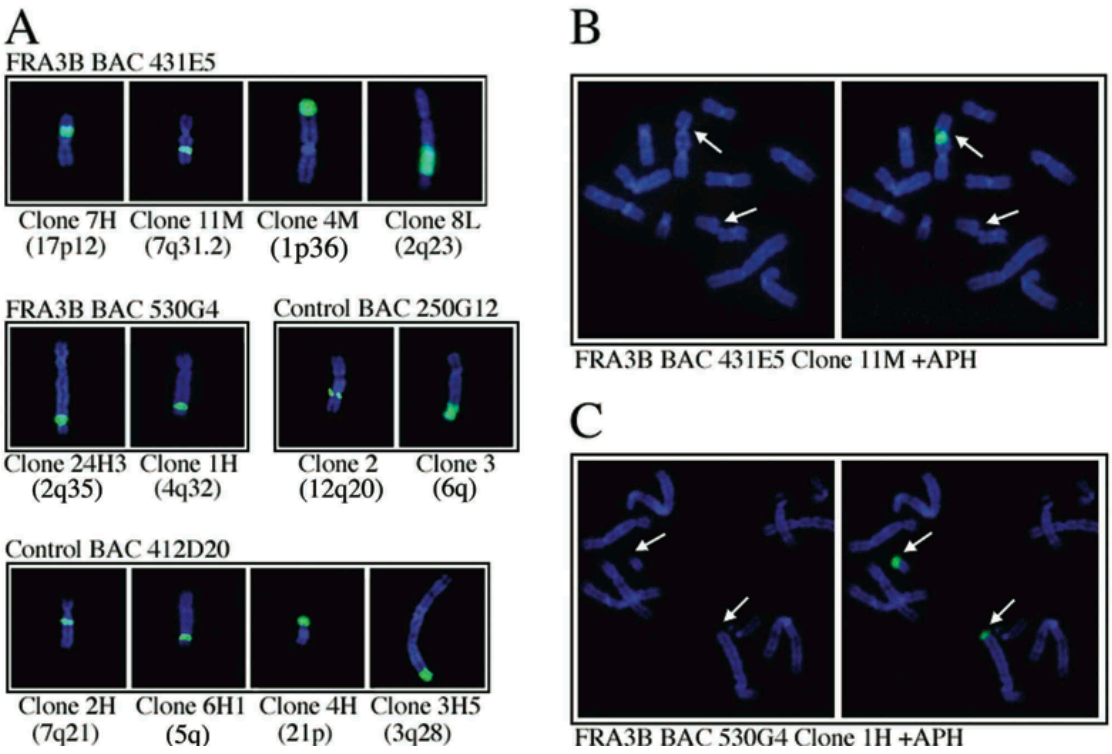
Map of the retrofitting vector pRetroES, the BAC backbone pBeloBAC11, and the resulting retrofitted construct. The GST-loxP-cre fusion gene drives recombination at the loxP sites and is subsequently separated from the tac promoter post recombination, stopping Cre expression. Replication of the pRetroES plasmid without recombination is prevented due to the ori$^{R6k\gamma}$, which will not function in most bacterial hosts. Homologous recombination is selected for by treating the bacterial hosts with both Ampicillin and chlorenphenicol.

Figure 2-3 Pulsed field gel electrophorisis analysis of BAC content



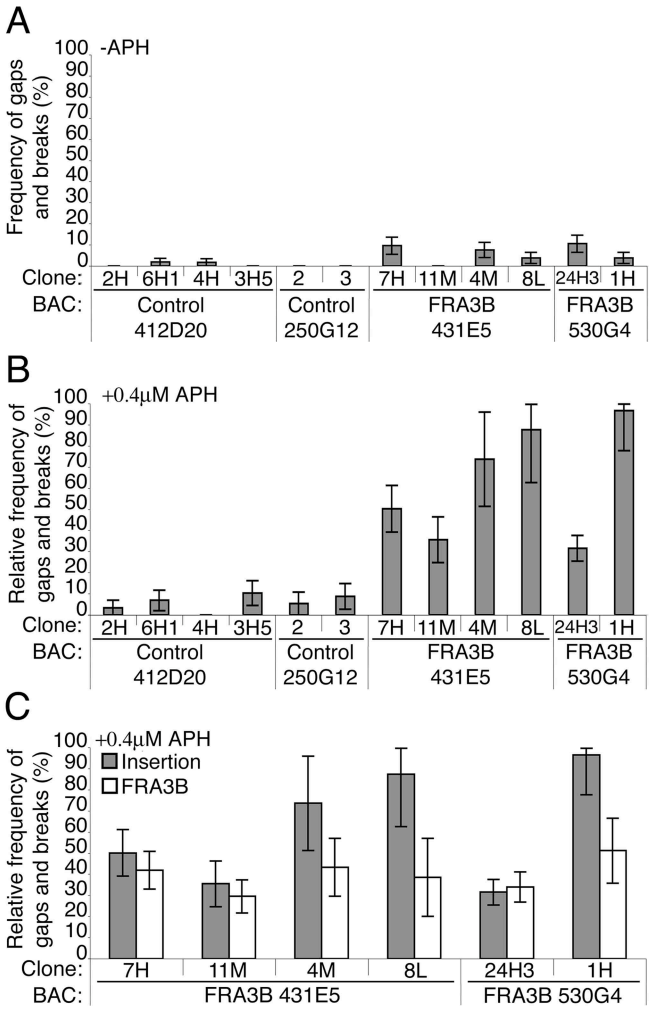
PFGE results for two FRA3B and two control BACs extracted by CsCl and digested overnight with the NotI restriction enzyme. The larger band in each land corresponds to the variable genomic insertion in each BAC, expected sizes are indicated bellow each BAC label. The smaller band corresponds to the expected size of the retrofitted vector backbone (11,825bp).

Figure 2-4 FISH signals of all cell clones containing BAC insertions



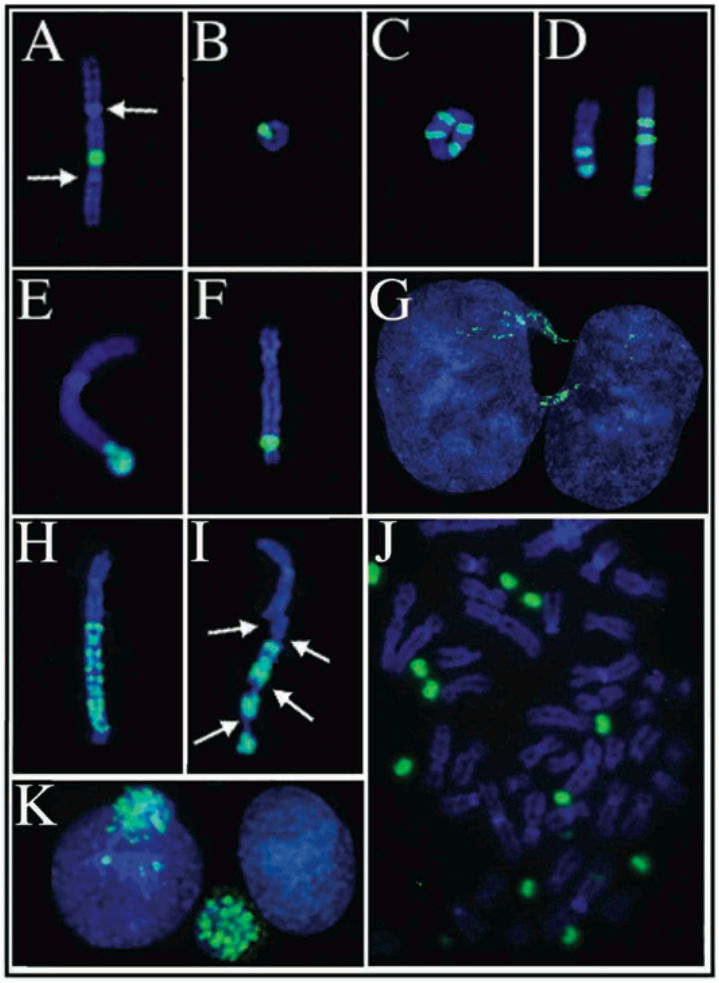
(**A**) Representative FISH signals showing the relative size and genomic location of the BAC insertions. (**B** and **C**) Two partial metaphase spreads demonstrating both a typical CFS-like break and a complete chromosome break respectively, as indicated by the arrows, at the sites of integration.

Figure 2-5 Frequency of gaps and breaks on metaphase chromosomes



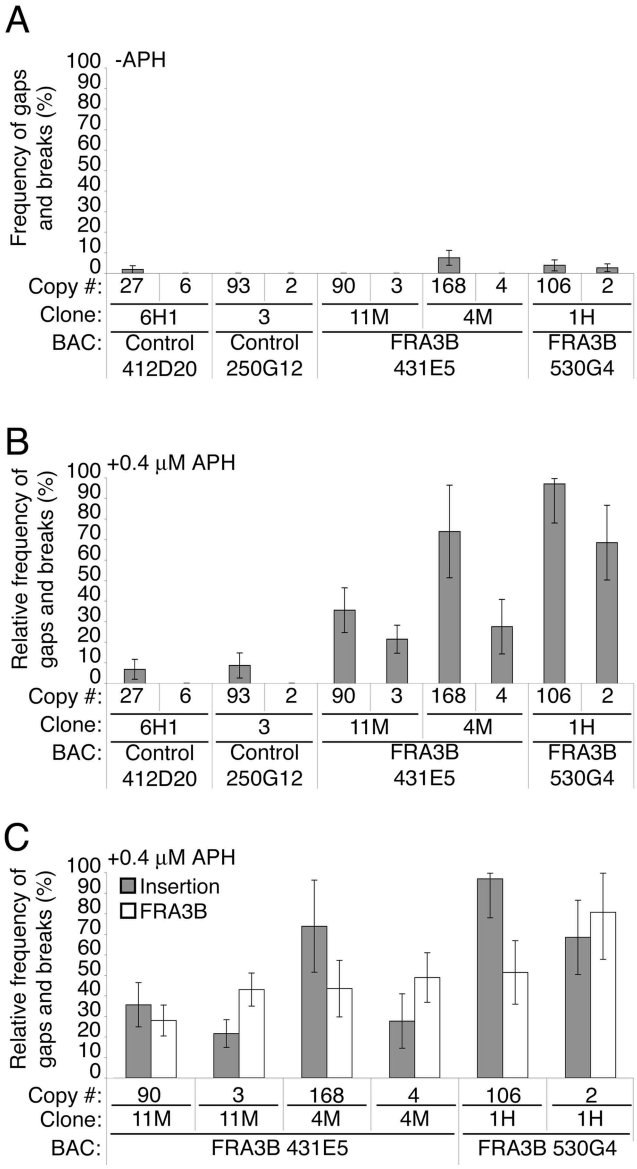
Cells containing FRA3B BAC insertions show elevated gaps and breaks at the integration site when treated with 0.4μm APH. All APH treated samples were normalized using total gaps and breaks as a measure. (**A**) Percent of integration sites with gaps or breaks in all untreated clones as seen in 50 metaphases for each clone. (**B**) The relative percentage of gaps and breaks at the site of integration after treatment with 0.4μm APH. (**C**) Relative percentage of gaps and breaks at both the integration site and the endogenous FRA3B locus for each clone.

Figure 2-6 Abnormal FISH signals associated with BAC integrations



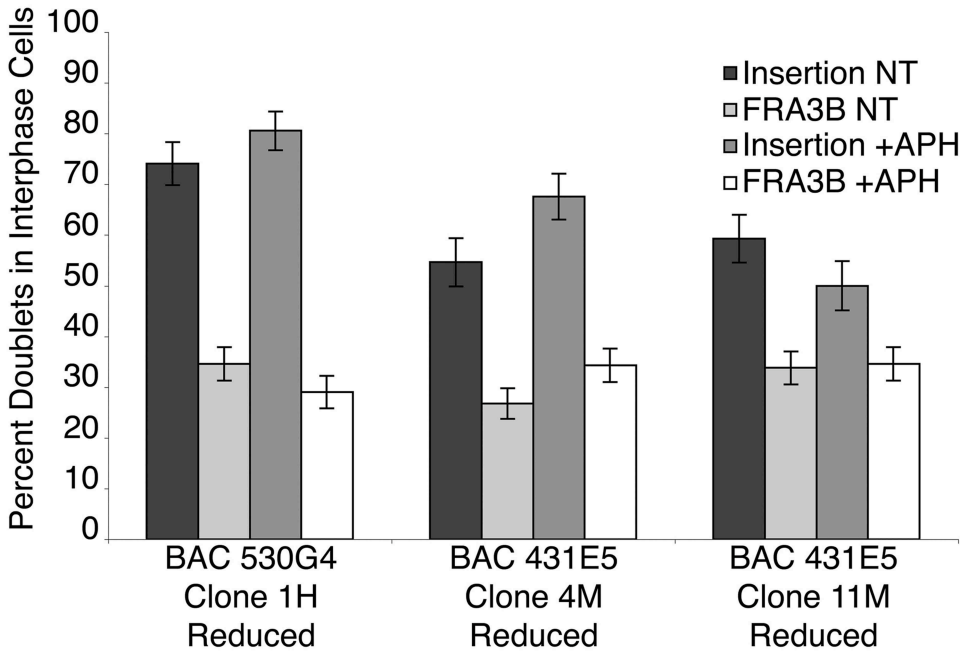
Chromosomal rearrangements at the sites of BAC integrations as seen on metaphase chromosomes. The FISH signal indicates integrated BAC sequence. (**A**) Dicentric chromosome observed in FRA3B BAC clone 431E5 4M. The arrows indicate the two centromeres. (**B**) Ring chromosome in control BAC clone 412D20 6H1. (**C**) Ring chromosome containing multiple FISH signals in the FRA3B BAC clone 530G4 24H3. (**D**) Multiple FISH signals due to amplification in the FRA3B BAC clone 530G4 24H3. (**E**) Chromosome containing fused sister chromatids observed in the FRA3B BAC clone 431E5 4M. (**F-K**) Clonal evolution of aberrations seen in the FRA3B BAC clone 530G4 24H3. (**F**) The FISH signal seen in the majority of cells from this clone. (**G**) A bridge containing multiple FISH signals as seen between two interphase nuclei. (**H**) A large amplification of the integrated BAC. The laddering seen is indicative of multiple rounds of BFBC. (**I**) An example of multiple breaks seen in the amplified region. (**J**) Extrachromosomal fragments presumably originating from unstable amplified integrated BAC. (**K**) Micronuclei containing FISH signals from multiple fragmented chromosomal elements.

Figure 2-7 Frequency of gaps and breaks on chromosomes in reduced clones



Cells containing FRA3B BAC integrations with reduced copy numbers show elevated gaps and breaks at the integration site when treated with 0.4μm APH. All APH treated samples were normalized relative to total gaps and breaks. The copy number for each clone is indicated. (**A**) Percent of integration sites broken as seen in fifty untreated metaphases for each clone. (**B**) The relative percentage of gaps and breaks at the site of integration after treatment with 0.4μm APH as found in both the original clones containing multiple BAC copy numbers and the reduced clones. (**C**) Relative percentage of gaps and breaks at both the integration site and the endogenous FRA3B locus for each reduced clone.

Figure 2-8 Percent of doublet FISH signals in interphase cells



The percent of interphase FISH signals seen as doublets for both the integrated BAC and the two endogenous FRA3B loci in cells containing BAC insertions with reduced copy numbers. 108 interphase cells were scored for each clone and treatment.

References

Alderton GK, Joenje H, Varon R, Borglum AD, Jeggo PA, O'Driscoll M. 2004. Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway. Hum Mol Genet 13:3127-3138.

Arlt MF, Casper AM, Glover TW. 2003. Common fragile sites. Cytogenet Genome Res 100:92-100.

Arlt MF, Miller DE, Beer DG, Glover TW. 2002. Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells. Genes Chromosomes Cancer 33:82-92.

Arlt MF, Xu B, Durkin SG, Casper AM, Kastan MB, Glover TW. 2004. BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function. Mol Cell Biol 24:6701-6709.

Barlow C, Hirotsune S, Paylor R, Liyanage M, Eckhaus M, Collins F, Shiloh Y, Crawley JN, Ried T, Tagle D, Wynshaw-Boris A. 1996. Atm-deficient mice: a paradigm of ataxia telangiectasia. Cell 86:159-171.

Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F, Zieger K, Guldberg P, Sehested M, Nesland JM, Lukas C, Orntoft T, Lukas J, Bartek J. 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434:864-870.

Bednarek AK, Keck-Waggoner CL, Daniel RL, Laflin KJ, Bergsagel PL, Kiguchi K, Brenner AJ, Aldaz CM. 2001. WWOX, the FRA16D gene, behaves as a suppressor of tumor growth. Cancer Res 61:8068-8073.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573-580.

Boldog F, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H, Drabkin HA. 1997. Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193-203.

Boldog FL, Waggoner B, Glover TW, Chumakov I, Le Paslier D, Cohen D, Gemmill RM, Drabkin HA. 1994. Integrated YAC contig containing the 3p14.2 hereditary renal carcinoma 3;8 translocation breakpoint and the fragile site FRA3B. Genes Chromosomes Cancer 11:216-221.

Borglum AD, Balslev T, Haagerup A, Birkebaek N, Binderup H, Kruse TA, Hertz JM. 2001. A new locus for Seckel syndrome on chromosome 18p11.31-q11.2. Eur J Hum Genet 9:753-757.

Brown EJ, Baltimore D. 2000. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev 14:397-402.

Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA. 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev 19:1040-1052.

Casper AM, Durkin SG, Arlt MF, Glover TW. 2004. Chromosomal instability at common fragile sites in Seckel syndrome. Am J Hum Genet 75:654-660.

Casper AM, Nghiem P, Arlt MF, Glover TW. 2002. ATR regulates fragile site stability. Cell 111:779-789.

Cesari R, Martin ES, Calin GA, Pentimalli F, Bichi R, McAdams H, Trapasso F, Drusco A, Shimizu M, Masciullo V, D'Andrilli G, Scambia G, Picchio MC, Alder H, Godwin AK, Croce CM. 2003. Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27. Proc Natl Acad Sci U S A 100:5956-5961.

Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M. 1997. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell 89:215-225.

Coquelle A, Rozier L, Dutrillaux B, Debatisse M. 2002. Induction of multiple double-strand breaks within an hsr by meganucleaseI-SceI expression or fragile site activation leads to formation of double minutes and other chromosomal rearrangements. Oncogene 21:7671-7679.

Corbin S, Neilly ME, Espinosa R, 3rd, Davis EM, McKeithan TW, Le Beau MM. 2002. Identification of unstable sequences within the common fragile site at 3p14.2: implications for the mechanism of deletions within fragile histidine triad gene/common fragile site at 3p14.2 in tumors. Cancer Res 62:3477-3484.

Critchlow SE, Jackson SP. 1998. DNA end-joining: from yeast to man. Trends Biochem Sci 23:394-398.

Dayn A, Malkhosyan S, Duzhy D, Lyamichev V, Panchenko Y, Mirkin S. 1991. Formation of (dA-dT)n cruciforms in Escherichia coli cells under different environmental conditions. J Bacteriol 173:2658-2664.

Denison SR, Callahan G, Becker NA, Phillips LA, Smith DI. 2003a. Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. Genes Chromosomes Cancer 38:40-52.

Denison SR, Wang F, Becker NA, Schule B, Kock N, Phillips LA, Klein C, Smith DI. 2003b. Alterations in the common fragile site gene Parkin in ovarian and other cancers. Oncogene 22:8370-8378.

DiTullio RA, Jr., Mochan TA, Venere M, Bartkova J, Sehested M, Bartek J, Halazonetis TD. 2002. 53BP1 functions in an ATM-dependent checkpoint pathway that is constitutively activated in human cancer. Nat Cell Biol 4:998-1002.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA, Jr., Butel JS, Bradley A. 1992. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature 356:215-221.

Dumon KR, Ishii H, Fong LY, Zanesi N, Fidanza V, Mancini R, Vecchione A, Baffa R, Trapasso F, During MJ, Huebner K, Croce CM. 2001. FHIT gene therapy prevents tumor development in Fhit-deficient mice. Proc Natl Acad Sci U S A 98:3346-3351.

Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. 2002. Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 99:16144-16149.

Durkin SG, Arlt MF, Howlett NG, Glover TW. 2006. Depletion of CHK1, but not CHK2, induces chromosomal instability and breaks at common fragile sites. Oncogene 25:4381-4388.

Durkin SG, Glover TW. 2007. Chromosome Fragile Sites. Annu Rev Genet.

Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW. 2008. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. Proc Natl Acad Sci U S A 105:246-251.

Elder FF, Robinson TJ. 1989. Rodent common fragile sites: are they conserved? Evidence from mouse and rat. Chromosoma 97:459-464.

Fang JM, Arlt MF, Burgess AC, Dagenais SL, Beer DG, Glover TW. 2001. Translocation breakpoints in FHIT and FRA3B in both homologs of chromosome 3 in an esophageal adenocarcinoma. Genes Chromosomes Cancer 30:292-298.

Fang Y, Tsao CC, Goodman BK, Furumai R, Tirado CA, Abraham RT, Wang XF. 2004. ATR functions as a gene dosage-dependent tumor suppressor on a mismatch repair-deficient background. Embo J 23:3164-3174.

Ferber MJ, Eilers P, Schuuring E, Fenton JA, Fleuren GJ, Kenter G, Szuhai K, Smith DI, Raap AK, Brink AA. 2004. Positioning of cervical carcinoma and Burkitt lymphoma translocation breakpoints with respect to the human papillomavirus integration cluster in FRA8C at 8q24.13. Cancer Genet Cytogenet 154:1-9.

Ferber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, McGovern R, Gostout BS, Cheung TH, Chung TK, Fu WY, Smith DI. 2003. Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma. Oncogene 22:7233-7242.

Finnis M, Dayan S, Hobson L, Chenevix-Trench G, Friend K, Ried K, Venter D, Woollatt E, Baker E, Richards RI. 2005. Common chromosomal fragile site FRA16D mutation in cancer cells. Hum Mol Genet 14:1341-1349.

Focarelli ML, Montagna C, Colombo R, Ried T, Vezzoni P, Musio A. 2006. SMC1 inhibition results in FRA3B expression but has no effect on its delayed replication. Mutat Res 595:23-28.

Glover TW, Berger C, Coyle J, Echo B. 1984. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:136-142.

Glover TW, Coyle-Morris JF, Li FP, Brown RS, Berger CS, Gemmill RM, Hecht F. 1988. Translocation t(3;8)(p14.2;q24.1) in renal cell carcinoma affects expression of the common fragile site at 3p14(FRA3B) in lymphocytes. Cancer Genet Cytogenet 31:69-73.

Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, Wilke CM, Dierick HA, Beer DG. 1998. The murine Fhit gene is highly similar to its human orthologue and maps to a common fragile site region. Cancer Res 58:3409-3414.

Glover TW, Stein CK. 1987. Induction of sister chromatid exchanges at common fragile sites. Am J Hum Genet 41:882-890.

Glover TW, Stein CK. 1988. Chromosome breakage and recombination at fragile sites. Am J Hum Genet 43:265-273.

Goodship J, Gill H, Carter J, Jackson A, Splitt M, Wright M. 2000. Autozygosity mapping of a seckel syndrome locus to chromosome 3q22. 1-q24. Am J Hum Genet 67:498-503.

Gorgoulis VG, Vassiliou LV, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, Venere M, Ditullio RA, Jr., Kastrinakis NG, Levy B, Kletsas D, Yoneta A, Herlyn M, Kittas C, Halazonetis TD. 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature 434:907-913.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. Cell 100:57-70.

Handt O, Baker E, Dayan S, Gartler SM, Woollatt E, Richards RI, Hansen RS. 2000. Analysis of replication timing at the FRA10B and FRA16B fragile site loci. Chromosome Res 8:677-688.

Hansen RS, Canfield TK, Fjeld AD, Mumm S, Laird CD, Gartler SM. 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 94:4587-4592.

Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD. 1993. Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell 73:1403-1409.

Harvey J, Judge C, Wiener S. 1977. Familial X-linked mental retardation with an X chromosome abnormality. J Med Genet 14:46-50.

Hellman A, Rahat A, Scherer SW, Darvasi A, Tsui LC, Kerem B. 2000. Replication delay along FRA7H, a common fragile site on human chromosome 7, leads to chromosomal instability. Mol Cell Biol 20:4420-4427.

Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T. 2006. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. Genome Res 16:1222-1230.

Howlett NG, Taniguchi T, Durkin SG, D'Andrea AD, Glover TW. 2005. The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. Hum Mol Genet 14:693-701.

Huebner K, Croce CM. 2003. Cancer and the FRA3B/FHIT fragile locus: it's a HIT. Br J Cancer 88:1501-1506.

Jackson SP. 2002. Sensing and repairing DNA double-strand breaks. Carcinogenesis 23:687-696.

Kilinc MO, Ninis VN, Ugur SA, Tuysuz B, Seven M, Balci S, Goodship J, Tolun A. 2003. Is the novel SCKL3 at 14q23 the predominant Seckel locus? Eur J Hum Genet 11:851-857.

Kim SY, Horrigan SK, Altenhofen JL, Arbieva ZH, Hoffman R, Westbrook CA. 1998. Modification of bacterial artificial chromosome clones using Cre recombinase: introduction of selectable markers for expression in eukaryotic cells. Genome Res 8:404-412.

Kostrub CF, Knudsen K, Subramani S, Enoch T. 1998. Hus1p, a conserved fission yeast checkpoint protein, interacts with Rad1p and is phosphorylated in response to DNA damage. Embo J 17:2055-2066.

Kuo MT, Vyas RC, Jiang LX, Hittelman WN. 1994. Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. Mol Cell Biol 14:5202-5211.

Lakin ND, Hann BC, Jackson SP. 1999. The ataxia-telangiectasia related protein ATR mediates DNA-dependent phosphorylation of p53. Oncogene 18:3989-3995.

Le Beau MM, Rassool FV, Neilly ME, Espinosa R, 3rd, Glover TW, Smith DI, McKeithan TW. 1998. Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. Hum Mol Genet 7:755-761.

Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD. 2005. Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. Cell 120:587-598.

Lewis KA, Mullany S, Thomas B, Chien J, Loewen R, Shridhar V, Cliby WA. 2005. Heterozygous ATR mutations in mismatch repair-deficient cancer cells have functional significance. Cancer Res 65:7091-7095.

Limongi MZ, Pelliccia F, Rocchi A. 2003. Characterization of the human common fragile site FRA2G. Genomics 81:93-97.

Liu A, Takakuwa T, Luo WJ, Fujita S, Aozasa K. 2006. Alterations in ATR in nasal NK/T-cell lymphoma and chronic active Epstein-Barr virus infection. Cancer Sci 97:605-610.

Liu Q, Guntuku S, Cui XS, Matsuoka S, Cortez D, Tamai K, Luo G, Carattini-Rivera S, DeMayo F, Bradley A, Donehower LA, Elledge SJ. 2000. Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev 14:1448-1459.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

Ludes-Meyers JH, Bednarek AK, Popescu NC, Bedford M, Aldaz CM. 2003. WWOX, the common chromosomal fragile site, FRA16D, cancer gene. Cytogenet Genome Res 100:101-110.

Ludes-Meyers JH, Kil H, Nunez MI, Conti CJ, Parker-Thornburg J, Bedford MT, Aldaz CM. 2007. WWOX hypomorphic mice display a higher incidence of B-cell lymphomas and develop testicular atrophy. Genes Chromosomes Cancer 46:1129-1136.

Lundin C, Erixon K, Arnaudeau C, Schultz N, Jenssen D, Meuth M, Helleday T. 2002. Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. Mol Cell Biol 22:5869-5878.

Magenis RE, Hecht F, Lovrien EW. 1970. Heritable fragile site on chromosome 16: probable localization of haptoglobin locus in man. Science 170:85-87.

Mangelsdorf M, Ried K, Woollatt E, Dayan S, Eyre H, Finnis M, Hobson L, Nancarrow J, Venter D, Baker E, Richards RI. 2000. Chromosomal fragile site FRA16D and DNA instability in cancer. Cancer Res 60:1683-1689.

McAllister BF, Greenbaum IF. 1997. How common are common fragile sites: variation of aphidicolin-induced chromosomal fragile sites in a population of the deer mouse (Peromyscus maniculatus). Hum Genet 100:182-188.

McAvoy S, Ganapathiraju SC, Ducharme-Smith AL, Pritchett JR, Kosari F, Perez DS, Zhu Y, James CD, Smith DI. 2007. Non-random inactivation of large common fragile site genes in different cancers. Cytogenet Genome Res 118:260-269.

Menoyo A, Alazzouzi H, Espin E, Armengol M, Yamamoto H, Schwartz S, Jr. 2001. Somatic mutations in the DNA damage-response genes ATR and CHK1 in sporadic stomach tumors with microsatellite instability. Cancer Res 61:7727-7730.

Michael D, Beer DG, Wilke CW, Miller DE, Glover TW. 1997. Frequent deletions of FHIT and FRA3B in Barrett's metaplasia and esophageal adenocarcinomas. Oncogene 15:1653-1659.

Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, Huebner K, Croce CM. 1999. Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. Proc Natl Acad Sci U S A 96:7456-7461.

Mirkin EV, Mirkin SM. 2007. Replication fork stalling at natural impediments. Microbiol Mol Biol Rev 71:13-35.

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B. 1998. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci U S A 95:8141-8146.

Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A, Ried T, Vezzoni P. 2005. SMC1 involvement in fragile site expression. Hum Mol Genet 14:525-533.

Myers JS, Cortez D. 2006. Rapid activation of ATR by ionizing radiation requires ATM and Mre11. J Biol Chem 281:9346-9350.

Nachman MW, Crowell SL. 2000. Estimate of the mutation rate per nucleotide in humans. Genetics 156:297-304.

Negrini M, Monaco C, Vorechovsky I, Ohta M, Druck T, Baffa R, Huebner K, Croce CM. 1996. The FHIT gene at 3p14.2 is abnormal in breast carcinomas. Cancer Res 56:3173-3179.

O'Driscoll M, Ruiz-Perez VL, Woods CG, Jeggo PA, Goodship JA. 2003. A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet 33:497-501.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K. 1996. The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84:587-597.

Ozeri-Galai E, Schwartz M, Rahat A, Kerem B. 2008. Interplay between ATM and ATR in the regulation of common fragile site stability. Oncogene 27:2109-2117.

Palakodeti A, Han Y, Jiang Y, Le Beau MM. 2004. The role of late/slow replication of the FRA16D in common fragile site induction. Genes Chromosomes Cancer 39:71-76.

Pelliccia F, Bosco N, Curatolo A, Rocchi A. 2008. Replication timing of two human common fragile sites: FRA1H and FRA2G. Cytogenet Genome Res 121:196-200.

Rassool FV, McKeithan TW, Neilly ME, van Melle E, Espinosa R, 3rd, Le Beau MM. 1991. Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. Proc Natl Acad Sci U S A 88:6657-6661.

Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S, Nancarrow JK, Woollatt E, Kremmidiotis G, Gardner A, Venter D, Baker E, Richards RI. 2000. Common chromosomal fragile site FRA16D sequence: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells. Hum Mol Genet 9:1651-1663.

Roeder GS, Fink GR. 1980. DNA rearrangements associated with a transposable element in yeast. Cell 21:239-249.

Rozier L, El-Achkar E, Apiou F, Debatisse M. 2004. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1:

possible association with an inherited disease and cancer. Oncogene 23:6872-6880.

Ruiz-Herrera A, Garcia F, Fronicke L, Ponsa M, Egozcue J, Caldes MG, Stanyon R. 2004. Conservation of aphidicolin-induced fragile sites in Papionini (Primates) species and humans. Chromosome Res 12:683-690.

Saintigny Y, Delacote F, Vares G, Petitot F, Lambert S, Averbeck D, Lopez BS. 2001. Characterization of homologous recombination induced by replication inhibition in mammalian cells. Embo J 20:3861-3870.

Schreiber W. 1979. Jimson seed intoxication: recognition and therapy. Mil Med 144:329-336.

Schwartz M, Zlotorynski E, Goldberg M, Ozeri E, Rahat A, le Sage C, Chen BP, Chen DJ, Agami R, Kerem B. 2005. Homologous recombination and nonhomologous end-joining repair pathways regulate fragile site stability. Genes Dev 19:2715-2726.

Shechter D, Costanzo V, Gautier J. 2004. Regulation of DNA replication by ATR: signaling in response to DNA intermediates. DNA Repair (Amst) 3:901-908.

Shiraishi T, Druck T, Mimori K, Flomenberg J, Berk L, Alder H, Miller W, Huebner K, Croce CM. 2001. Sequence conservation at human and mouse orthologous common fragile regions, FRA3B/FHIT and Fra14A2/Fhit. Proc Natl Acad Sci U S A 98:5722-5727.

Siprashvili Z, Sozzi G, Barnes LD, McCue P, Robinson AK, Eryomin V, Sard L, Tagliabue E, Greco A, Fusetti L, Schwartz G, Pierotti MA, Croce CM, Huebner K. 1997. Replacement of Fhit in cancer cells suppresses tumorigenicity. Proc Natl Acad Sci U S A 94:13771-13776.

Smeets DF, van de Klundert FA. 1990. Common fragile sites in man and three closely related primate species. Cytogenet Cell Genet 53:8-14.

Smith L, La Thangue NB. 2005. Signalling DNA damage by regulating p53 co-factor activity. Cell Cycle 4:30-32.

Soulie J, De Grouchy J. 1981. A cytogenetic survey of 110 baboons (Papio cynocephalus). Am J Phys Anthropol 56:107-113.

Stiff T, Walker SA, Cerosaletti K, Goodarzi AA, Petermann E, Concannon P, O'Driscoll M, Jeggo PA. 2006. ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. Embo J 25:5775-5782.

Stone DM, Jacky PB, Hancock DD, Prieur DJ. 1991. Chromosomal fragile site expression in dogs: I. Breed specific differences. Am J Med Genet 40:214-222.

Strunnikov AV, Jessberger R. 1999. Structural maintenance of chromosomes (SMC) proteins: conserved molecular properties for multiple biological functions. Eur J Biochem 263:6-13.

Sutherland GR. 2003. Rare fragile sites. Cytogenet Genome Res 100:77-84.

Sutherland GR, Parslow MI, Baker E. 1985. New classes of common fragile sites induced by 5-azacytidine and bromodeoxyuridine. Hum Genet 69:233-237.

Thorland EC, Myers SL, Gostout BS, Smith DI. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. Oncogene 22:1225-1237.

Tibbetts RS, Cortez D, Brumbaugh KM, Scully R, Livingston D, Elledge SJ, Abraham RT. 2000. Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. Genes Dev 14:2989-3002.

Tsantoulis PK, Kotsinas A, Sfikakis PP, Evangelou K, Sideridou M, Levy B, Mo L, Kittas C, Wu XR, Papavassiliou AG, Gorgoulis VG. 2007. Oncogene-induced replication stress preferentially targets common fragile sites in preneoplastic lesions. A genome-wide study. Oncogene.

Wang F, Denison S, Lai JP, Philips LA, Montoya D, Kock N, Schule B, Klein C, Shridhar V, Roberts LR, Smith DI. 2004. Parkin gene alterations in hepatocellular carcinoma. Genes Chromosomes Cancer 40:85-96.

Wang L, Darling J, Zhang JS, Huang H, Liu W, Smith DI. 1999. Allele-specific late replication and fragility of the most active common fragile site, FRA3B. Hum Mol Genet 8:431-437.

Wang ND, Testa JR, Smith DI. 1993. Determination of the specificity of aphidicolin-induced breakage of the human 3p14.2 fragile site. Genomics 17:341-347.

Wang Z, Engler P, Longacre A, Storb U. 2001. An efficient method for high-fidelity BAC/PAC retrofitting with a selectable marker for mammalian cell transfection. Genome Res 11:137-142.

Warburton PE, Giordano J, Cheung F, Gelfand Y, Benson G. 2004. Inverted repeat structure of the human genome: the X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes. Genome Res 14:1861-1869.

Wilke CM, Guo SW, Hall BK, Boldog F, Gemmill RM, Chandrasekharappa SC, Barcroft CL, Drabkin HA, Glover TW. 1994. Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. Genomics 22:319-326.

Wilke CM, Hall BK, Hoge A, Paradee W, Smith DI, Glover TW. 1996. FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. Hum Mol Genet 5:187-195.

Wilson DM, 3rd, Thompson LH. 2007. Molecular mechanisms of sister-chromatid exchange. Mutat Res 616:11-23.

Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol Chapter 2:Unit 2 4.

Yang J, Zhang W. 2008. WWOX tumor suppressor gene. Histol Histopathol 23:877-882.

Zhang H, Freudenreich CH. 2007. An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in S. cerevisiae. Mol Cell 27:367-379.

Zhao H, Piwnica-Worms H. 2001. ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. Mol Cell Biol 21:4129-4139.

Zhu M, Weiss RS. 2007. Increased common fragile site expression, cell proliferation defects, and apoptosis following conditional inactivation of mouse Hus1 in primary cultured cells. Mol Biol Cell 18:1044-1055.

Zhu Y, McAvoy S, Kuhn R, Smith DI. 2006. RORA, a large common fragile site gene, is involved in cellular stress response. Oncogene 25:2901-2908.

Zlotorynski E, Rahat A, Skaug J, Ben-Porat N, Ozeri E, Hershberg R, Levi A, Scherer SW, Margalit H, Kerem B. 2003. Molecular basis for expression of common and rare fragile sites. Mol Cell Biol 23:7143-7151.

Chapter III

Mice with hypomorphic Atr expression exhibit increased APH-induced DNA damage and checkpoint abnormalities.

Abstract

[2]The ATR checkpoint pathway responds to DNA damage during the S/G2 phases of the cell cycle and is activated early in tumorigenesis. Investigation of ATR's role in development and tumorigenesis is complicated by the lethality of homozygous knock-out mice and the limited effects of heterozygous deficiency. To overcome this limitation, we sought to create mice with a hypomorphic Atr mutation based on the ATR mutation in the human disease Seckel syndrome-1 (SCKL1). Homozygous SCKL1 mice were generated by targeted knock-in of the A→G SCKL1 mutation. Western blot and RT-PCR analysis established that homozygotes have no reduction in Atr protein or increase in mis-splicing as is seen in humans. Thus, the A→G substitution alone is not sufficient to reproduce in mice the effects that are seen in humans. However, homozygous SCKL1 mice retaining the *neo* cassette used for targeting have an estimated 66-82% reduction in total Atr protein levels due to mis-splicing into the *neo* cassette. Under conditions of APH-induced replication stress,

[2] This chapter has been previously submitted for publication: Ragland RL, Arlt MF, Hughes ED, Saunders TL, Glover TW (2009) Mice with hypomorphic Atr expression exhibit increased APH-induced DNA damage and checkpoint abnormalities. Mammalian Genome *Submitted*.

primary fibroblasts from homozygous mice displayed an increase in overall chromosomes damage and an increase in gaps and breaks at specific common fragile sites. In addition, mutant cells display a significant delay in checkpoint induction and an increase in DNA damage as assayed by Chk1 phosphorylation and γ-H2ax levels, respectively. These mice provide a novel model system for studies of Atr deficiency and replication stress.

Introduction

The ataxia telangiectasia-mutated and Rad3-related (ATR) kinase is a member of the phosphoinositide 3-kinase-related kinase family of proteins, which includes ataxia-telangiectasia-mutated (ATM). These kinases mediate critical signal transduction roles in the cellular DNA damage response and are essential for maintenance of genomic stability. ATR responds primarily to DNA damage caused by agents that impede replication forks, such as UV irradiation (UV), hydroxyurea (HU), and aphidicolin (APH) (Abraham, 2001). In response to DNA damage, ATR functions to stall the cell cycle via phosphorylation of downstream effecter proteins, including several genes with cancer involvement, such as BRCA1, CHK1, p53, and H2AX (Lakin *et al.*, 1999; Liu *et al.*, 2000; Smith & La Thangue, 2005; Zhao & Piwnica-Worms, 2001; Zhou & Elledge, 2000). ATM primarily responds to agents that generate double strand DNA breaks, such as ionizing radiation (IR), and functions via phosphorylation of a number of downstream targets, including CHK2 (Bakkenist & Kastan, 2003; Matsuoka *et al.*, 2000; Shiloh, 2006; Zhou & Elledge, 2000). Although ATR and ATM primarily respond to different forms of DNA damage, there is a considerable amount of functional overlap of between these two proteins. It has been shown that ATR is activated in an ATM-dependent

manner following IR (Myers & Cortez, 2006) and that ATM is activated in an ATR-dependent manner following UV irradiation (Jazayeri et al., 2006; Stiff et al., 2006).

While homozygous loss of *ATR* is not compatible with cellular viability, heterozygous loss or mutation is not lethal and has been found in a number of sporadic mismatch repair-deficient stomach, colon, and endometrial tumors (Lewis *et al.*, 2005; Menoyo *et al.*, 2001; Vassileva *et al.*, 2002) and also in nasal natural killer T-cell lymphomas, suggesting that *ATR* may function as a haploinsufficient tumor suppressor (Lee et al., 2006). In addition, mismatch repair-deficient cancer cell lines containing heterozygous ATR mutations are hypersensitive to UV, IR, cisplatin, and topoisomerase inhibitors indicative of a defective DNA damage response (Lewis et al., 2005).

Given the similarity in function between the ATM and ATR kinases, the number of known tumor suppressor genes that ATR regulates, and the finding of mutations in some tumors it is possible that ATR functions as a tumor suppressor gene. However, unlike ATM, experimental studies of the role of ATR in tumorigenesis and development is complicated by lethality of homozygous knock-out mice and the limited effects of heterozygous deficiency (Brown & Baltimore, 2000; Cortez *et al.*, 2001).Two groups have studied cancer incidence in heterozygous Atr mice. Brown and Baltimore (Brown & Baltimore, 2000) reported a small (3-4 fold) increase in late-life tumors in in Atr heterozygotes ($Atr^{+/-}$) and in Atr and Atm double heterozygotes ($Atr^{+/-}Atm^{+/-}$) compared to Atm heterozygotes ($Atm^{+/-}$). In addition, Fang et al (Fang et al., 2004) found that 10 of 13 $Atr^{+/-}/Mlh^{-/-}$ mice developed tumors compared to 2 of 7 $Mlh^{-/-}$ mice. The tumors found were primarily lymphomas and intestinal adenocarcinomas. Deletion of ATR in adult mice using Cre/lox technology led to defects in tissue homeostasis and the appearance of

age-related phenotypes, such as hair graying, alopecia, osteoporosis, fibrosis, and others (Ruzankina et al., 2007). In these mice, tissues with continuous cell proliferation and no ATR expression displayed acute cellular loss. This loss was associated with a reduction in tissue specific stem and progenitor cells, and an exhaustion of tissue renewal and homeostatic capacity. These studies suggested that reduced regenerative capacity in adults via deletion of ATR is sufficient to cause the premature appearance of age related phenotypes.

Given the difficulty of studying ATR *in vivo*, almost all of our current knowledge about the function of ATR and the consequence of deficiency of this gene is based on *in vitro* cell culture studies. However, an alternative approach to this problem presents itself from studies of the human genetic disorder Seckel syndrome 1 (SCKL1). SCKL1 is an autosomal recessive disorder caused by a point mutation in exon 9 of *ATR*, (2101A→G) (O'Driscoll et al., 2003). This mutation results in the expression of high levels of an alternative ATR splice variant that lacks exon 9 and introduces a stop codon in exon 10, causing premature termination of the transcript and hypomorphic expression from the mutant allele (O'Driscoll et al., 2003). Individuals with this disease display severe microcephaly and both growth and mental retardation. Because only two families with a total of four affected children have been identified with this mutation, it is not yet known if there is an increased risk for tumor formation (Goodship et al., 2000). However, cell lines derived from these SCKL1 individuals display a deficiency in the ATR checkpoint response and an increased sensitivity to DNA damage. Alderton et. al. (Alderton et al., 2004) reported an increase in micronuclei formation, an increase in nuclear fragmentation, impaired phosphorylation of CHK1(Ser317), and a failure to arrest in G2/M after IR treatment. It has also been shown that SCKL1 cells

have increased aphidicolin-induced chromosome breakage, predominantly at common fragile sites, which are particularly sensitive to this form of replication stress (Casper et al., 2004).

Here we report the creation and initial characterization of a knock-in hypomorphic Atr mouse model. Mice containing the 2102 A→G mutation did not exhibit an increase in aberrant Atr splicing or reduced Atr expression as is seen in SCKL1. However, homozygous mice that retained the *neo* cassette used for gene targeting exhibited a 66-82% reduction in the levels of functional Atr, increased genomic instability, delayed Atr checkpoint response, an increased Atm checkpoint response, and decreased mitotic index. These mice are novel reagents for studies of Atr deficiency, replication stress, and genomic instability on development and tumorigenesis.

Results

Analysis of Atr splicing in normal mouse fibroblasts and $Atr^{Sckl1/Sckl1}$ ES cells.

The SCKL1 mutation in humans leads to an estimated 90% decrease in the normal ATR splice form. In order to determine if Atr splices in an analogous manner in mice we performed RT-PCR with primers in exons 8 and 10 of mouse Atr (Figure 1a, primers p4 and p5) in normal mouse fibroblasts and sequenced the resulting products. This experiment yielded two bands, the first of which corresponded in size and sequence to a normal splice product containing exons 8, 9, and 10. The second, fainter band corresponded in size and sequence to a mis-spliced mRNA product lacking exon 9 (data not shown). The presence of this fainter band lacking exon 9 showed that, like humans, normal mouse fibroblasts produce small amounts of mis-spliced Atr. These findings supported the hypothesis that creation of a point mutation in the mouse equivalent to the

human SCKL1 mutation should disrupt normal splicing and allow us to generate a viable mouse model with hypomorphic Atr expression. In order to accomplish this, we used knock in technology to alter the endogenous mouse exon 9 in embryonic stem (ES) cells to contain the A→G mutation that is analogous to the human 2101 A→G mutation found in SCKL1 (Figure 1a).

Prior to the creation of mice containing the A→G mutation, we examined heterozygous ES cells for any alterations in splicing caused by the presence of this mutation. Using RT-PCR, ES cells were analyzed using primers located in exon 8 and 10 (Figure 1a primers p4 and p5). We did not observe an increase in intensity of the band equivalent to loss of exon 9 as seen in SCKL1 cells. However, some laddering of bands corresponding to the production of larger splice products was observed when primers spanning the *neo* cassette were used (data not shown). In addition, sequencing of the normal splice products revealed an under representation of the RNA products from the allele containing the mutant guanine nucleotide. These results suggested that the A→G substitution in mouse cells does not function in a manner analogous to the human SCKL1 mutation. However, these data also showed a reduction in production of normally spliced Atr from the targeted allele and the production of novel mis-spliced products.

Generation and analysis of Atr$^{Sckl1/Sckl1}$ mice.

To examine the effect of the A→G mutation in mice, we bred a male mouse heterozygous for the allele containing both the *neo* cassette and the A→G mutation to two rounds of female B6.Cg-Tg(ACTFLPe)9205Dym/J mice containing the FLPe transgene to remove the *neo* cassette via FRT recombination (Rodriguez et al., 2000).

This resulted in mice heterozygous for the Sckl1 mutation and lacking the *neo* cassette ($Atr^{Sckl1/+}$). These mice were backcrossed onto a C57BL/6J background for five generations before production of homozygous offspring. From two backcrossed heterozygote matings ($Atr^{Sckl1/+}$ x $Atr^{Sckl1/+}$) 19 offspring were produced, 7 of which were wild type ($Atr^{+/+}$), 8 of which were heterozygous ($Atr^{Sckl1/+}$), and 4 of which were homozygotes ($Atr^{Sckl1/Sckl1}$). These mice were viable and overtly normal in regards to their physical appearance.

Primary tail fibroblasts grown from tail biopsies were analyzed by western blot and RT-PCR for Atr protein and mRNA levels, respectively. No difference in Atr protein expression was seen between fibroblasts from one $Atr^{+/+}$, one $Atr^{Sckl1/+}$, and three $Atr^{Sckl1/Sckl1}$ mice (Figure 2a). In addition, RT-PCR analysis of total mRNA extracted from primary tail fibroblasts, spleen, kidney, and liver tissue showed no increase in the quantity of the aberrant splice product lacking exon 9 or laddering of splice products (Figure 2b). Because the presence of the A→G mutation was confirmed by both PCR and sequence analysis, this finding showed that the Sckl1 mutation in mice by itself did not alter the expression of Atr at either the protein or mRNA level in the tissues examined.

Generation of $Atr^{SckN/SckN}$ mice.

Although $Atr^{Sckl1/Sckl1}$ mice did not show a change in splicing or reduction in Atr protein expression, the targeted ES cells used to produce these mice did show altered RNA expression. The heterozygous nature of the ES cells and the presence of the *neo* cassette in the targeted allele are the primary differences between the ES cells previously

examined and primary $Atr^{Sckl1/Sckl1}$ fibroblasts. This, together with the RT-PCR results from the ES cells, suggested that *neo* could be interfering in correct splicing of Atr and could provide an alternative method for creation of a mouse hypomorphic for Atr expression. To examine the effect of this allele on Atr expression, the original male chimera, heterozygous for both *neo* and the Sckl1 mutation ($Atr^{SckN/+}$), was bred to female C57BL/6J mice creating several heterozygous offspring. These heterozygotes were crossed to each other resulting in five litters of pups consisting of seven wild type pups ($Atr^{+/+}$), 27 heterozygotes ($Atr^{SckN/+}$), and 16 homozygotes ($Atr^{SckN/SckN}$). These mice are therefore mixtures of the 129SvEvTac and C57BL/6J backgrounds.

$Atr^{SckN/SckN}$ mice have reduced Atr protein levels.

In order to determine the effect of the Atr^{SckN} allele on Atr expression, primary tail fibroblast cell lines established from biopsies from six $Atr^{+/+}$, six $Atr^{SckN/+}$, and seven $Atr^{SckN/SckN}$ mice were analyzed by western blot (Figure 3a-c). In addition, these cell lines were examined by quantitative real time RT-PCR using a probe designed to only detect full length transcripts and located downstream from both the *neo* cassette and the A→G mutation. Fibroblasts cultured from six $Atr^{+/+}$ animals all had similar levels of Atr expression (Figure 3a). However, fibroblasts from seven $Atr^{SckN/SckN}$ mice all showed strong reduction in Atr levels, ranging from an estimated 66% - 82% reduction in protein levels compared to $Atr^{+/+}$ littermate control fibroblasts (Figure 3b). Interestingly, fibroblasts from the six $Atr^{SckN/+}$ mice analyzed were also found to have strongly reduced (57% - 81%) Atr expression (Figure 3b). The mRNA levels were determined by quantitative real time RT-PCR and correlated well with estimates of protein reduction

made based on band intensity in western blot analysis (Figure 3d). The variation of Atr expression within each genotype is not unexpected based on similar results obtained from other mouse models using *neo* as a means to generate hypomorphic expression of a given allele (Levin & Meisler, 2004; Meyers *et al.*, 1998; Nagy *et al.*, 1998).

In addition to fibroblasts, we examined Atr expression in the spleen. Examination of spleen tissue derived from one $Atr^{+/+}$, one $Atr^{SckN/+}$, and one $Atr^{SckN/SckN}$ mouse showed a reduction in Atr protein expression in the tissues derived from the $Atr^{SckN/SckN}$ mouse (Figure 3e).

Correct splicing of *Atr* is interrupted by *neo*.

Our RT-PCR results in the targeted ES cells and previously published results showing that the presence of a *neo* cassette can be used to generate mice with variable expression and hypomorphic expression of targeted alleles suggested that normal Atr splicing in our mice could be interrupted by mis-splicing into the *neo* cassette (Levin & Meisler, 2004; Meyers *et al.*, 1998; Nagy *et al.*, 1998). We therefore examined total mRNA extracted from cultured primary tail fibroblasts for aberrant splicing using RT-PCR analysis.

When primers located in exons 8 and 10 of Atr (Figure 4, primers p4 and p5) were used, a small increase in the aberrant splice product corresponding to loss of exon 9 was observed in fibroblasts from both $Atr^{SckN/+}$ and $Atr^{SckN/SckN}$ mice (Figure 4b). Although this increase is not as robust as was found in human patients with SCKL1, it suggests that some aberrant splicing similar to what is seen in SCKL1 patients is occurring. In addition, as was seen in the ES cells, sequencing results from $Atr^{SckN/+}$

fibroblasts demonstrated that the mRNA from the mutated Atr^{SckN} allele containing the guanine nucleotide was underrepresented in the normal splice band (Figure 4c). When primers designed to capture the 5' end of splicing into *neo* (Figure 4a, primers p1 and p3) were used, three distinct products were found (Figure 4d). We cloned and sequenced two of these products and found that, in the reverse orientation, *neo* contains a functional cononical splice accepter and a small internal intron. Similarly, when primers designed to capture the 3' end of a splice into *neo* (Figure 4a, primers p2 and p5) were used, at least four distinct splice products were found (Figure 4d). We cloned and sequenced two of these four products. A cononical splice donor site was found at the 3' end of *neo* and the two splice products cloned correspond to either the presence or absence of Atr exon 9 in these splice forms (Figure 4c). These results are consistent with the literature and demonstrate that *neo* is interfering with correct splicing of Atr, thus creating a hypomorphic allele and a series of animals with variable Atr expression (Levin & Meisler, 2004; Meyers *et al.*, 1998; Nagy *et al.*, 1998).

$Atr^{SckN/SckN}$ mice are sensitive to APH mediated replication stress.

Deficiency of Atr leads to a number of readily measurable cellular phenotypes. These include an increase in the number of gaps and breaks on metaphase chromosomes and at common fragile sites following treatment with the DNA polymerase inhibitor APH (Casper *et al.*, 2004; Casper *et al.*, 2002). To determine if the reduction in Atr protein levels found in $Atr^{SckN/SckN}$ mice was sufficient to cause this effect, we examined fibroblasts from one $Atr^{+/+}$ mouse, one $Atr^{SckN/+}$ mouse, and four $Atr^{SckN/SckN}$ mice for total gaps and breaks on metaphase chromosomes and for gaps and breaks at specific

common fragile sites. Four $Atr^{SckN/SckN}$ mice were examined: 1353, 1354, 1359, and 1356 with 78%, 71%, 68%, and 66% reduction in Atr protein, respectively. Fibroblasts were either left untreated or were treated with 0.1 or 0.4 μM APH for 24 hours and metaphase spreads analyzed for chromosome gaps and breaks. Although no significant difference was found between samples at the lower APH treatment (0.1μM), at 0.4μM APH, there was a statistically significant increase in the total number of gaps and breaks in both the $Atr^{SckN/+}$ fibroblasts ($p= 0.01$) and all of the $Atr^{SckN/SckN}$ fibroblasts ($p<0.0001$) compared to the $Atr^{+/+}$ fibroblasts (Figure 5 b).

The mouse common fragile sites Fra14A2 and Fra8E1, which are orthologous to human CFSs FRA3B and FRA16D, respectively, were examined for gaps and breaks by FISH analysis using YAC and BAC probes to these loci. There was a statistically significant increase ($p<0.005$) in the number of gaps and breaks at both CFSs in both of the two $Atr^{SckN/SckN}$ fibroblasts with the greatest reductions in Atr (mouse 1353 and 1354) as compared to the control $Atr^{+/+}$ fibroblasts when treated with 0.4μM APH. Cells derived from $Atr^{SckN/SckN}$ mice with a more modest reduction in Atr (mouse 1356 and 1359) were not as sensitive to APH treatment and were only statistically different ($p<0.05$) from wild type controls at the CFS Fra8E1 suggesting an Atr dosage threshold may be in effect. These results are consistent with previously reported sensitivity of human SCKL1 cells to APH-induced replication stress (Casper *et al.*, 2004; Casper *et al.*, 2002) and suggest that at the cellular levels of Atr expression in the $Atr^{SckN/SckN}$ mouse recapitulated well what occurs cytogenetically in human cells with low Atr expression.

Because $Atr^{SckN/+}$ mice were found to have a similar reduction in Atr levels as $Atr^{SckN/SckN}$ mice, we compared total gaps and breaks after 0.4μM APH treatment in four

additional $Atr^{SckN/+}$ fibroblast cell lines. We found that, in all cases, fibroblasts from $Atr^{SckN/+}$ mice displayed significantly more gaps and breaks after APH treatment than did $Atr^{+/+}$ fibroblasts ($p<0.01$) but less breakage than $Atr^{SckN/SckN}$ fibroblasts ($p<0.001$) (Figure 5d). These results showed that the $Atr^{SckN/+}$ fibroblasts, while as reduced in Atr expression by western analysis, are not as sensitive to APH-induced chromosome breaks as the $Atr^{SckN/SckN}$ fibroblasts.

We observed a decrease in mitotic index in fibroblasts derived from both $Atr^{SckN/+}$ and $Atr^{SckN/SckN}$ mice compared to those from $Atr^{+/+}$ mice. Data from over 2000 cells were therefore examined from all genotypes and the percent of nuclei in metaphase was determined. No difference in mitotic index in untreated fibroblasts was found between genotypes suggesting that all cell types were dividing at similar rates under normal conditions. However, after 24 hours of 0.4μM APH treatment, all four $Atr^{SckN/SckN}$ cell lines and the one $Atr^{SckN/+}$ examined showed a statistically significant decrease in mitotic index ($p<0.05$) as compared to $Atr^{+/+}$ controls (Figure 5a). These results suggest that $Atr^{SckN/SckN}$ and $Atr^{SckN/+}$ fibroblasts are more sensitive to APH than their $Atr^{+/+}$ counterparts.

Checkpoint signaling after replication stress is altered in $Atr^{SckN/SckN}$ cells.

Because cells taken from human patients with SCKL1 display a reduced checkpoint response in the form of delayed Chk1 phosphorylation, we examined primary tail fibroblasts cultured from our mice for Chk1 phosphorylation after APH-induced replication stress. Cells were plated at a density of $5X10^5$ cells/10cm dish, treated with 0.4μM APH for various time points, and harvested for western blot analysis (Figure 6a-

d). Fibroblasts derived from $Atr^{+/+}$ mice phosphorylated Chk1 as early as 15 minutes post APH treatment and that the amount of Chk1 phosphorylation increased at each time point until 16-hours post treatment where measurable phosphorylated Chk1 reached saturation by western blot analysis (Figure 6). Strikingly, fibroblasts derived from $Atr^{SckN/SckN}$ mice showed little to no phosphorylation of Chk1 until the 16-hour time point at which time Chk1 phosphorylation appeared to be saturated (Figure 6). Fibroblasts derived from a $Atr^{SckN/+}$ mouse had an intermediate phenotype, displaying only a moderate decrease in Chk1 phosphorylation at the 15 and 30 minute time points as compared to $Atr^{+/+}$ cells. These experiments were repeated with equivalent results.

Phosphorylated H2ax (γ-H2ax) is activated by both Atr and Atm and can be used as an approximate measure of DNA damage (Burma *et al.*, 2001; Ward & Chen, 2001). In order to observe the effect of reduced Atr expression on γ-H2ax levels we examined γ-H2ax by western blot analysis in fibroblasts. Fibroblasts cultured from an $Atr^{+/+}$ mouse showed no discernable γ-H2ax when untreated and only a low level of phosphorylation until 16-hours post APH treatment at which time the western blot detection of γ-H2ax reached saturation (Figure 6a,e). In contrast to these findings, fibroblasts derived from $Atr^{SckN/SckN}$ mice showed a low level of increased γ-H2ax when untreated and an earlier saturation of γ-H2ax, 8-hours post APH treatment rather than 16 hours (Figure 6b,e). Similar to results from analysis of Chk1 phosphorylation, fibroblasts derived from an $Atr^{SckN/+}$ mouse had a phenotype that is between the $Atr^{+/+}$ and $Atr^{SckN/SckN}$ fibroblasts exhibiting slightly elevated levels of γ-H2ax as compared to $Atr^{+/+}$ mice when left untreated and at 8 hours post APH treatment (Figure 6c,e).

Given that Chk1 is phosphorylated at a much later time point than γ-H2ax in $Atr^{SckN/SckN}$ cells, the basis for the early formation of γ-H2ax was unclear. Because ATM has been shown to be capable of phosphorylating H2AX in the absence of ATR in human cells (Burma et al., 2001), we reasoned that the Atm pathway is responsible for the early γ-H2ax formation in our mouse cells. To examine this possibility, we measured the phosphorylation levels of Chk2 a downstream target of Atm after APH treatment in fibroblasts by western blot. Fibroblasts cultured from an $Atr^{+/+}$ mouse showed no discernible phosphorylation of Chk2 even at 24 hours post APH treatment (Figure 6a). However, $Atr^{SckN/SckN}$ fibroblasts showed strong phosphorylation of Chk2 at all time points post APH treatment and even when untreated (Figure 6b). Fibroblasts derived from an $Atr^{SckN/+}$ mouse again displayed an intermediate phenotype showing phosphorylation of Chk2 at all time points post APH treatment but very little phosphorylated Chk2 when untreated (Figure 6c).

Discussion

We have shown that insertion of a *neo* cassette into intron 7 of Atr interrupts normal splicing, creating an alternative Atr hypomorphic mouse model. However, the introduction of the Atr 2101 A→ G mutation identified in SCKL1 is not sufficient to produce aberrant splicing and reduced Atr expression in mice. Several groups have previously observed that the introduction of a *neo* cassette can be used to generate mice with variable hypomorphic expression of targeted alleles in mice, thus mimicking an allelic series (Levin & Meisler, 2004; Meyers *et al.*, 1998; Nagy *et al.*, 1998). Here it is

shown that the Atr^{SckN} allele acts similarly and homozygous $Atr^{SckN/SckN}$ mice provide novel reagents to examine the effects of low Atr expression in viable adult mice.

In humans hypomorphic expression of Atr has been associated with a Seckel syndrome phenotype that includes microcephaly, dwarfism, increased DNA damage, and impaired checkpoint response (Alderton *et al.*, 2004; O'Driscoll *et al.*, 2007; O'Driscoll *et al.*, 2003). Homozygous loss of Atr, in mice, is early embryonic lethal and leads to chromosome breaks and failure to proliferate in cell culture (Brown & Baltimore, 2000), whereas deletion of Atr in adult mice leads to loss of replicative capacity and a premature aging phenotype (Ruzankina et al., 2007). Heterozygous loss of Atr can cause an increased mortality rate, an increase in late life tumor formation, and an increase in tumor formation on a mismatch repair deficient background but is not associated with dwarfism or microcephaly (Brown & Baltimore, 2000; Fang *et al.*, 2004). These findings demonstrate a strong correlation between Atr expression levels and severity of phenotype. This is demonstrated in our mice at both the organism and cellular levels. The estimated expression levels of Atr protein in our homozygous mice falls in a range between those seen in $Atr^{+/-}$ heterozygotes and human cells homozygous for the SCKL1 allele. Although quantitative measurements were not performed, $Atr^{SckN/SckN}$ mice were not overtly dwarfed, microcephallic, prematurely aged, and as of 10 months of age do not have any obvious spontaneous tumor formation. While it is possible that these particular abnormalities are not present due to unknown differences between humans and mice, it is more likely that the levels of Atr expression in $Atr^{SckN/SckN}$ mice are simply not low enough to cause these abnormalities. Given that the abnormalities associated with Atr deficiency appear to be very strongly affected by gene dosage, crosses to mice with null

alleles (Brown & Baltimore, 2000) could enhance the abnormalities described here and possibly induce others not currently observed.

On the cellular level, Atr expression levels correlated with the DNA damage response and severity of the chromosomal phenotype. This was seen in relation to total gaps and breaks on metaphase chromosomes (Figure 5b), gaps and breaks at specific CFSs (Figure 5c), ability to activate checkpoints (Figure 6), and levels of γ-H2ax and phosphorylated Chk2 after replication stress (Figure 6).

$Atr^{SckN/SckN}$ fibroblasts treated with APH showed activation of γ-H2ax at an earlier time point and have an increased inherent level of γ-H2ax in untreated fibroblast as compared to $Atr^{+/+}$ fibroblasts. Concurrent with this finding, Chk2 is constitutively phosphorylated in $Atr^{SckN/SckN}$ fibroblasts. It has been shown that, in human cells, ATM is capable of phosphorylating H2AX in response to double strand DNA breaks (Burma et al., 2001). Furthermore, it has been shown that chromosome breaks occurring as a result of APH treatment, if left unrepaired, can proceed to double strand breaks and activate the ATM pathway (Ozeri-Galai et al., 2008). It is therefore likely that due to the hypomorphic expression of Atr, the intrinsic and induced DNA damage in $Atr^{SckN/SckN}$ fibroblasts is not recognized by the dysfunctional Atr and Chk1 pathway and thus proceeds to double strand DNA breaks that activate the Atm and Chk2 pathway.

By using a single dose of UV irradiation, others have shown that human cells deficient in ATR expression do not show the expected decrease in mitotic index as is seen in wild type cells (Alderton et al., 2004) due to failure to activate of the G2/M checkpoint and cell cycle arrest after UV treatment. Our observation of a decreased

mitotic index in $Atr^{SckN/SckN}$ fibroblasts as compared to $Atr^{+/+}$ fibroblasts is most likely indicative of activation of the Atr checkpoint and delayed Atr checkpoint response.

Although we found a low level of the Atr splice variant lacking exon 9 in normal mouse fibroblasts from C57BL6 mice analogous to that found in humans, the point mutation alone that is the basis for SCKL1 in humans does not result in a similar increase in mis-splicing or reduction in Atr in our mice. There are several possibilities for this result with the most likely explanation being broader sequence differences between mouse and human ATR. Although the sequence of both mouse and human ATR is conserved at the 2101 A→G SCKL1 mutation site and surrounding nucleotides, the whole of exon 9 is only 90.8% similar and surrounding introns are highly divergent. The nature of the SCKL1 mutation suggests that this site functions as either an exon splice enhancer or an exon splice silencer with the 2101 A→G mutation either disrupting or enhancing its function but broader sequences likely modulate this response (O'Driscoll et al., 2003). While the exact function of this site is not fully known, it is likely that sequence differences between exon 9 or surrounding introns of human and mouse ATR account for the lack of increased mis-splicing found in our $Atr^{Sckl1/Sckl1}$ mice and that substitution of a larger region of human ATR may result in the desired altered splicing.

The loss or mutation of genes central to sensing and repairing DNA damage is an enabling characteristic in tumorigenesis (Hanahan & Weinberg, 2000). In addition, mutation of a number of DNA damage response genes has been found to be associated with developmental abnormalities and a premature ageing phenotype. Because complete loss of Atr is lethal and heterozygous mice have a limited phenotype the mice described

here with hypomorphic Atr expression serve as valuable reagents for studies of the biological effects of ATR deficiency *in vivo*.

Materials and methods

Generation of mice with the Atr^{Sckl1} and Atr^{SckN} alleles

The RPCI-22 Mouse BAC library was screened by the Hospital for Sick Children in Toronto, Ontario. Subcloning of the 5.6 kb genomic fragment of mouse Atr into the backbone of the targeting vector Bluescript KS+ was performed using the restriction enzyme SpeI. The Neo cassette was excised from the pNeo ZTK2 plasmid using XhoI and SalI and was cloned into the targeting vector via a SpeI site. Finally, site directed mutagenesis using primers containing the desired mutation and digestion with DpnI was performed to select for the newly replicated plasmid vector containing the A→G point mutation. The final targeting vector was linearized using a single KpnI site prior to electroporation into mouse W4 ES cells (Auerbach et al., 2000). Transfected ES cells were selected for incorporation of the vector with G418 selection and cloned as described (Hughes et al., 2007). Screening of ES cell clones for homologous recombination of the targeting vector was performed using PCR and Southern blot analysis for both the 3' and 5' ends of the construct.

Screening of ES cell clones and production of mice

480 ES cell clones were screened in duplicate by PCR analysis for correct incorporation of the targeting vector with primers unique to a correct homologous recombination event and flanking the 5' border of the vector. All ES cell clones that

were positive by PCR analysis were subsequently analyzed by both 5' and 3' Southern blots. Briefly, total genomic DNA was extracted from the ES cell clones and 5μg/clone was digested with the BglII restriction enzyme overnight. Digests were separated on a 0.8% agarose gel at 30 volts overnight and transferred in 20X SSC to a hybond N+ membrane (Amersham Biosciences). Membranes were UV crosslinked and probed with the indicated probes (Figure 1a). Positive ES cell clones were then expanded and chromosomes were counted to confirm euploid numbers.

ES cells were injected into C57BL/6J X (C57BL/6J X DBA/2) blastocysts to produce chimeric mice. Chimeras then mated to C57BL/6J mice to achieve germline transmission of the Atr^{SckN} allele. Mice that had inherited the targeted allele were mated with B6.Cg-Tg(ACTFLPe)9205Dym/J mice to delete the FRT flanked *neo* cassette (Rodriguez et al., 2000).

Cell Culture and CFS analysis

Primary mouse tail fibroblasts were grown in RPMI medium containing 15% fetal bovine serum (FBS) Non-essential amino acids, L-Glutamine, and penicillin/ streptomycin (Gibco). Cells were incubated at 37°C and 5% CO_2.

In order to induce chromosome gaps and breaks and common fragile site expression, cells were treated with 0.4μM APH for 24 hours prior to metaphase harvest. Metaphases were harvested by treating the cells with 50ng/ml Colcemid for 45 minutes. Cells were then placed in 0.075 M KCl at 37°C for 18 minutes and fixed in Carnoy fixative (3:1 methanol:glacial acetic acid) for two hours before replacing the fixative several times. Fixed cells were dropped on slides and aged at room temperature for 48

hours. Once the slides were aged fluorescent in situ hybridization (FISH) was performed as descried below.

Fluorescent in situ hybridization (FISH)

Established FISH protocols were followed (Wilke et al., 1996). Probes were labeled with biotin 14-dATP using the BioNic Translation kit (Invitrogen, Carlsbad, Ca.). Bound probe was detected using fluorescein isothiocynate (FITC) conjugated avidin-DCS (Vector Laboratories, Burlingame, CA.) followed by fluorescein conjugated anti-avidin immunoglobulin G (IgG). Chromosomes were then stained using 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Burlingame, CA.). Fluorescent signals were detected using a Ziss Axioscope epifluorescence microscope.

Mitotic index analysis

The mitotic index of the mouse fibroblast cell lines was determined by creating metaphase slides as described and then counting random fields of vision as seen through a 100X objective until a total of 2000 or more cells were scored. The total number of cells in metaphase was divided by the total number of cells scored to give the mitotic index for each cell type.

Reverse transcription polymerase chain reaction (RT-PCR)

Five micrograms of total RNA was reverse transcribed in a total volume of 10μl using oligo DT reverse primers. Superscript II Reverse Transcriptase (Invitrogen) was used according to the manufacturer's instructions for first strand cDNA generation. Two

microliters of the resulting reaction were used for each PCR reaction. PCR cycling conditions were as follows: 95° for five minutes, followed by 35 cycles of 95° for 30 seconds, 55° for 30 seconds, and 72° for 45 seconds, the final cycle consisted of 72° for 5 minutes.

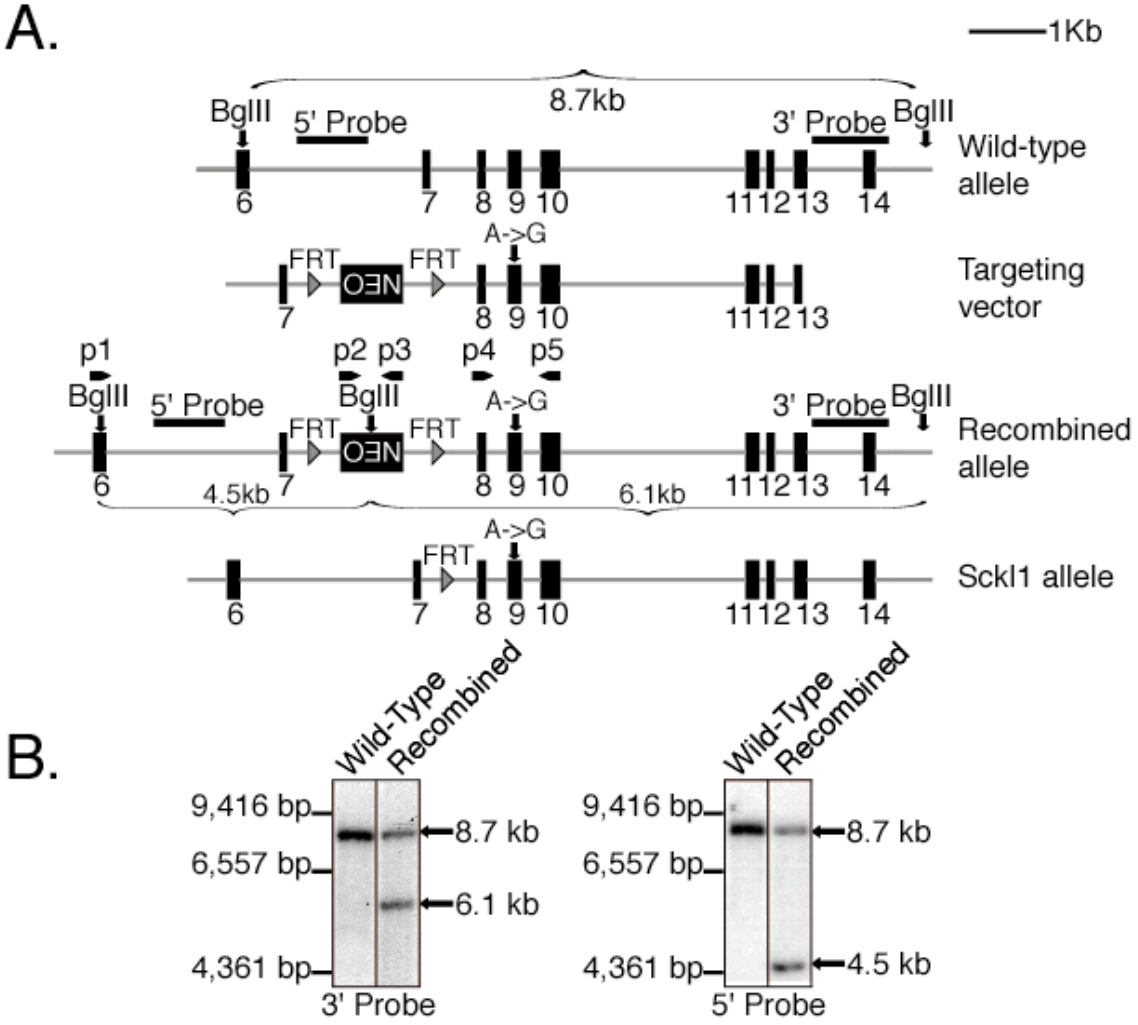
Phosphorylation assay

Primary mouse tail fibroblasts were plated at a density of $0.5X10^6$ cells/10cm dish. Plated cells were incubated for 24 hours prior to treatment with 0.4μM APH to induce checkpoint activation. Cell pellets were harvested at the indicated time points after APH addition, flash frozen and stored at -80°c until western blotting was performed.

Western blotting

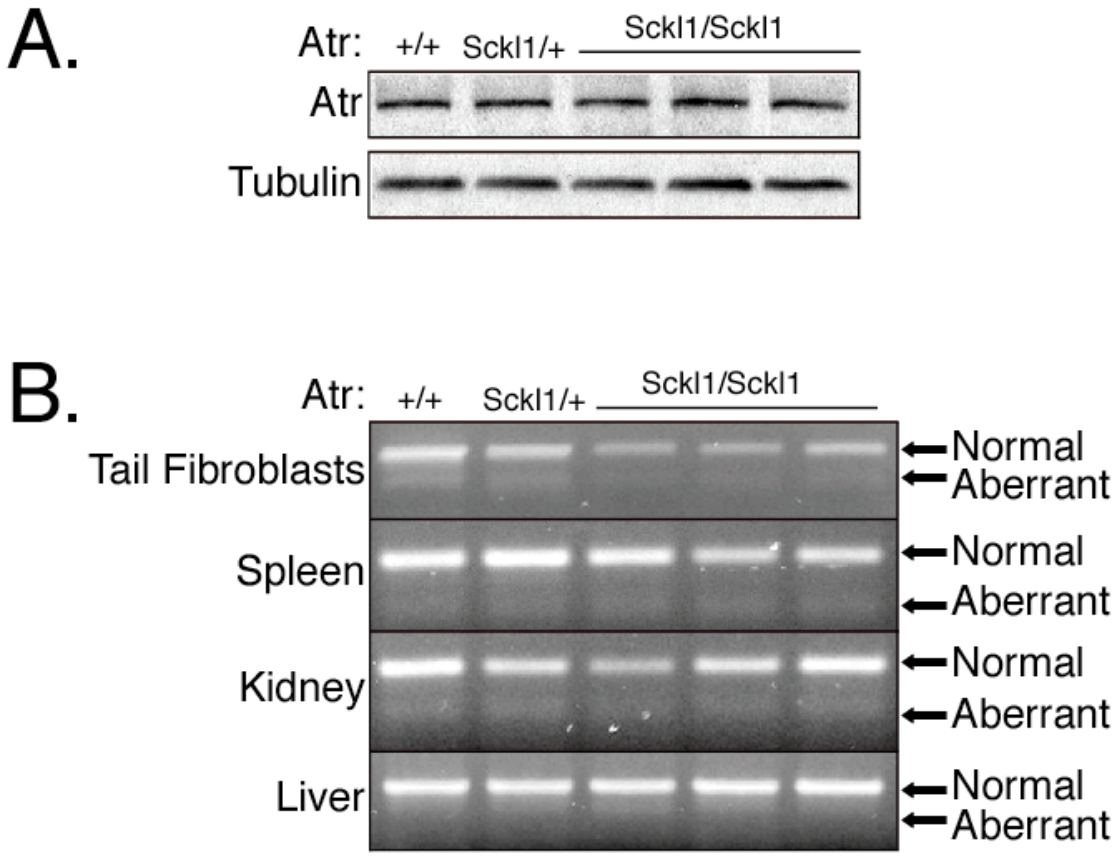
Whole cell extracts were created using a 1%SDS lysis buffer and resolved using a 3-8% Tris-Acetate gel (Invitrogen) for Atr or a 4-12% Bis-Tris (Invitrogen) gel for all other proteins examined. The protein was then transferred to a nitrocellulose membrain and blocked in a 5% milk dissolved in a tris-buffered saline and tween 20 (TBST) solution. The membrane was then probed using the following primary antibodies; mouse monoclonal anti-Chk1 1:1000 (cell signaling), rabbit polyclonal anti-Chk1 (Ser345) 1:1000 (cell signaling), rabbit polyclonal anti-H2AX 1:1000 (cell signaling), rabbit polyclonal anti-H2AX (Ser139) (Cell Signaling), rabbit polyclonal anti-Atr 1:40,000 (AbCam), or mouse monoclonal anti-Tubulin 1:5000 (NeoMarkers) diluted in a TBST solution and incubated overnight at 4°. Using appropriate HRP-conjugated secondary antibodies followed this procedure.

Figure 3-1 Schematic representation of the generation of the Atr^{Sckl1} allele



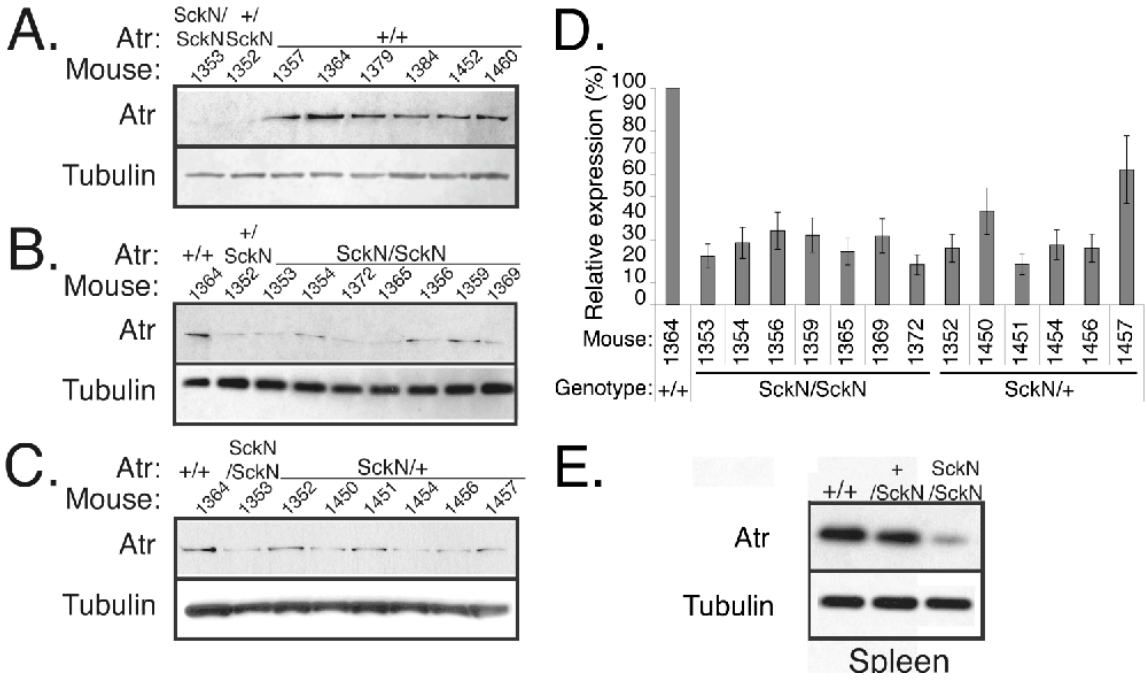
Generation of the knock in mouse model of Sckl1. **(A)** To scale schematic representation of the wild type Atr locus, targeting vector, recombined allele, and the final knock in allele post FLP expression. Primers used for splicing analysis are marked as P1-P5. The location of both the 5' and the 3' ~1 Kb Southern blot PCR probes is indicated. **(B)** Conformation of correct targeting of the vector to the mouse Atr locus in ES cells was confirmed by southern blot analysis of total genomic DNA digested overnight with the BglII enzyme. Shown here are both the 3' and 5' Southerns for both the wild type allele and the recombined allele.

Figure 3-2 Western and RT-PCR of various tissues from $Atr^{Sckl1/Sckl1}$ mice



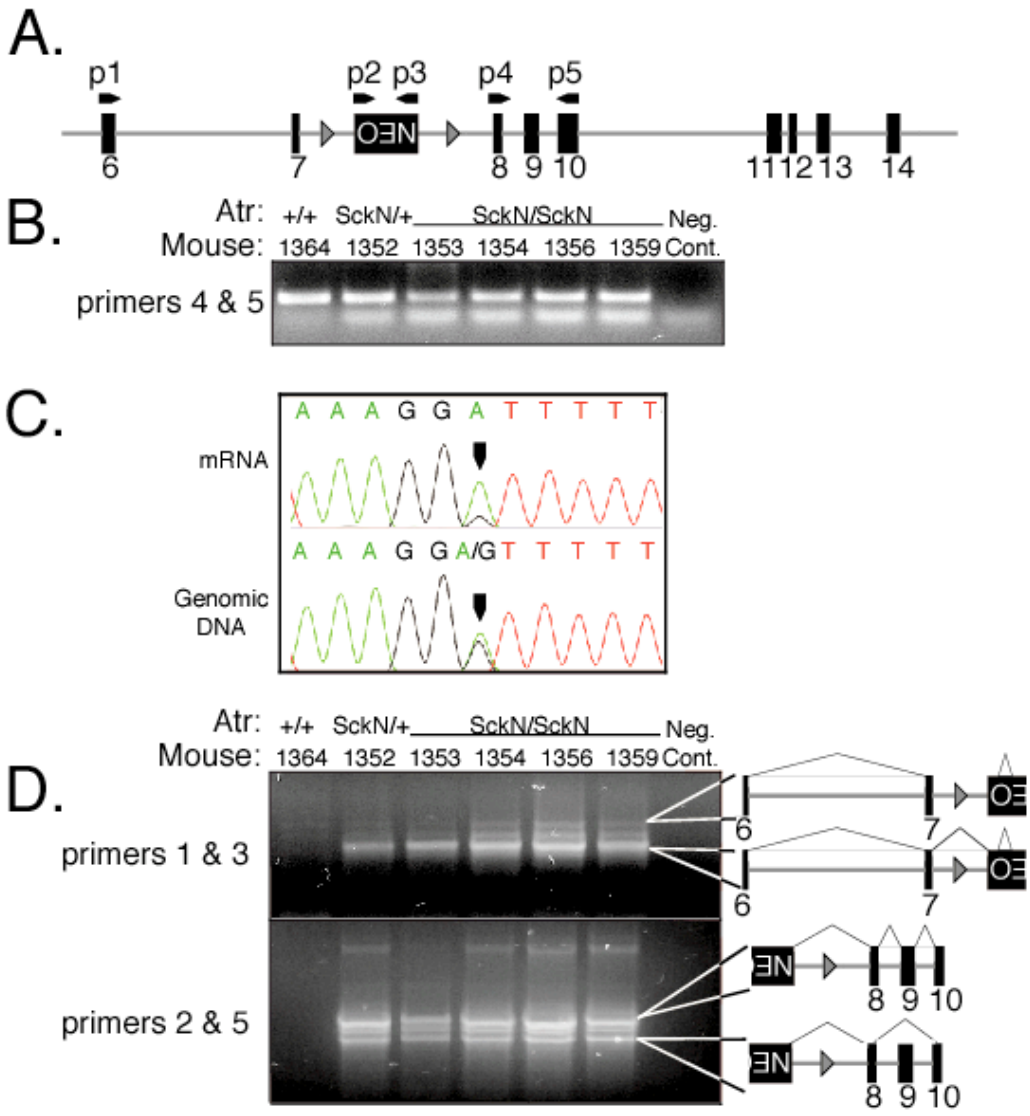
Mice backcrossed for five generations to the C57Bl/6J background containing the Sckl1 mutation in exon 9 of Atr do not show any alteration in protein or RNA expression. **(A)** Western blot preformed on cultured mouse tail fibroblast cell protein extracts and probed for Atr. Tubulin is used as a loading control for protein amounts. **(B)** Total RNA was extracted from the labeled mouse tissues and an RT-PCR from exon 8 to exon 10 was preformed. No increase in the amount of the aberrant splice product lacking exon 9 was found.

Figure 3-3 Western blot analysis of Atr protein in mice containing the Atr^{SckN} allele



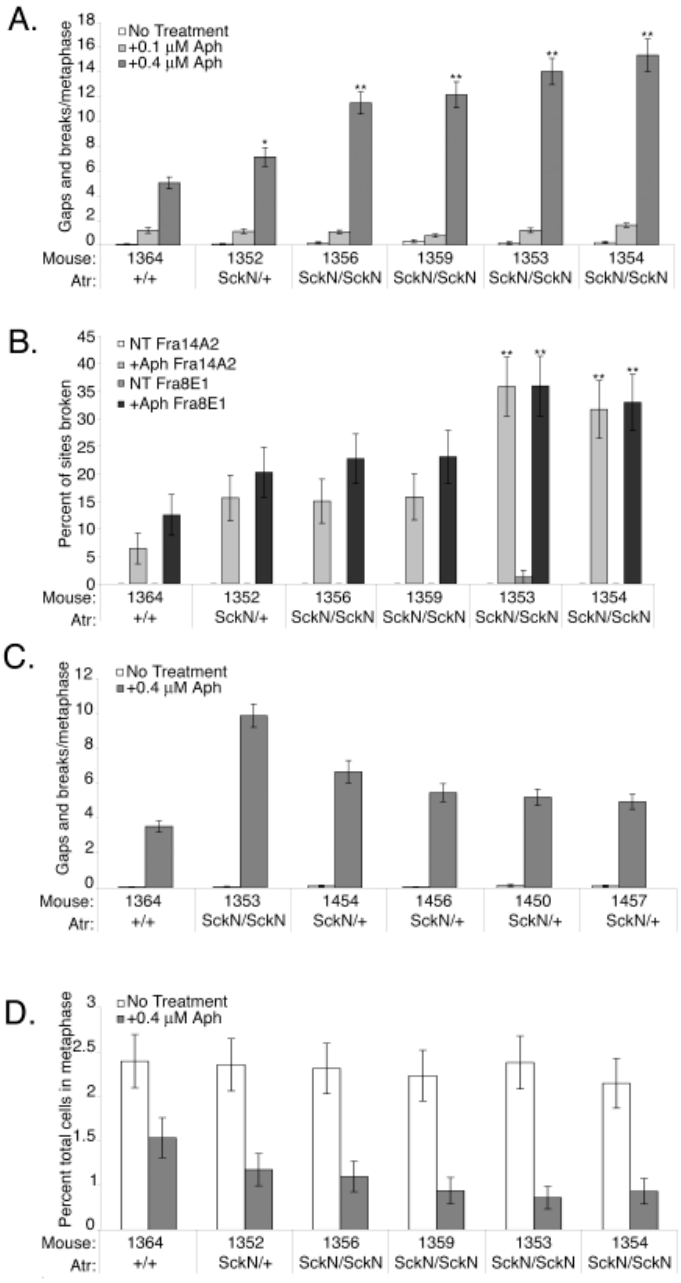
Western blot analysis of Atr protein taken from whole cell extracts of cultured primary mouse tail fibroblasts. **(A)** Six $Atr^{+/+}$ mice taken from three different litters and their respective controls. **(B)** Seven $Atr^{SckN/SckN}$ mice taken from three different litters and their respective controls. **(C)** Six $Atr^{SckN/+}$ mice taken from three different litters and their respective controls. (**D**) Graphical analysis comparing the real time PCR results for one $Atr^{+/+}$ seven $Atr^{SckN/SckN}$. **(E)** Atr protein levels as found in either brain or spleen tissue in one $Atr^{+/+}$, one $Atr^{SckN/+}$, and one $Atr^{SckN/SckN}$ mouse.

Figure 3-4 RT-PCR analysis of primary tail fibroblasts



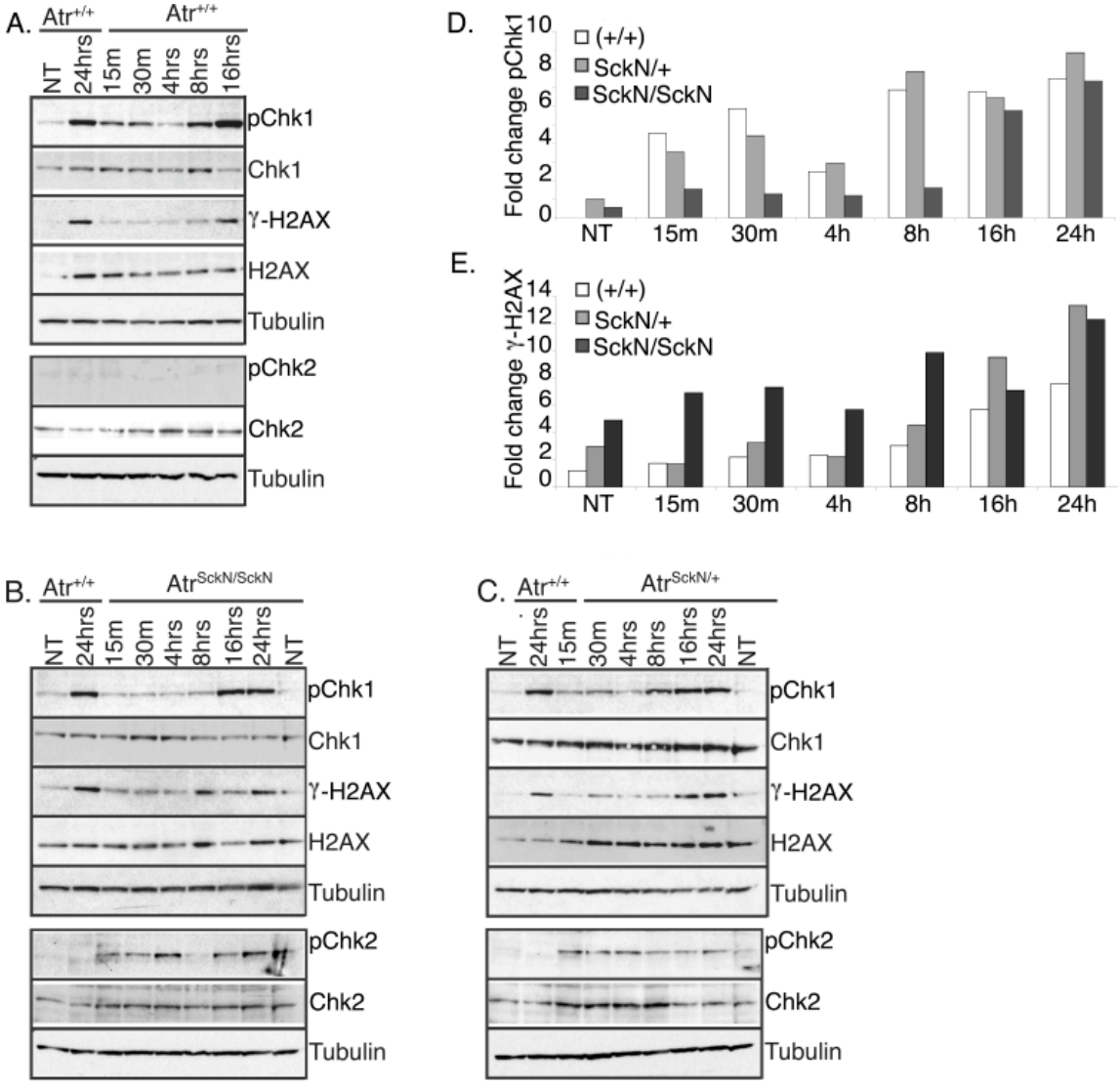
RT-PCR analysis of total RNA taken from primary mouse tail fibroblasts. **(A)** Schematic representation of a portion of the Atr^{SckN} allele and the respective location of the primers used in this analysis. **(B)** RT-PCR results using primers designed to amplify any products lacking exon 9. (**C**) Chromatogram of mRNA and genomic DNA form a heterozygous mouse (1352 $Atr^{SckN/+}$) showing an under representation of normal splice product from the targeted allele. **(D)** RT-PCR results using primers designed to capture any RNA products that contain either the 3' or the 5' end of the Neo cassette. On the right are schematic representations of the splice products that were captured and sequenced.

Figure 3-5 Metaphase analysis of primary tail fibroblasts



(A) Average total gaps and breaks per metaphase scored for a panel of homozygous mice. **(B)** Gaps and breaks as found at mouse common fragile sites as a percentage of total sites scored. **(C)** Average total gaps and breaks per metaphase scored for a panel of heterozygous mice. **(D)** Mitotic index for all primary mouse tail fibroblasts examined.

Figure 3-6 Western blot analysis of primary tail fibroblasts



Western blot analysis for Chk1, H2ax, and Chk2 phosphorylation after a given time of 0.4μM Aph treatment using protein extracts taken from primary mouse tail fibroblasts. **(A)** Western blot analysis of mouse 1364 ($Atr^{+/+}$) treated with 0.4μM Aph for the indicated times and controls. **(B)** Western blot analysis of mouse 1353 ($Atr^{SckN/SckN}$) treated with 0.4μM Aph for the indicated times and controls. **(C)** Western blot analysis of mouse 1352 ($Atr^{SckN/+}$) treated with 0.4μM Aph for the indicated times and controls. **(D-E)** Graphical representation of protein expression levels, based on band density, for Chk1 and H2ax phosphorylation.

References

Abraham R. T. (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev 15, 2177-96.

Alderton G. K., Joenje H., Varon R., Borglum A. D., Jeggo P. A., and O'Driscoll M. (2004) Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway. Hum Mol Genet 13, 3127-38.

Auerbach W., Dunmore J. H., Fairchild-Huntress V., Fang Q., Auerbach A. B., Huszar D., and Joyner A. L. (2000) Establishment and chimera analysis of 129/SvEv- and C57BL/6-derived mouse embryonic stem cell lines. Biotechniques 29, 1024-8, 1030, 1032.

Bakkenist C. J., and Kastan M. B. (2003) DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature 421, 499-506.

Bakkenist C. J., and Kastan M. B. (2004) Initiating cellular stress responses. Cell 118, 9-17.

Brown E. J., and Baltimore D. (2000) ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev 14, 397-402.

Burma S., Chen B. P., Murphy M., Kurimasa A., and Chen D. J. (2001) ATM phosphorylates histone H2AX in response to DNA double-strand breaks. J Biol Chem 276, 42462-7.

Casper A. M., Durkin S. G., Arlt M. F., and Glover T. W. (2004) Chromosomal instability at common fragile sites in Seckel syndrome. Am J Hum Genet 75, 654-60.

Casper A. M., Nghiem P., Arlt M. F., and Glover T. W. (2002) ATR regulates fragile site stability. Cell 111, 779-89.

Cortez D., Guntuku S., Qin J., and Elledge S. J. (2001) ATR and ATRIP: partners in checkpoint signaling. Science 294, 1713-6.

Fang Y., Tsao C. C., Goodman B. K., Furumai R., Tirado C. A., Abraham R. T., and Wang X. F. (2004) ATR functions as a gene dosage-dependent tumor suppressor on a mismatch repair-deficient background. Embo J 23, 3164-74.

Goodship J., Gill H., Carter J., Jackson A., Splitt M., and Wright M. (2000) Autozygosity mapping of a seckel syndrome locus to chromosome 3q22. 1-q24. Am J Hum Genet 67, 498-503.

Hanahan D., and Weinberg R. A. (2000) The hallmarks of cancer. Cell 100, 57-70.

Hughes E. D., Qu Y. Y., Genik S. J., Lyons R. H., Pacheco C. D., Lieberman A. P., Samuelson L. C., Nasonkin I. O., Camper S. A., Van Keuren M. L., and Saunders T. L. (2007) Genetic variation in C57BL/6 ES cell lines and genetic instability in the Bruce4 C57BL/6 ES cell line. Mamm Genome 18, 549-58.

Jazayeri A., Falck J., Lukas C., Bartek J., Smith G. C., Lukas J., and Jackson S. P. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. Nat Cell Biol 8, 37-45.

Lakin N. D., Hann B. C., and Jackson S. P. (1999) The ataxia-telangiectasia related protein ATR mediates DNA-dependent phosphorylation of p53. Oncogene 18, 3989-95.

Lee M. Y., Tan T. D., Feng A. C., and Liu M. C. (2006) Clinicopathological analysis of 598 malignant lymphomas in Taiwan: seven-year experience in a single institution. Am J Hematol 81, 568-75.

Levin S. I., and Meisler M. H. (2004) Floxed allele for conditional inactivation of the voltage-gated sodium channel Scn8a (NaV1.6). Genesis 39, 234-9.

Lewis K. A., Mullany S., Thomas B., Chien J., Loewen R., Shridhar V., and Cliby W. A. (2005) Heterozygous ATR mutations in mismatch repair-deficient cancer cells have functional significance. Cancer Res 65, 7091-5.

Liu Q., Guntuku S., Cui X. S., Matsuoka S., Cortez D., Tamai K., Luo G., Carattini-Rivera S., DeMayo F., Bradley A., Donehower L. A., and Elledge S. J. (2000) Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev 14, 1448-59.

Matsuoka S., Rotman G., Ogawa A., Shiloh Y., Tamai K., and Elledge S. J. (2000) Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. Proc Natl Acad Sci U S A 97, 10389-94.

Menoyo A., Alazzouzi H., Espin E., Armengol M., Yamamoto H., and Schwartz S., Jr. (2001) Somatic mutations in the DNA damage-response genes ATR and CHK1 in sporadic stomach tumors with microsatellite instability. Cancer Res 61, 7727-30.

Meyers E. N., Lewandoski M., and Martin G. R. (1998) An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. Nat Genet 18, 136-41.

Myers J. S., and Cortez D. (2006) Rapid activation of ATR by ionizing radiation requires ATM and Mre11. J Biol Chem 281, 9346-50.

Nagy A., Moens C., Ivanyi E., Pawling J., Gertsenstein M., Hadjantonakis A. K., Pirity M., and Rossant J. (1998) Dissecting the role of N-myc in development using a single targeting vector to generate a series of alleles. Curr Biol 8, 661-4.

O'Driscoll M., Dobyns W. B., van Hagen J. M., and Jeggo P. A. (2007) Cellular and clinical impact of haploinsufficiency for genes involved in ATR signaling. Am J Hum Genet 81, 77-86.

O'Driscoll M., Ruiz-Perez V. L., Woods C. G., Jeggo P. A., and Goodship J. A. (2003) A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet 33, 497-501.

Ozeri-Galai E., Schwartz M., Rahat A., and Kerem B. (2008) Interplay between ATM and ATR in the regulation of common fragile site stability. Oncogene 27, 2109-17.

Rodriguez C. I., Buchholz F., Galloway J., Sequerra R., Kasper J., Ayala R., Stewart A. F., and Dymecki S. M. (2000) High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet 25, 139-40.

Ruzankina Y., Pinzon-Guzman C., Asare A., Ong T., Pontano L., Cotsarelis G., Zediak V. P., Velez M., Bhandoola A., and Brown E. J. (2007) Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. Cell Stem Cell 1, 113-26.

Shiloh Y. (2006) The ATM-mediated DNA-damage response: taking shape. Trends Biochem Sci 31, 402-10.

Smith L., and La Thangue N. B. (2005) Signalling DNA damage by regulating p53 co-factor activity. Cell Cycle 4, 30-2.

Stiff T., Walker S. A., Cerosaletti K., Goodarzi A. A., Petermann E., Concannon P., O'Driscoll M., and Jeggo P. A. (2006) ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. Embo J 25, 5775-82.

Vassileva V., Millar A., Briollais L., Chapman W., and Bapat B. (2002) Genes involved in DNA repair are mutational targets in endometrial cancers with microsatellite instability. Cancer Res 62, 4095-9.

Ward I. M., and Chen J. (2001) Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. J Biol Chem 276, 47759-62.

Wilke C. M., Hall B. K., Hoge A., Paradee W., Smith D. I., and Glover T. W. (1996) FRA3B extends over a broad region and contains a spontaneous HPV16

integration site: direct evidence for the coincidence of viral integration sites and fragile sites. Hum Mol Genet 5, 187-95.

Zhao H., and Piwnica-Worms H. (2001) ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. Mol Cell Biol 21, 4129-39.

Zhou B. B., and Elledge S. J. (2000) The DNA damage response: putting checkpoints in perspective. Nature 408, 433-9.

Chapter IV

Discussion and Future Directions

Overview

Over the course of this thesis, I have sought to answer two major questions in the field of CFS biology. These questions are: "What makes common fragile sites fragile?", and "What is the relationship between common fragile site expression and cancer?" I have begun to examine the role the DNA sequence may play in the instability of CFSs. In addition, I have developed a novel mouse model that may serve as a reagent to allow future researchers to dissect the relationship between Atr expression, replication stress, increased CFS expression, and tumorigenesis.

The work presented in chapter II represents a step forward in our fundamental understanding of the causes of CFS instability. Although investigators in the CFS field have hypothesized for years that the sequence of these sites could play an important and causal role in their instability, until this study, no one had directly tested this hypothesis in human cells. Based on the experiments presented in chapter II of this thesis, we can now suggest that sequences contained in the FRA3B BACs used are sufficient to cause significant genomic instability at formally non-fragile ectopic loci. Furthermore, the instability caused by the integration of these CFS sequences is not reliant on late

replication timing. These data suggest, that while late replication timing may play a role in the fragility of endogenous CFSs, it is not necessary for the CFS-like instability found at these ectopic integrations. The question of what specific sequences are most important to this phenomenon has not been answered in full and will be discussed later in this chapter.

The creation and characterization of a mouse model hypomorphic for Atr expression, as presented in chapter III of this thesis, provides a potentially important tool for identifying the role of Atr, the cellular response to replication stress, and CFS instability in tumorigenesis. However, the most important experiment, that of showing a change in the timing or frequency of tumor formation in homozygous $Atr^{SckN/SckN}$ mice, has not been performed at this time. Many genes in the Atr pathway, such as BRCA1, CHK1, and genes with similar functions to ATR, such as ATM and P53, are all strongly linked to tumorigenesis. These findings in conjunction with the available literature on Atr and tumor formation in mice, as previously discussed (Brown and Baltimore, 2000; Fang et al., 2004), suggest that at a low enough expression of Atr, viable mice should have an increase in tumorigenesis.

The major remaining question is, can we reduce the levels of functional Atr low enough to still have a viable mouse while inducing an increase in tumorigenesis? It is estimated that our current mouse model has up to an 82% reduction in the levels of functional Atr protein. At this time, it appears that this level of reduction is not sufficient to cause these mice to exhibit some of the features associated with SCKL1, such as microcephaly and dwarfism. Nor has it caused the obvious formation of spontaneous tumors by ~10 months of age.

Others have observed age related phenotypes in adult Cre/Lox Atr knock out mice (Ruzankina et al., 2007). In these mice, tissues with continuous cell proliferation and no ATR expression displayed acute cellular loss. This loss was associated with a reduction in tissue specific stem and progenitor cells, and an exhaustion of tissue renewal and homeostatic capacity. These studies suggest that reduced regenerative capacity in adults via deletion of ATR is sufficient to cause the premature appearance of age related phenotypes. While we have not observed age related phenotypes in our mice, the phenomenon of selective loss of cells with low Atr expression may help to explain the variation in Atr expression seen between individuals in our studies.

Despite these and many previous advances in our understanding of CFS biology, there are a number of fundamental questions that remain unanswered. We do not know what specific sequences are important for the instability of CFSs, nor do we know whether late replication is causal or merely coincidental to the instability of endogenous CFS loci. In addition, other potentially more fundamental questions of CFS biology remain unanswered. What is the nature of a metaphase gap and/or break? What are the boundaries of a CFS? These questions and others will be addressed in the remainder of this chapter.

Nature of the chromosome "break"

In is not understood what a cytogenetically broken or "expressed" CFS represents. The aberrations seen on metaphase chromosomes that we define as expressed CFSs have been described by various groups as gaps, breaks, restrictions, constrictions, and/or poor staining (Durkin and Glover, 2007). Associated with these labels, there are a variety of

different visible phenomena commonly grouped together as expressed CFSs (Figure 4-1). What we see through the microscope may correlate to single strand DNA breaks, double strand DNA breaks, uncondensed chromatin, unreplicated DNA, or perhaps other unknown phenomena.

To fully understand the nature of CFSs and their response to APH-induced replication stress, the comprehension of what an expressed CFSs represents is important. As discussed in chapter I of this thesis, there is evidence to suggest that CFS expression begins as unreplicated single strand DNA regions that are processed as single strand DNA breaks, coated by RPA, and recognized by the ATR pathway. If a single strand DNA break is left unrepaired or is improperly repaired, it can proceed to a double strand DNA break and is then recognized primarily by the ATM pathway (Ozeri-Galai et al., 2008). Clearly some chromosomes are visibly completely broken and physically separate at the CFS. This suggests that double strand DNA breaks are indeed one of the phenomena that we define as an expressed CFS. Other expressed CFSs appear to be single chromatid breaks or pinches in the chromosome (Figure 4-1). Whether or not these phenomena represent single strand DNA breaks, double strand DNA breaks, decondensed, or unreplicated DNA in not yet understood and needs further examination.

How different are CFSs from other non-fragile loci?

CFSs are defined as loci that are preferentially sensitive to various forms of replication stress, such as low dose APH treatment, forming cytogenetically visible aberrations as seen on metaphase chromosomes. It is not entirely understood what causes these loci to respond differently to replication stress than other non-fragile loci. While

differences such as sequence content and replication timing have been implicated as being causal to the instability of CFSs, on close examination, it is unclear how different CFSs truly are from non-fragile sites in regards to these factors. This is exemplified by the findings presented in chapter II of this thesis; wherein an examination of sequence motifs present at two control and two FRA3B BACs did not uncover any noteworthy differences. In addition, our current working definition of CFSs complicates the elucidation of any differences that may exist between CFS and non-CFS loci.

At the extremes, FRA3B and FRA16D account for 25-34% of all gaps and breaks found after low dose APH treatment in human lymphocytes (Glover et al., 1984). These numbers decline drastically when one considers more than these two sites. Currently, more than 70 loci are defined as CFSs, yet 85% of all gaps and breaks found after low dose APH treatment in human lymphocytes occur at just 20 of these sites. This means that more than 50 sites, defined as CFSs, account for only 15% of all such gaps and breaks. These data suggest there is a large difference in the fragility of the top 20 and the bottom 50 CFSs and further suggest that the mechanisms causing fragility at the most fragile CFSs may be different from those found at the least fragile. This has made a global examination of possible differences between CFSs and non-CFSs difficult. Creating a more precise definition of CFSs or separating CFSs into subcategories based on their overall instability, could help to elucidate underlying mechanisms and contributory factors that would otherwise be obfuscated.

What are the boundaries of a CFS?

Complicating the discussion of what makes CFSs different from other non-fragile loci is the issue of defining distinct borders for specific CFSs. Cloned CFSs are commonly mapped by repeated FISH experiments with multiple tiled DNA probes demarcating whether a single probe had most breaks crossing, distal, or proximal, to the FISH signal. This has left the boundaries and the center of fragility for even the cloned CFSs somewhat in question. Furthermore, many CFSs are not as well defined as those that have been cloned in this manner and are only localized to a chromosome band that may be tens of megabases in size. An examination of the literature reveals that some groups have defined CFSs as a relatively limited region around the presumed center of the site (1-3Mb) (Boldog et al., 1997; Mishmar et al., 1998; Ried et al., 2000; Shiraishi et al., 2001; Arlt et al., 2002; Limongi et al., 2003; Ferber et al., 2004). Whereas, others have a looser definition of the boundaries of CFSs, stating that any sequence contained in the chromosome band where the CFS is located is part of the site itself (Helmrich et al., 2006; Tsantoulis et al., 2007). These differences in the definition of CFS boundaries have confounded even the most basic of facts. Some groups claim CFSs are AT rich and others state the opposite. For these reasons, the studies presented in this thesis have focused on the sequences present at the very center of the most commonly expressed CFS FRA3B in an effort to maximize any sequence differences that may be present.

What is the role of late replication and other factors at CFSs?

It is well established that late replication timing is associated with many CFSs and treatment with low dose APH causes both the expression of CFSs and pushes replication

timing even later. However, it is not currently established that this phenomenon is causal to the instability of these sites. The experiments described in Chapter II show that late replication does not appear to be necessary for the fragility of our ectopic FRA3B BAC integrations. Although it appears that late replication is not necessary for the instability of these ectopic sites, this finding does not eliminate the possibility that late replication could play a role in the instability of endogenous CFSs. Besides the association of late replication with CFSs, Wang et al., (Wang el al., 1999) found that there was an allele specific relationship between late replication and overall fragility at the FRA3B locus. Furthermore, the greater the dose of APH, the later replication is pushed for the genome and the more loci become "fragile" up to the point of complete replication halt. These data suggest that late replication may be at least partially causal to the instability of endogenous CFSs and that if replication of a given locus is pushed late enough, this may be sufficient to cause almost any loci to become unstable. There is evidence that additional factors such as improper catenation and transcription at the time of replication may also be contributory to the instability of CFSs.

Etopiside is an inhibitor of the enzyme topoisomerase II and functions to stabilize double strand DNA breaks that are caused during the decatenation process preventing reannealing of the broken ends. In a recent series of unpublished experiments, it has been shown that the addition of low dose etopiside is capable of inducing gaps and breaks at CFSs (Arlt et. al unpublished data). These studies suggest that expressed CFSs could be a result of improperly catenated DNA present at M phase of the cell cycle. It was also shown in these studies that in cell clones containing small copy number FRA3B BAC integrations, which are fragile but not late replicating, the addition of etopiside was not

sufficient to induce gaps and breaks at these sites. Therefore, etopiside appears to induce gaps and breaks at CFSs in a late replication dependent manner. This finding suggests that improper catenation of DNA present late in the cell cycle may be casual to CFS instability.

In addition to late replication timing and catenation, there is an association between CFSs and large transcriptionally active genes (Helmrich et al., 2006). It is hypothesized that large transcriptionally active genes could be transcribed at the same time as they are being replicated thus interfering with the correct replication of DNA at these sites and causing the expression of associated CFSs. This hypothesis is supported by the fact that the transcriptionally active CFS FRAXB is expressed on both homologous chromosomes in females, whereas the X-inactivated CFS FRAXC was only found expressed on the transcriptionally active X chromosome (Austin et al., 1992). Despite this evidence, it is not clear that transcriptionally active genes are necessary or contributory to CFS instability. Many CFSs are not associated with large transcriptionally active genes nor are many large transcriptionally active genes associated with CFSs.

What is the role of specific sequences at CFSs?

One of the leading hypotheses as to what makes a CFS different from the rest of the genome is the sequence of CFSs themselves. Chapter II indicates that transference of sequences taken from the CFS FRA3B is sufficient to cause instability at ectopic non-fragile loci. It has also been shown that deletions of large regions of FRA3B do not completely eliminate the fragility of this endogenous locus (Durkin et al., 2008).

However, evidence exists suggesting that the genomic sequence of CFSs may not be all that is necessary for the instability of endogenous CFSs. The issue of sequence contribution is complicated by the amount of sequence referenced in the previous statements. The FRA3B deletions examined did not eliminate what may be considered to be the entire CFS, thus some of the sequences contributing to the fragility of FRA3B may still remain. In support of this assertion, it was found that a tumor cell line containing a deletion of the entire CFS FRAXB locus had completely lost the fragility of FRAXB (Arlt et al., 2002). In addition, the BAC insertions examined in Chapter II did create a new ectopic fragile site but it remains unclear what specific sequences were contributory and how many copies of those sequences are truly necessary to cause this instability.

As measured by the phyloP program (http://compgen.bscb.cornell.edu/phast) the sequence of the FRA3B BACs are overall more highly conserved in mammalian evolution than that of the control BACs. While small regions of strong conservation are present in all of the BACs examined none of these regions correspond with flexibility peaks (Figure 4-2). This suggests that the specific flexibility peaks found in our BACs are not conserved features of CFSs even thought the instability of these sites is. This finding does not rule out the possible contribution of flexibility peaks to CFS instability but rather suggests that it is not necessary to conserve specific peaks to conserve fragility.

While flexibility peaks are not highly conserved short stretches (>100bp) of highly conserved sequence were found in both the FRA3B and control BACs (Figure 4-2). Searching for these small sequences showed that they are conserved features present at several different loci in of the rest of the genome. Other loci containing these small sequences were not associated with CFSs. Furthermore, these small conserved sequences

were present in the control as well as the FRA3B BACs, suggesting that these sequences likely have some evolutionarily conserved function that, alone, is not sufficient to induce CFS-like instability.

Based on comparison of the control and experimental BACs, it appears that, unlike rare fragile sites, no one sequence motif examined is responsible for the observed instability. A more likely hypothesis is that many different sequence motifs spread throughout a given CFS region can all independently contribute to the instability of the site. These sequence motifs, in combination with other factors such as late replication, likely together cause the overall instability of CFSs (Figure 4-3). This hypothesis is in accordance with the mouse human hybrid deletion findings (Durkin et al., 2008) and may help to explain why, in chapter II, it was found that our control BAC sequences are not radically different from our FRA3B BAC sequences. The leading motif present in the literature, that of flexibility peaks containing perfect AT repeat sequences as a cause of CFS fragility is not supported these findings. Thousands of loci through out the entire human genome, the majority of which are not fragile, contain flexibility peaks and perfect AT repeat sequences. These findings are discussed at length in chapter II and support the hypothesis that no one sequence motif is responsible for CFS instability. Despite the apparent role of sequence in the instability of CFSs, there appears to be no direct relationship between the presumed size of a given fragile site and the relative fragility of that site, as larger CFSs are not necessarily more fragile (McAvoy et al., 2007). This finding may appear to be contrary to what was discussed in chapter II.

In chapter II evidence was presented that the more FRA3B sequence present in a given area, the more fragile the area becomes. There are several possible explanations

for this discrepancy. A sequence from one CFS may be relatively more or less fragile than a sequence from another site when compared to each other. This scenario could arise based on the frequency of sequence motifs contributing to CFS instability in a given region. Said a different way, CFSs that are less fragile may contain more sequence that is also relatively less fragile. Alternatively, other factors such as late replication timing may be present and contributory at one site and not at another potentially increasing the fragility of a site without changing the physical size of the site. Before any definitive conclusions can be made, a more through examination of these factors is required.

Biological functions of CFSs

The very nature of CFSs argues against their evolutionary conservation. These sites are particularly prone to DNA damage after replication stress and are associated with known tumor suppressor genes whose loss could be an initiating event in tumorigenesis. This leads to the next point of inquiry: how can it be that these sites are conserved at both the sequence and expression level, given the potentially detrimental effects of the instability of these sites? Given the potentially negative impact of CFS expression on a cell, it is clear that these sites must perform some function vitally important to cell survival.

One possibility is that unknown regulatory sequences vital to the survival of the cell are present at CFSs. While such an explanation would account for the conservation of these sites on a sequence level, it does not explain the conservation of the instability of these sites, nor why other regions containing known regulatory sequences are not fragile. A more likely explanation is that CFSs act as a DNA stress monitoring mechanism in the

cell. This is commonly referred to as the "canary in the mineshaft" hypothesis.

This hypothesis posits that CFSs are the principal sites monitored by the DNA repair machinery, as they are the most sensitive to such damage and the cell fate decision is determined based on initial events at these sites. As previously discussed, it was found that in hybrid mouse cells containing a human chromosome 3, low dose APH was capable of causing a high frequency of large deletions at the FRA3B locus (Durkin et al., 2008). However, when a similar study was performed using normal human cells, no preferential association between deletions and CFS loci was found (Arlt et. al. unpublished data). This suggests that, in artificial conditions where there is no selection against deletions at FRA3B (such as is seen in mouse human hybrid cells) we observed a high frequency of deletions at this locus; while under normal cellular conditions, the preferential deletion of endogenous CFSs was not observed suggesting that in normal human cells there is active selection against deletions of FHIT and FRA3B. These data are in support of the hypothesis that cells preferentially monitor CFSs for damage and determine cell fate based on events at these sites.

There are many ways in which such a hypothesis could be further tested. One method involves crosslinking ATR or other proteins important to sensing DNA damage, shearing the DNA, and pulling down the resulting DNA fragments for microarray analysis, a process known as ChIP on Chip analysis. The results of such an experiment ought to determine whether a preferential association of proteins involved in DNA repair with CFS loci exists in normal untreated cells. If an association was found, it would suggest that the cell is preferentially monitoring CFS for DNA damage.

Future directions, factors influencing the instability of CFSs

Using the unique resources generated in the course of these studies, there are several possible lines of investigation that could be pursued at this time. A more traditional mapping approach could be used. The sequences contained in the two FRA3B BACs examined in chapter II, could be broken into smaller pieces which could be subsequently cloned into plasmid or retroviral vectors. These smaller sequence fragments could be integrated ectopically into the human genome and FISH experiments used to determine the stability of these novel integrants.

As a retrovirus inserts itself into the genome in a single copy, a retroviral vector used in this way would overcome the issue of multiple insertion copies. If one or more of the FRA3B retroviral integrations were fragile, we would have a much smaller region of sequence to examine to find our sequence of interest. On the other hand, if all of the FRA3B retroviral integrations were fragile, it would suggest, as hypothesized above, that no one sequence is solely responsible for the instability of CFSs. Rather, instability is determined based on the presence of many sequence spread through out the region. In addition, if this result were obtained, examination of smaller sequence fragments in additional clones using similar methods could help to identify specific sequences of importance. Finally, if none of the FRA3B retroviral integrants were fragile, it might suggest that the instability of CFSs has more to do with the amount and concentration of unstable sequence present in a given region than it has to do with the presence of one particularly unstable sequence. If this result were obtained, we could place the smaller FRA3B sequences that were found not to be fragile into plasmid vectors that would likely integrate ectopically in multiple copies, as the BACs did. This experiment would serve to

confirm or refute the hypothesis that the instability of CFSs is more dependent on the amount and concentration of unstable sequence than the presence of one particularly unstable sequence.

Based on work done by Durkin et al. (Durkin et al., 2008), we have seen that deleting large regions of sequence from an endogenous CFS can bring about a reduction in the overall instability of that site. While we have observed a reduction in fragility, no clone has shown a complete loss of instability This suggests that either some sequences capable of causing instability, or additional non-sequence factors unaffected by these deletions, such as late replication timing, are still present and causal to the remaining instability. The existing deletion clones that have a mild reduction in fragility could be subjected to a second dose of APH treatment resulting in clones containing larger deletions. These clones could then be tested using our standard FISH analysis for their response to replication stress. Ideally, we would be able to identify clones that have overlapping deletions, one retaining fragility, and one not. Such a result could lead to the identification of critical sequences necessary for CFS fragility.

The two sets of experiments proposed above, in addition to allowing for the examination of sequences important to CFS instability, could potentially help us to determine the extent to which late replication contributes to the instability of CFSs. Based on the replication timing experiments discussed in chapter II, it appears that ectopic integrations of FRA3B sequences do not retain their replication timing. This finding has helped us to separate the contribution of sequence and late replication in the role that they play at these ectopic loci. Examination, via FISH, of the clones generated in these two proposed experiments could potentially yield clones where the endogenous

FRA3B locus is not late replicating but is still fragile. This would further support the hypothesis that replication timing is not necessary for instability. Alternatively, we could generate several cell clones containing different ectopic integration of the same sequence where the integrated sequence is only unstable in specific replication conditions. Such a result would support the hypothesis that late replication is sufficient to cause CFS-like instability. As these clones will be generated to perform the above sets of experiments, subsequent examination of replication timing using these existing clones should not prove difficult.

There are several methods that could be employed to examine the potential role of transcription in CFS expression. Methods that could be used include: examination of the fragility of CFSs associated with genes in cell lines containing methylated and unmethalyaed genes; knock out of the gene promoters; overexpression of a specific transcriptional repressors; or the examination of cell types that do not normally express genes associated with a particular CFSs are all methods that could be used. If transcription during replication were causal to CFSs instability, then the overall expression of the CFSs examined should be reduced in cell lines where the associated gene is under expressed or not expressed at all.

Future directions, Atr mouse model

Ideally, a mouse that is fully backcrossed to the C57BL/6 line and contains one hypomorphic allele and one knock out allele, $Atr^{SckN/-}$, should be used to accurately examine the contribution of Atr deficiency in tumor formation. Such a cross would provide the ideal mouse model and could serve to reduce the total functional protein

levels from an estimated 18% of total normal protein to an estimated 9%. Although quantitative measurements were not performed in detail, $Atr^{SckN/SckN}$ mice do not appear to be overtly dwarfed, microcephallic, and as of 10 months of age, do not have any obvious spontaneous tumor formation or age related abnormalities. While it is possible that these particular abnormalities are not present due to unknown differences between humans and mice, it is more likely that the levels of Atr expression in $Atr^{SckN/SckN}$ mice are simply not low enough to cause these abnormalities. While it is likely that a cross to mice with a null allele could enhance the abnormalities described here, and possibly induce others not currently observed, there is a possibility that lowering the functional dose of Atr could cause these mice to become non-viable.

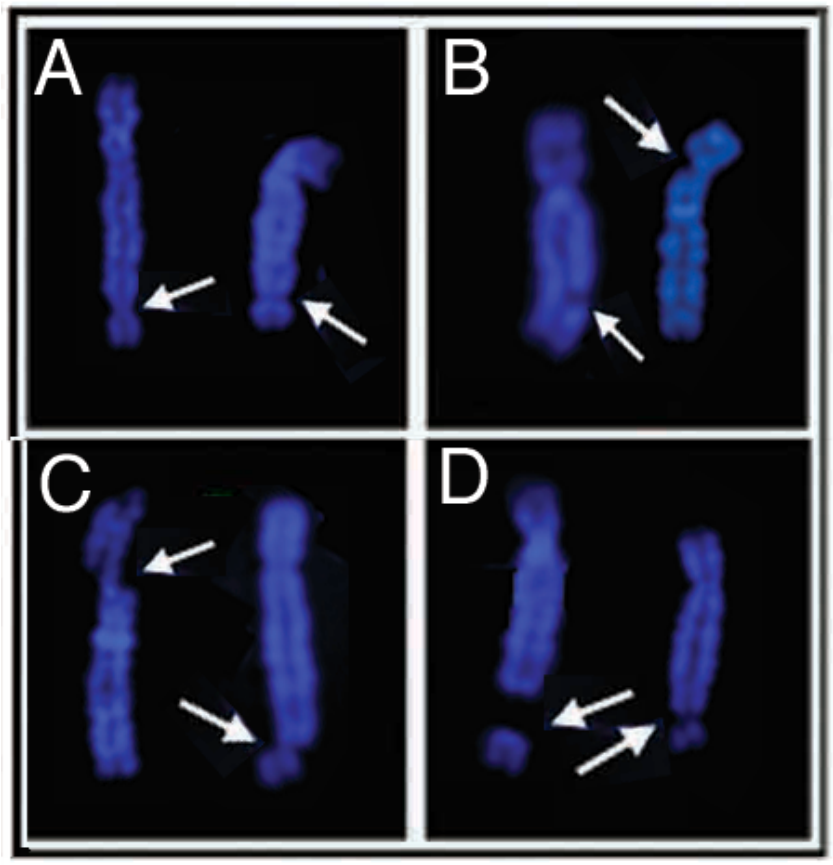
These mice would then either be left untreated or put on a folic acid deficient diet and/or irradiated with UV. Examination of tumor formation would be performed via necropsy after several months of treatment. If these low doses of Atr are not sufficient to induce an increase in tumorigenesis, our mice could be crossed to $Atm^{+/-}$ mice. We have shown in chapter III that the Atm pathway appears to be compensating, to some degree, for the deficiencies of the Atr pathway in our mice. A reduction in the functionality of the Atm pathway, in conjunction with a deficient Atr pathway, should serve to exacerbate any phenotype that these mice may have. In addition, Brown et al. showed a small two fold increase in late life tumor formation in $Atr^{+/-}/Atm^{+/-}$ compound heterozygotes as compared to $Atm^{+/-}$ or $Atr^{+/-}$ mice alone suggesting that our mice bred in such a cross should display earlier formation and higher levels of tumorigenesis than was found in this study (Brown and Baltimore, 2000). Alternatively, it has been shown that there is a strong increase in tumorigenesis when $Atr^{+/-}$ mice are placed on a mis-match repair

deficient background, Mlh1$^{-/-}$. We should get similar or stronger results if we were to cross our mice to a mis-match repair deficient line (Fang et al., 2004).

Conclusions

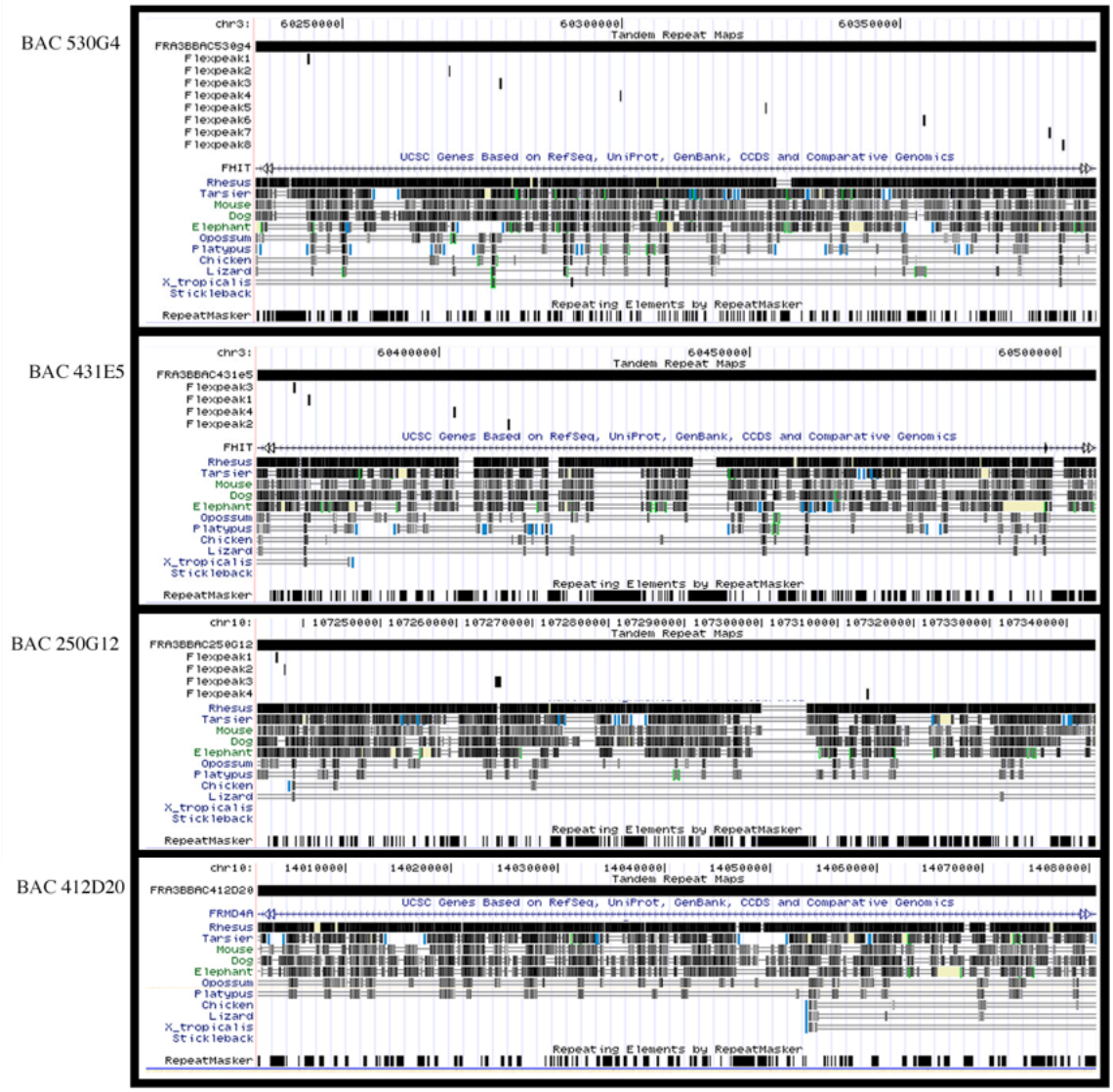
The ongoing study of CFS biology is of great importance because CFSs are a normal part of every human genome, and are strongly associated with tumorigenesis. While the work presented here represents small steps forward in our overall understanding of CFSs, these steps are both necessary and fundamental to our ongoing exploration of CFS biology. Although CFSs have been shown to be associated with late replication and with transcriptionally active genes, whether these factors are causal or coincidental remains scientifically unclear and needs further examination. These data presented and discussed in this thesis suggest a model whereby CFSs are the "perfect storm" of many different factors. Factors such as late replication, transcription during replication, and difficult to replicate through sequences, each of which are capable of contributing a small amount of genomic instability by themselves. When found together, these factors, in the right combination or concentration can cause loci to become "fragile" (Figure 4-3).

Figure 4-1 Example chromosomal phenomenon scored as gaps and breaks



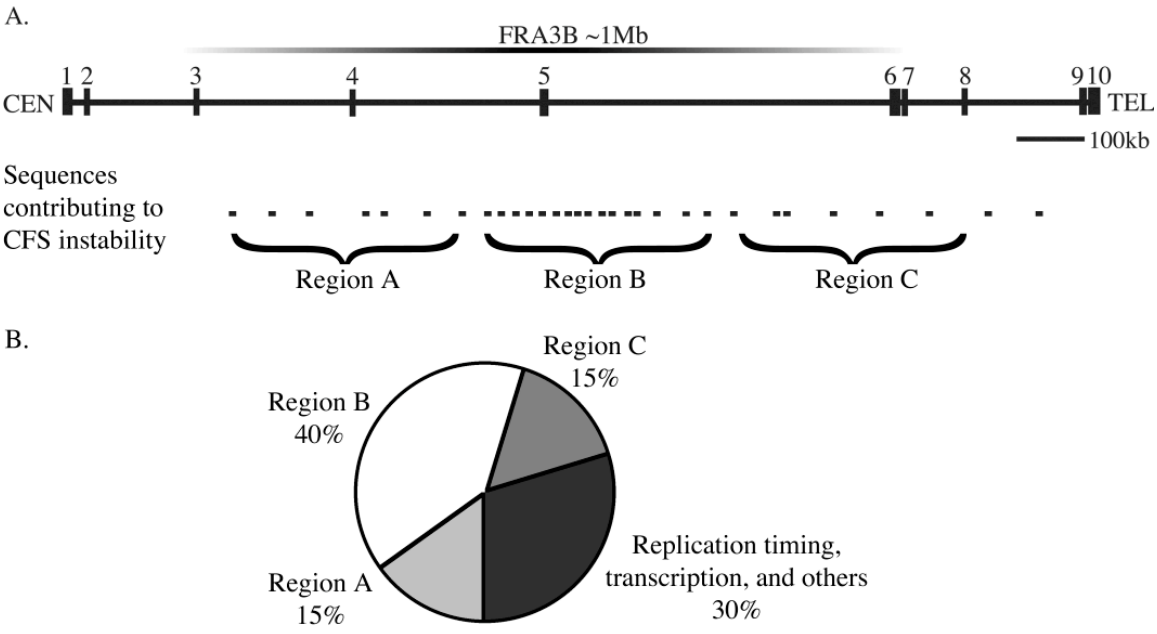
Examples of chromosomes with (**A**) constrictions, (**B**) gaps, (**C**) breaks, (**D**) and complete chromosome breaks. A FISH probe with signal on either side of the complete chromosome breaks identifies these chromosome pieces as belonging to the same chromosome and is not shown in these pictures.

Figure 4-2 Conservation of flexibility peaks in FRA3B and control BACs



Interspecies sequence conservation of BACs examined in chapter II. The top black bar represents the BAC examined. Smaller black boxes below BAC are the flexibility peaks found. While areas of strong conservation exist they are not correlated with flexibility peaks.

Figure 4-3 Model of factors contributing to CFS instability.



To scale physical model of the CFS FRA3B and factors that may influence its instability. (**A**) To scale representation of FRA3B/FHIT locus with hypothetical sequences contributing to the instability of the site shown. (**B**) Pie chart depicting the possible contribution of various regions and factors to the overall instability of the site as a whole.

References

Arlt MF, Miller DE, Beer DG, Glover TW. 2002. Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells. Genes Chromosomes Cancer 33:82-92.

Austin MJ, Collins JM, Corey LA, Nance WE, Neale MC, Schieken RM, Brown JA. 1992. Aphidicolin-inducible common fragile-site expression: results from a population survey of twins. Am J Hum Genet 50:76-83.

Boldog F, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H, Drabkin HA. 1997. Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193-203.

Brown EJ, Baltimore D. 2000. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev 14:397-402.

Durkin SG, Glover TW. 2007. Chromosome Fragile Sites. Annu Rev Genet.

Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW. 2008. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. Proc Natl Acad Sci U S A 105:246-251.

Fang Y, Tsao CC, Goodman BK, Furumai R, Tirado CA, Abraham RT, Wang XF. 2004. ATR functions as a gene dosage-dependent tumor suppressor on a mismatch repair-deficient background. Embo J 23:3164-3174.

Ferber MJ, Eilers P, Schuuring E, Fenton JA, Fleuren GJ, Kenter G, Szuhai K, Smith DI, Raap AK, Brink AA. 2004. Positioning of cervical carcinoma and Burkitt lymphoma translocation breakpoints with respect to the human papillomavirus integration cluster in FRA8C at 8q24.13. Cancer Genet Cytogenet 154:1-9.

Glover TW, Berger C, Coyle J, Echo B. 1984. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:136-142.

Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T. 2006. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. Genome Res 16:1222-1230.

Limongi MZ, Pelliccia F, Rocchi A. 2003. Characterization of the human common fragile site FRA2G. Genomics 81:93-97.

McAvoy S, Ganapathiraju SC, Ducharme-Smith AL, Pritchett JR, Kosari F, Perez DS, Zhu Y, James CD, Smith DI. 2007. Non-random inactivation of large common fragile site genes in different cancers. Cytogenet Genome Res 118:260-269.

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B. 1998. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci U S A 95:8141-8146.

Ozeri-Galai E, Schwartz M, Rahat A, Kerem B. 2008. Interplay between ATM and ATR in the regulation of common fragile site stability. Oncogene 27:2109-2117.

Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S, Nancarrow JK, Woollatt E, Kremmidiotis G, Gardner A, Venter D, Baker E, Richards RI. 2000. Common chromosomal fragile site FRA16D sequence: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells. Hum Mol Genet 9:1651-1663.

Ruzankina Y, Pinzon-Guzman C, Asare A, Ong T, Pontano L, Cotsarelis G, Zediak V, Velez M, Bhandoola A, and Brown E 2007. Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. Cell Stem Cell 1:113-26.

Shiraishi T, Druck T, Mimori K, Flomenberg J, Berk L, Alder H, Miller W, Huebner K, Croce CM. 2001. Sequence conservation at human and mouse orthologous common fragile regions, FRA3B/FHIT and Fra14A2/Fhit. Proc Natl Acad Sci U S A 98:5722-5727.

Tsantoulis PK, Kotsinas A, Sfikakis PP, Evangelou K, Sideridou M, Levy B, Mo L, Kittas C, Wu XR, Papavassiliou AG, Gorgoulis VG. 2007. Oncogene-induced replication stress preferentially targets common fragile sites in preneoplastic lesions. A genome-wide study. Oncogene.