GENETIC AND FUNCTIONAL ANALYSIS OF AUTOSOMAL LOCI THAT CONTROL CHROMOSOME-WIDE DNA REPLICATION TIMING AND MONOALLELIC GENE EXPRESSION

by

Nathan G. Donley

A DISSERTATION

Presented to the Department of Cell and Developmental Biology

and the Oregon Health & Science University School of Medicine

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

May 2013

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

_____

This is to certify that the PhD dissertation of

Nathan G. Donley

has been approved

_____

Mathew Thayer, Ph.D., Mentor

_____

Soren Impey, Ph.D., Committee Chair

_____

Melissa Wong, Ph.D., Committee Member

_____

R. Stephen Lloyd, Ph.D., Committee Member

_____

Richard Maurer, Ph.D., Committee Member

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

5-AZA	--	5-azacytidine
AI	--	autosomal inactivation
Amp	--	ampicillan
APC/C	--	anaphase-promoting complex/cyclosome
APRT	--	adenine phosphoribosyltransferase
ASAR6	--	asynchronous replication and autosomal RNA on chromosome 6
ATM	--	ataxia-telangiectasia mutated
ATR	--	ataxia-telangiectasia and RAD3 related
ATRIP	--	ATR-interacting protein
BAC	--	bacterial artificial chromosome
bp	--	base pairs
BrdU	--	bromodeoxyuridine
BUB3	--	budding uninhibited by benzimidazole 3
BUBR1	--	budding uninhibited by benzimidazole receptor 1
C	--	Celsius
Cam	--	chloramphenicol
Cdc6	--	cell division cycle 6

Cdc7	--	cell division cycle 7
Cdc20	--	cell division cycle 20
Cdc25	--	cell division cycle 25
Cdc45	--	cell division cycle 45
CDK	--	cyclin-dependent kinase
CDT1	--	chromatin licensing and DNA replication factor 1
CEP	--	centromeric probe
CHK1	--	checkpoint kinase-1
CHK2	--	checkpoint kinase-2
CIN	--	chromosome instability
CNA	--	DNA copy number alteration
CO-FISH	--	Chromosome Orientation-Fluorescence In Situ Hybridization
Cre	--	causes recombination
CSIN	--	chromosome structure instability
DAP	--	2,6-diaminopurine
dH_2O	--	distilled H_2O
DMC	--	delay in mitotic chromosome condensation
Dre	--	D6 recombinase
DRT	--	delay in replication timing
DSB	--	DNA double-strand break
ESC	--	embryonic stem cell
FISH	--	fluorescence *in situ* hybridization

FHL5	--	four and a half LIM domains 5
FHL5ost	--	FHL5 opposite strand transcript
FLP	--	flippase
Fosmid	--	F-factor Cosmid cloning vector
FoSTeS	--	fork stalling and template switching
Frt	--	flippase recognition target
FUT9	--	fucosyltransferase 9 (alpha (1,3) fucosyltransferase)
galK	--	galactokinase
GFP	--	green fluorescence protein
HSF	--	human skin fibroblast
Hyg	--	hygromycin B
Hyg^R	--	hygromycin B resistance
ICD	--	intrachromosomal deletions
ICT	--	interchromosomal translocation
INCENP	--	inner centromere protein
IR	--	ionizing radiation
I/S center	--	inactivation/stability center
Kan	--	kanamycin
Kb	--	kilobase
kV	--	kilovolt
LAM-PCR	--	linear amplification-mediated PCR
LB	--	Lysogeny Broth
LCR	--	locus control region

LINE-1	--	long interspersed nuclear element-1
LOH	--	loss of heterozygosity
LOJ	--	loss of junction
*lox*P	--	locus of cross-over (x-over) P1
M9	--	M9-minimal media
MAD2	--	mitotic-arrest deficient 2
MANEA	--	mannosidase, endo-alpha
Mb	--	megabase
MCC	--	mitotic checkpoint complex
MCM2-7	--	minichromosome maintenance complex
MCM10	--	minichromosome maintenance 10
mL	--	milliliter
mM	--	milliMolar
MMBIR	--	microhomology mediated break induced replication
MOI	--	multiplicity of infection
ncRNA	--	non-coding RNA
Neo	--	neomycin
ng	--	nanogram
NHEJ	--	non-homologous end joining
ODP	--	origin decision point
ORC	--	origin recognition complex
P-clone	--	parental cell clone
PBLs	--	primary blood lymphocytes

PCR	--	polymerase chain reaction
pre-IC	--	pre-initiation complex
pre-mRNA	--	pre-messenger RNA
pre-RC	--	pre-replicative complex
PW	--	Prader-Willi
rAAV	--	recombinant adeno-associated virus
R-clone	--	recombinant clone
ReTISH	--	Replication Timing-Specific Hybridization
rDNA	--	ribosomal DNA
RNA pol II	--	RNA polymerase II
rox	--	region of x-over
RPA	--	replication protein A
RT-PCR	--	reverse transcription-PCR
SAC	--	spindle assembly checkpoint
SCAPER	--	S phase cyclin A-associated protein in the ER
SINE	--	small interspersed nuclear element
Str	--	streptomycin
Tet	--	tetracycline
TDP	--	timing decision point
Tsix	--	antisense transcript to XIST
TSS	--	transcription start site
U	--	unit
UCSC	--	University of California, Santa Cruz

UFL1	--	E3 UFM1-protein ligase 1
μF	--	microFarad
μg	--	microgram
μM	--	microMolar
WCP	--	whole-chromosome paint
Xa	--	active X chromosome
XCI	--	X chromosome inactivation
Xi	--	inactive X chromosome
XIC	--	X chromosome inactivation center
Xist	--	X-inactivation specific transcript
YAC	--	yeast artificial chromosomes
YY1	--	yin yang 1

ACKNOWLEDGEMENTS

There are so many people I'd like to thank for my success in graduate school. The first is my mentor Matt Thayer. Matt, thank you so much for your help and support through these years. There were many times when I would get caught up in small details and start to doubt my project and myself, but you were always there to help me see the big picture. Most importantly, you were there when I needed your input and not when I didn't, the latter being the most important. You had confidence in my abilities and gave me the freedom to work things out on my own. You've taken some risks in your career to pursue a very difficult research topic and it is truly inspirational to see where this project started and where it is today. Thanks for being a great scientist that I can look up to.

This project would not have been possible without the cytogenetic expertise of Leslie Smith. Most of these cell based assays are very hard to master, but Leslie makes them look easy. Without her data, my thesis would not consist of very much, so this work is just as much Leslie's as it is mine. Also thanks for being a great lab companion.

I would also like to thank Eric Stoffregen. Eric took it upon himself to help me get acclimated to the lab and learn a bunch of new techniques. He also helped develop a lot of the techniques that I relied on in my research. But most of

all he was a good friend who was always willing to lend a helping hand.

Thanks to my thesis committee. Dr. Soren Impey, Dr. Stephen Lloyd, Dr. Melissa Wong, Dr. Maureen Hoatlin and Dr. Richard Maurer. I appreciate all of your input and your suggestions have really made this research project a better one. Thanks for taking time out of your already busy schedules to see me through to the end of my schooling.

Before I started school, I was a lab technician in Jan Christian's lab for two years. Jan could have just had me washing dishes and making stock solutions, but she allowed me to carry out experiments and eventually gave me my own research project. This really gave me confidence in my abilities as a scientist. Devorah Goldman, who was a post-doc in the lab, was also instrumental in fostering my interest in science. Both Jan and Devorah taught me how to work in a lab and develop ideas and interpret results. If I worked in any other lab I may not have decided to pursue a career in this field.

I'd like to thank my in-laws, Art and Linda. You have given Elise and I so much help since we got married. We were both students for the longest time and without your generosity we would have come on really hard times. Your farm was a great place for us to come and relax when life got really stressful. Thank you for everything.

I am so lucky to have such a great family. Mom, Dad, Alexis, Greer, Jay and Shani, you have all been such a wonderful family to have. I don't know where I would be in my life without every single one of you. Mom and Dad, you have been so supportive of everything I do. Your lives are an inspiration to me.

Alexis and Greer, I could not have asked for better sisters. It has been a great pleasure to see both of you grow up and mature into really wonderful people. Jay, you have been a great brother and your children are lucky to have such a dedicated father. Shani, I can't help but think that all the time we spent walking through rivers and swamps collecting specimens for your graduate thesis and nursing sick animals back to health at Wildcare had some impact on me becoming a scientist. Thanks for exposing me to nature and science at a young age and being a great sibling.

Last but not least, I want to thank my lovely wife Elise. We got married right before I started school and I could not think of a better person to have by my side for these last five and a half years. We have weathered some tough times together but we have always been as solid as a rock. I love you so much and thank you for always being there for me.

ABSTRACT

The majority of human cancers display a wide range of genetic abnormalities, including changes in chromosome number and structure, that are essentially nonexistent in normal tissue. While some of these abnormalities directly affect genes involved in cell growth and survival, the vast majority play an unknown role in carcinogenesis. It has been shown that certain chromosomal rearrangements present in tumor cells display a significant delay in replication timing (DRT) and a subsequent delay in mitotic chromosome condensation (DMC). Importantly, these chromosomes are very unstable and undergo frequent rearrangements, resulting in an overall increase in the rate of mutagenesis. The Thayer lab has proposed that DRT/DMC is caused by the disruption of a *cis*-acting 'inactivation/stability center' (I/S center) that functions to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of individual chromosomes. To date, two chromosomal loci, the *ASAR6* locus on chromosome 6 and the *Xist* locus on the X chromosome, have been identified as candidate I/S centers; however, the exact function of these loci and the existence of I/S centers on other chromosomes is currently unknown. In this thesis, I provide data to support the hypothesis that all

mammalian chromosomes contain I/S centers. Specifically, I show that *ASAR6*, like *Xist*, is able to dominantly disrupt replication timing when ectopically integrated into an autosome. Furthermore, ectopic *ASAR6* integration leads to ASAR6 RNA coating and chromosome-wide gene silencing of the integrated autosome. I also identify a new locus on chromosome 15 that, when deleted or disrupted, results in DRT/DMC and genomic instability. Therefore, three loci have now been identified that are necessary for the proper replication timing, mitotic chromosome condensation, and stability of their respective chromosomes. I finish with some preliminary data that opens up the avenue for a more in-depth genetic analysis of these loci.

CHAPTER ONE

Introduction

Eukaryotic cells are often exposed to various exogenous and endogenous stressors that can lead to DNA damage or mutagenesis. Many physical barriers, such as the nuclear envelope and chromatin, are present to prevent these agents from reaching and interacting with the cell's DNA. Eukaryotic cells have also evolved multiple mechanisms to efficiently repair DNA that has been damaged. Both the physical barriers and the DNA repair pathways generally keep the intrinsic mutation rate in quiescent cells very low. However, when a cell enters the cell cycle and begins growing and dividing, it presents a major problem for the maintenance of an error-free genomic landscape. During the cell cycle, DNA must be stripped of chromatin and both of the strands separated from one another in preparation for DNA replication. The replication machinery must replicate billions of base pairs with the possibility that errors will be generated, and the nuclear envelope must break down and reform again in each daughter cell. In response to these challenges, cells have developed very strict cell cycle controls to ensure the process of cell duplication occurs faithfully.

In Chapter One, I will discuss three processes that are tightly regulated to minimize mutagenesis during the cell cycle: DNA replication, chromosome condensation, and cell cycle checkpoints. I will also introduce a chromosomal phenotype called DRT/DMC and discuss the effect it has on these three cell cycle processes and the consequences this poses for the cell. A description of our current understanding of how DRT/DMC is generated will conclude this chapter.

Cell Cycle Regulation: Ensuring the Stability of the Genome

The cell cycle takes place in four stages: G1, S, G2 and M. In G1 phase, the cell grows in size and prepares for DNA replication. S phase is when DNA replication takes place followed by further growth and preparation for mitosis during the G2 phase. M phase, or mitosis, is when the newly replicated chromosomes condense and segregate. Cytokinesis also occurs, generating two daughter cells. Cell cycle checkpoints monitor the cell's progress through these four stages to ensure that this program is faithfully accomplished.

DNA Replication:

In bacteria, DNA replication is initiated at a single site along each chromosome [1]. The size and complexity of the eukaryotic genome, however, necessitates the usage of multiple initiation sites. In mammals, the number of replication initiation sites, or replication origins, has been estimated to be 30,000-50,000 [2]. A six-subunit origin recognition complex (ORC) binds each replication origin and remains bound throughout most of the cell cycle [3]. Although there is no consensus sequence for most eukaryotic origins, ORCs do not bind to random sites along each chromosome. Regions of DNA that are utilized as replication origins in one cell cycle are generally utilized as origins in subsequent cell cycles indicating that ORCs bind to regions of DNA with some specificity [4].

Interestingly, the cell begins preparation for DNA replication in telophase of the prior cell cycle [5]. This is when cell division cycle 6 (Cdc6) and chromatin

licensing and DNA replication factor 1 (CDT1) bind to ORCs and recruit the six-subunit minichromosome maintenance complex (MCM2-7). MCM2-7 has intrinsic helicase activity and is needed for efficient elongation during DNA synthesis [3]. The ORC/Cdc6/CDT1/MCM2-7 complex is referred to as the pre-replicative complex (pre-RC). Not all pre-RCs will go on to become active replication origins. In mid-G1, at the origin decision point (ODP), some pre-RCs are chosen to become initiators of DNA replication while others remain inactive throughout S-phase [6, 7]. Following cyclin-dependent kinase (CDK) and cell division cycle 7 (Cdc7) activation and modification of MCM2-7, minichromosome maintenance 10 (MCM10) and cell division cycle 45 (Cdc45) bind to a subset of the pre-RCs, transforming them into pre-initiation complexes (pre-ICs) [8]. Shortly after the pre-IC is formed, DNA polymerase α and primase are recruited to the origin and DNA synthesis begins in a bidirectional manner. DNA replication proceeds from each origin until the replication forks from two neighboring origins meet and the nascent DNA strands are ligated [3].

While DNA replication will initiate from most active origins within S-phase, the timing at which initiation takes place can vary widely between different origins. Adjacent origins tend to begin DNA replication at the same time resulting in large, synchronously replicating chromosomal domains called “replicon clusters” [9, 10]. Some replicon clusters will begin replication at the onset of S-phase while others will begin later during the middle or near the end of S-phase. This coordination of the temporal control of DNA replication is referred to as the replication-timing program. The replication-timing program is established shortly

after mitosis at a point in G1 phase, preceding the ODP, called the timing decision point (TDP) [11, 12]. The TDP is established coincidently with a global reorganization of chromatin into specified regions within the nucleus [12].

Not surprisingly, it turns out that 3-dimensional chromosome architecture in the nucleus is highly coordinated with DNA replication timing. In most, if not all, eukaryotic organisms, early replicating DNA resides in the interior of the nucleus while the later replicating regions remain at the nuclear and nucleolar periphery [11, 13, 14]. Molecular analyses have also revealed that late-replicating regions tend to cluster with other late-replicating regions in the nucleus and vice-versa [15]. Other associations have been observed with the genome sequence, structure and replication timing. For example, early-replicating regions tend to positively correlate with gene expression, G+C rich sequences, light-staining Giemsa bands, and active chromatin marks while late-replicating regions tend to be gene-poor, A+T rich, and have repressive chromatin marks [13, 16, 17]. It should be pointed out that while these correlations are significant they are not absolute, as some expressed genes and active chromatin marks reside in late-replicating regions [10].

In complement to these global replication-timing correlations, studies analyzing specific loci have uncovered *cis*-acting regulatory sequences that can directly influence the timing of origin replication. Early studies probing the effects of genome organization on replication timing found that local chromosomal rearrangements can cause changes in the replication timing of these regions [18, 19]. For example, a deletion in the area upstream of the human β-globin gene

called the locus control region (LCR) renders that gene late replicating in erythroid cells, a cell type in which it normally replicates early [20, 21]. Intriguingly, transgenes containing the LCR replicate early in erythroid cells regardless of where in the genome they integrate and can even change the replication timing of the surrounding region [22]. Telomeres have also been identified as *cis*-acting regions that influence replication timing. Moving a late replicating origin away from a telomere or an early replicating origin close to one results in a complete change of the replication timing of that origin [23, 24]. Recently, *cis*-acting regions have been identified that not only govern replication timing of localized regions but also of entire chromosomes. These regions are the focus of this thesis and will be discussed in more detail in subsequent sections.

The biological significance of having a replication-timing program in place is still currently unknown, however the existence of defective replication timing in many different diseases indicates that this is a vital cellular process. Whether it is present in inherited genetic diseases [25-27] or in cancer [28, 29], a replication-timing defect tends to result in abnormal chromosome condensation and genomic instability [30, 31].

Chromosome Condensation:

Following DNA replication, each chromosome undergoes extensive compaction, which facilitates the process of chromosome segregation and cell division. This compaction reduces the size of the DNA fiber 4-50 fold compared

to interphase chromatin and ensures that the segregating chromosomes can reach each spindle pole before cytokinesis occurs [32]. At the G2-M phase transition, the chromosome passenger complex, consisting of Aurora B, inner centromere protein (INCENP), and survivin, forms along the length of each chromosome [33]. The Aurora B kinase phosphorylates serine 10 on histone H3 throughout the entire length of the chromosome [34]. Concomitant with this phosphorylation is the binding of a large multi-subunit protein complex called condensin, which introduces supercoils into the DNA fiber [33, 35]. Each chromosome completes condensation by the end of prophase, and by metaphase the nuclear envelope breaks down and the centrosomes nucleate microtubule strands that bind to the kinetochores of each chromosome. Sister chromatids are pulled to opposite spindle poles and cytokinesis occurs generating two daughter cells with identical genomic content [36].

The seamless transition from DNA replication to chromosome condensation during the cell cycle indicates that the two may be functionally linked [37]. Early studies that fused mitotic cells to cells in varying stages of interphase (forcing the interphase chromosomes to undergo premature chromosome condensation) gave the first indication that DNA replication is a necessary precursor to chromosome condensation. When mitotic cells were fused to cells that were in G1 or G2 phase, the G2 phase chromosomes condensed to a much greater extent than the chromosomes that were in G1, indicating that progression through S phase is necessary to achieve the proper level of compaction [38]. In addition, when mitotic cells were fused to cells in S-

phase, the prematurely condensed chromosomes gave a hybrid appearance with regions that were tightly condensed (regions that had already replicated) intermixed with regions that were more extended (regions that had not yet replicated) [38, 39]. Additional evidence comes from genetic studies demonstrating that mutations in many different genes involved in the initiation or timing of DNA replication can lead to chromosome condensation defects [30, 40, 41]. This has led to a model in which the newly replicated DNA begins some level of compaction directly out of the replication complex and exists as a "chromosome territory" for the duration of S and G2 phases [37, 42]. The final stages of compaction then take place in prophase as mentioned above. Taking into account the sequential nature of replication and condensation as well as the functional relation, any defect in DNA replication that is not corrected will likely cause a subsequent defect in chromosome condensation.

Not surprisingly, defects in chromosome condensation tend to have the same consequence as defects in DNA replication: genomic instability. Deregulation of genes involved in chromosome condensation has been shown to cause genomic instability, abnormal mitosis, and may play a role in cancer development [43-45]. Fortunately, the cell has another line of defense if these cell cycle processes fail.

Cell Cycle Checkpoints:

The last, and perhaps most important, regulatory mechanism of cell cycle progression that will be discussed in this section is the cell cycle checkpoint. The

purpose of the cell cycle checkpoint is to halt or delay progression through the cell cycle in response to the abrogated function of a prior process [46]. It is generally thought that each phase of the cell cycle has its own checkpoint or checkpoints. The G1 phase checkpoint, referred to as the restriction point, functions to halt progression into S phase in the presence of DNA damage or unfavorable microenvironmental conditions [47]. The S phase has two separate checkpoints that ensure the genome is damage-free and completely replicated before entry into G2 or M phase. The G2 checkpoint will stop the cell from entering mitosis in the presence of DNA damage and the M phase checkpoint, also called the spindle assembly checkpoint (SAC), will keep the cell in metaphase until all of the chromosomes are aligned properly along the spindle [46, 48]. There are some redundancies between different checkpoints, but each checkpoint uses a distinct signaling cascade to affect cell cycle progression. Generally, the end result of each signaling cascade is to inhibit phase-specific cyclin/CDK function, which effectively stalls the cell at a specific phase of the cell cycle [49]. In this section, I will focus solely on the S phase checkpoints and the SAC because the other checkpoints are not relevant to the understanding of the phenotype that is the focus of this thesis.

There are two S phase checkpoints: the replication-dependent replication checkpoint (sometimes called the S-M checkpoint), and the replication-independent intra-S-phase checkpoint [48]. The replication checkpoint becomes activated in the presence of stalled replication forks that arise during DNA replication. Replication forks can become stalled by the inhibition of DNA

polymerase, encountering certain types of DNA damage and various other mechanisms [48]. Regardless of how they are formed, stalled replication forks lead to the generation of single-stranded DNA that is recognized by replication protein A (RPA). This brings the ataxia-telangiectasia and RAD3 related (ATR)-ATR-interacting protein (ATRIP) complex and other mediator proteins to the stalled fork site [50]. This complex recruits the effector protein checkpoint kinase-1 (CHK1) to the DNA site where it is phosphorylated on serine 345 [46, 51]. This activated CHK1 in turn phosphorylates cell division cycle 25 (Cdc25), which prevents S phase and M phase cyclin/CDK complexes from becoming activated [50]. The inhibition of these cyclin/CDK complexes prevents the initiation of DNA replication from origins that have not fired yet and also prevents the cell from moving into M phase [48, 50, 52]. This complex remains bound to the DNA to stabilize the stalled fork until the blockage is remedied [52, 53]. In effect, this checkpoint prevents the cell from entering mitosis while there is DNA damage or before DNA replication has completed. The other S phase checkpoint is the replication-independent intra-S-phase checkpoint. Being replication-independent, the site of signal transduction does not begin at an active replication fork but rather at the site of a double-strand break (DSB) outside of an actively replicating region. This checkpoint works to repair DSBs that are incurred during S phase. Following a DSB, local changes in the chromatin landscape activate the ataxia-telangiectasia mutated (ATM) protein [48]. This ATM activation induces checkpoint kinase-2 (CHK2) to phosphorylate Cdc25, targeting it for proteasomal degradation [54]. This inhibits S phase cyclin/CDK complexes and prevents

replication initiation from unfired origins until the damage is repaired, thus delaying progression through S phase [54, 55].

Once the cell has passed the S phase checkpoints and begins mitosis, another checkpoint must be passed to complete the cell cycle. The SAC monitors the progression from prometaphase to metaphase to ensure that the chromosomes are aligned properly along the metaphase plate before the onset of anaphase. At the beginning of mitosis the mitotic checkpoint complex (MCC), consisting of mitotic-arrest deficient 2 (MAD2), budding uninhibited by benzimidazole receptor 1 (BUBR1) and budding uninhibited by benzimidazole 3 (BUB3), forms at the kinetochore of each chromosome [56, 57]. The MCC then binds cell division cycle 20 (Cdc20), a component of the anaphase-promoting complex/cyclosome (APC/C) ubiquitin ligase [58]. This sequesters the APC/C and prevents it from ubiquitinating downstream target genes [59]. During prometaphase, microtubule fibers begin extending out from the centrosomes and attach to the kinetochore of each chromosome [60]. The MCC remains bound to the kinetochore until a bipolar attachment of microtubules is achieved on each sister chromosome pair [59]. This tension-generating bipolar attachment initiates the dissociation of the MCC-APC/C complex from the kinetochore. This dissociation allows APC/C to polyubiquitinate securin and cyclin B, which targets them for proteasomal degradation [57, 59]. Upon the degradation of these proteins, the cell proceeds into anaphase and telophase completing the cell cycle. This checkpoint ensures that each daughter cell receives the same chromosomal content.

With such an important role in maintaining genome integrity, it has been known for a long time that defects in the function of cell cycle checkpoints can lead to genomic instability. Most of our knowledge comes from genetic studies analyzing what happens when checkpoints no longer function properly. Deregulation of genes involved in the replication checkpoint or the SAC tends to impair the functioning of those checkpoints, which often leads to genomic instability [61-63]. However, an intact functional checkpoint can also be bypassed before DNA replication, DNA repair or kinetochore attachment is complete. This is a process called “checkpoint adaptation” and has been described in many organisms from yeast to humans [64, 65]. During checkpoint adaptation, a fully functional checkpoint response is initiated and maintained in reaction to DNA damage or chromosome misalignment, yet the cell is able to circumvent that response and proceed with the rest of the cell cycle before the damage or misalignment is resolved [66]. It has also been shown that checkpoint adaptation can lead to genomic instability [67]. So whether the cell has an impaired checkpoint response due to genetic deregulation of checkpoint genes or adapts to a fully functional checkpoint, the typical downstream effect is genomic instability.

DRT/DMC: A Breakdown of Cell Cycle Regulation

DNA Replication and Chromosome Condensation:

In 1967, Dr. Harald zur Hausen identified a novel chromosomal phenotype that was present in multiple different leukemia cell lines [68]. He observed that one or a few chromosomes in these cells exhibited a chromosome-wide delay in mitotic condensation and DNA replication. This delayed condensation and replication phenotype was subsequently observed in cells from patients with Bloom's syndrome, Fanconi's anemia and other developmental abnormalities [69-73]. These studies, however, only established that delayed chromosomes were present in some cells and gave no indication as to how or why these chromosomes were delayed.

In 2001, the Thayer lab published the first extensive characterization of chromosome-wide delay in replication and condensation [74]. It was found that the rhabdomyosarcoma cell line RH30 and the small-cell lung carcinoma cell line CRL-5845 each contain one or two chromosomes that display a chromosome-wide delay in DNA replication timing (DRT). DRT is characterized by a 2-3 hour delay in the initiation and completion of DNA replication along the entire chromosome (Fig. 1.1) [74]. Thus, any chromosome that displays DRT will begin replication 2-3 hours after the onset of S phase and complete replication in the G2 phase. In extreme cases, DNA replication in mitosis has been observed on chromosomes that display DRT [75]. This phenotype typically affects only one or two chromosomes in the cell while the other chromosomes replicate normally

[74]. Interestingly, chromosomes that display DRT maintain a banded pattern of DNA replication, indicating that a hierarchy of DNA replication timing remains intact on these chromosomes [74]. This suggests that the actual replication-timing program persists on DRT chromosomes with early, middle and late replicating regions still present, but the entire program just begins 2-3 hours late.

DRT is usually accompanied by a chromosome-wide delay in mitotic chromosome condensation (DMC) [74]. A chromosome that displays DMC exhibits at least two of the three following characteristics: it is at least twice as long as any other chromosome in the same mitotic spread, it is less than half as wide as any other chromosome in the same mitotic spread, and/or it contains a bend of greater than 180° [74]. DMC is likely a result of the DRT phenotype, as DNA replication in G2 or M phase would be expected to delay chromosome condensation. The DMC phenotype is only present on chromosomes that display DRT, so the other chromosomes in the cell condense normally following DNA replication [74]. DMC is associated with a delay in the recruitment of Aurora B kinase to the chromosome [75]. Consistent with a role in mitosis-specific histone phosphorylation, delayed Aurora B recruitment leads to a delay in the phosphorylation on serine 10 of histone H3 [74, 75]. Because histone H3 phosphorylation begins in late G2 phase [76], this lack of H3 phosphorylation indicates that DMC chromosomes are in a G2-state of condensation when the cell is in mitosis.

Chromosomes that display DRT/DMC are not present in normal blood lymphocytes and are typically seen only in cells that have incurred a high level of

DNA damage such as irradiated cells or cells from patients with chromosome breakage syndromes [69, 71, 74]. Thus, it would make sense that DNA damage is one possible cause of DRT/DMC. Consistent with this, DRT/DMC has only been observed on rearranged chromosomes such as translocated chromosomes, ring chromosomes and chromosomes that contain large deletions [73, 74, 77-79]. Furthermore, DNA damaging agents, such as ionizing radiation (IR), DNA recombinases and endonucleases, can induce DRT/DMC on normal chromosomes [77, 80, 81]. The exact cause of DRT/DMC will be discussed later but, as far as is known, some type of genomic rearrangement is necessary in *cis* to cause this phenotype.

DRT/DMC has been observed on nearly every human chromosome in many different human cell lines and primary cells [74, 77, 80, 81]. Mouse and hamster chromosomes can also display DRT/DMC [80, 82]. Therefore, it is likely that DRT/DMC can be induced on any mammalian chromosome and possibly other eukaryotic chromosomes as well.

Cell Cycle Checkpoints:

Strictly speaking, DRT/DMC should not cause any real problems in the cell. If the cell-cycle checkpoints are functioning properly then the cell should delay mitosis until DNA replication has been completed and delay anaphase until all chromosomes are aligned properly along the metaphase plate. This would, in effect, prevent any problems associated with delayed replication and condensation and have no real consequence for the cell other than an increase

in the time it takes to complete the cell cycle. However, it is known that this phenotype has significant consequences for the cell in the form of genomic instability (discussed in the following subsection).

It has been shown that cells with chromosome-wide replication delay experience checkpoint adaptation at both the DNA replication checkpoint and the SAC [75]. During S phase, CHK1 is phosphorylated on S345 and binds DRT/DMC chromosomes indicating that the ATR-CHK1 signaling cascade in the DNA replication checkpoint is intact [75]. Despite the replication checkpoint being triggered, DNA replication can still be detected in mitosis in these cells [75]. This suggests that at least some cells with DRT/DMC chromosomes can undergo checkpoint adaptation at the replication checkpoint, leading to DNA replication in mitosis. Furthermore, the MCC component MAD2 remains bound to DRT/DMC chromosomes at a time when normal chromosomes are MAD2 negative [75]. Regardless of SAC activation, defects in chromosome segregation are readily apparent in these cells [75]. This indicates that at least some cells with DRT/DMC chromosomes can also adapt to the SAC and undergo cytokinesis without proper chromosome alignment.

Genomic Instability:

DRT/DMC results in at least two distinct types of genomic instability. The first is chromosome instability (CIN), which is characterized by an increase in the rate at which cells gain or lose entire chromosomes [75, 80, 83]. Cells with DRT/DMC chromosomes tend to have chromosome number imbalances [75, 80].

Furthermore, irradiated cells that have DRT/DMC chromosomes tend to be hyperdiploid whereas those that lack DRT/DMC chromosomes are generally diploid [80]. Inducing DRT/DMC on a chromosome in karyotypically normal cells results in aneuploidy, with each cell containing a different number of chromosomes [75]. Thus, cells with DRT/DMC chromosomes display frequent gains or losses of entire chromosomes resulting in dramatic aneuploidy affecting the entire karyotype [75, 80]. In addition, cells containing DRT/DMC chromosomes have abnormal mitotic spindles, abnormal centrosome number, and an increased frequency of endoreduplication [75]. It is unclear how DRT/DMC on a single chromosome is causing these events, but these factors can certainly explain the CIN observed in cells with this phenotype.

The second type of instability observed in cells with DRT/DMC chromosomes is chromosome structure instability (CSIN), which is characterized by an increase in the rate that new chromosomal rearrangements occur [74, 77, 81]. DRT/DMC chromosomes participate in numerous translocation events with other chromosomes in the cell and translocation intermediates can also be seen in mitotic spreads [74]. In addition, it has been demonstrated that DRT/DMC increases the rate of secondary chromosomal rearrangement by 30-80 fold [77]. The DRT/DMC induced CSIN is not random, as most of the chromosome rearrangements occur on the delayed chromosome(s). Although the structural instability is primarily observed on the delayed chromosome, other chromosomes can participate in inter-chromosomal translocations with the delayed

chromosome, indicting that DRT/DMC on one chromosome can destabilize the structural integrity of all chromosomes within the cell [77].

There are currently two models for how DRT/DMC can cause CSIN: 1) DRT/DMC results in delayed mitotic spindle attachment leading to chromosome mis-segregation and the formation of micronuclei, which can lead to CSIN [84, 85] and/or 2) DRT/DMC results in checkpoint adaptation and the onset of mitotic chromosome condensation prior to the completion of DNA synthesis leading to stalled replication forks, multiple DSBs and DNA repair via error-prone mechanisms [85].

Unlike other mechanisms that cause genomic instability, DRT/DMC tends to be a transient phenomenon. The inherent instability of DRT/DMC chromosomes makes them prone to extreme fragmentation over relatively few cell divisions, which results in highly rearranged chromosomes that no longer display DRT/DMC [74]. This feature of DRT/DMC makes it an underappreciated, yet potentially important force driving mutagenesis in certain disease states.

DRT/DMC as a Mechanism for Genomic Instability in Cancer

Cancer develops when normal cells acquire genetic and epigenetic alterations that lead to uncontrolled growth and the ability to evade cell death. These genetic and epigenetic alterations are generally thought to drive carcinogenesis by deregulating key pathways that control cell growth and proliferation [86]. In nearly all cases, deregulation of a single gene is not sufficient to cause cancer, making it necessary for a cell to acquire extensive genetic and epigenetic deregulation to develop a neoplastic phenotype [87]. Genetic analysis of many tumor types has revealed that malignant cells typically contain a very large number of independent genetic alterations. A recent sampling of various tumor cell types has revealed more than 2,000 recurrent chromosomal aberrations [88, 89]. In addition to these recurrent aberrations, over 100,000 non-recurrent aberrations have been catalogued [90]. Studies examining DNA copy number alterations (CNAs) have determined that most cancers contain multiple CNAs [91, 92]. Other reports have also indicated that 10-25% of the genome has lost heterozygosity in the typical breast, colon, pancreas, and prostate cancer cell [93-96]. Furthermore, a study done on individual colon cancer cells has revealed an average of 11,000 genomic aberrations per cell, highlighting the extent to which these genetic changes occur [97].

To explain the sheer number of genetic and epigenetic alterations that are observed in cancer cells, current models suggest that the acquisition of genomic instability is an integral part of carcinogenesis [97-100]. Genetic alterations can

arise during cancer progression through normal cellular processes, mutagenesis and as a result of genomic instability. Mutagenesis refers to a process by which genetic changes occur, either spontaneously or as a consequence of exposure to mutagens, resulting in a change in the DNA sequence. Genomic instability, on the other hand, refers to an increase in the rate of mutagenesis per unit time. While normal cells have a very low intrinsic mutation rate, a cell that displays genomic instability has a higher mutation rate and an increased likelihood of accumulating the necessary genetic and epigenetic changes required for malignant growth. Genomic instability as a facilitator of carcinogenesis is an attractive model because it not only accounts for the genetic heterogeneity observed in many tumors [101], but also the large number of mutations that seemingly provide no growth advantage (passenger mutations) [97, 102].

Despite its importance in carcinogenesis, there are still gaps in our knowledge of what causes genomic instability. Historically, it was thought to be a direct result of the deregulation of *trans*-acting factors (molecules that are physically separate from DNA, such as proteins). For instance, it has been shown that deregulation of p53 and Mdm2 can cause genomic instability in mammalian cells [61, 62]. Deregulation of numerous other proteins involved in cell cycle checkpoint control [63, 103], centrosome function [104], and DNA replication [105, 106] can also result in genomic instability. Additionally, over 130 different proteins have been shown to play a role in maintaining genome integrity in yeast [107]. As the molecular function of many of these *trans*-acting factors is

known, the mechanisms by which they affect genomic instability are generally well understood.

Nevertheless, there are certain aspects of genomic instability that cannot be explained by the action of *trans*-acting factors alone. Studies done on radiation-induced instability have noted that 10-25% of surviving cells in an irradiated population will display genomic instability, even when low doses are administered [108, 109]. This frequency of induced genomic instability is too high to be explained exclusively by the deregulation of a single protein or even an entire family of proteins. It has also been shown that the transmission of genomic instability from an irradiated, unstable parental clone to a sibling subclone can occur in a non-Mendelian fashion; some siblings exhibit a diminished degree of instability and others exhibit a higher degree [110, 111]. This argues against a simple model where genetic mutations in *trans*-acting factors acquired during radiation treatment are solely responsible for genomic instability. Furthermore, analysis of irradiated cells that display genomic instability has shown that the chromosomal rearrangements that are present are not randomly distributed throughout the karyotype [112, 113]. Since *trans*-acting factors are physically separate from DNA and exert their effect on instability randomly with respect to the entire genome, a non-random distribution of rearrangements supports the notion that other mechanisms are involved.

This has led to the idea that *cis*-acting mechanisms may play a key role in the acquisition of genomic instability [110, 114]. One such *cis*-acting mechanism is DRT/DMC. As mentioned in the previous section, DRT/DMC results in

genomic instability in the form of CIN and CSIN [75, 77]. DRT/DMC is a *cis*-acting mechanism, meaning that only the affected chromosome displays replication delay, condensation delay and instability [74, 77]. This phenotype can be induced by exposure to IR and other DNA damaging agents, indicating that DRT/DMC may be responsible for some of the previously mentioned behavior observed in irradiated cells [77, 80]. Interestingly, an analysis of multiple types of human tumors revealed that eight of ten tumor cell lines and five of thirteen primary tumor samples contained DRT/DMC chromosomes [74]. The demonstration that DRT/DMC results in genomic instability and is present in primary tumor cells suggests that it is a common source of genomic instability in human cancer.

The instability observed following DRT/DMC has a very unique cytogenetic signature, with most of the chromosomal rearrangements affecting the delayed chromosome [77]. This single-chromosome instability is reminiscent of two newly described instability signatures, “chromothripsis” and “kataegis”, which are present in some cancers [115, 116]. Chromothripsis and kataegis appear to be cataclysmic events in which a chromosome, chromosome arm or local region on a chromosome is fragmented or heavily mutated in a relatively short period of time. This clustering of mutational events occurs in *cis* and results in entire chromosomes or local regions that have undergone extreme fragmentation and mutagenesis [115, 116]. In the case of chromothripsis, the sequences at the rearrangement junctions show either a lack of homology or microhomology between the joined segments, suggesting that the DNA was

repaired by non-homologous end joining (NHEJ) [115]. In addition, the complex chromosome rearrangements associated with genomic disorders in humans were recently found to resemble chromothripsis [117, 118].

Sequencing the breakpoints at these complex rearrangements identified characteristic features, including small templated insertions of nearby sequences and microhomologies, suggestive of replicative processes. These observations led the Lupski group to propose the term ‘chromoanasynthesis’ as an alternative descriptor to chromothripsis for the shattering and reassembly of chromosomes via replicative mechanisms [117]. The Lupski group proposed a microhomology mediated break induced replication (MMBIR) and a related fork stalling and template switching (FoSTeS) model for the origin of these complex rearrangements [119]. The distinction between MMBIR/FoSTeS and NHEJ is that the microhomology junctions in MMBIR/FoSTeS are followed by stretches of DNA sequence derived from elsewhere, usually nearby. The MMBIR/FoSTeS models involve stalled DNA replication forks that are resolved by replication restart using short stretches of homology [119]. Furthermore, stalled DNA replication forks can also be resolved into DSBs providing a substrate that can be repaired by NHEJ [120].

Thus, if multiple stalled replication forks form on a single chromosome and are resolved via MMBIR, FoSTeS, or strand breakage followed by NHEJ, it would leave the chromothripsis and kataegis instability signatures. One possibility is that DRT/DMC is responsible for the formation of multiple stalled replication forks on a single chromosome. As mentioned in the previous section, there are two

models for how DRT/DMC can induce CSIN. One is that DRT/DMC results in checkpoint adaptation and the onset of mitotic chromosome condensation prior to the completion of DNA synthesis leading to stalled replication forks [85]. This would generate multiple stalled replication forks on a single chromosome and lead to multiple rearrangements generated at the stalled replication forks via NHEJ, MMBIR, and/or FoSTeS type mechanisms. The other way DRT/DMC can induce CSIN is by delaying mitotic spindle attachment, which can lead to chromosome mis-segregation and the formation of micronuclei [85]. Interestingly, it has been found that inducing micronuclei by nocodazole treatment can lead to extreme fragmentation of single chromosomes [84]. Therefore, DRT/DMC is not only present in some cancer cells but it can also explain certain unique instability signatures that are present in some cancer cells.

Genetic Analysis of DRT/DMC

Early studies on DRT/DMC indicated that it causes genomic instability and is present in some cultured and primary cancer cells, but gave no indication as to what caused the phenotype in the first place [68, 74]. The first clue into the cause of DRT/DMC came from the observation that all of the chromosomes that displayed DRT/DMC in cancer cells were translocation derivatives [74]. This indicated that a chromosomal rearrangement could be an initiating event for this phenotype. Subsequent studies confirmed this by generating interchromosomal translocations (ICTs) in normal cells via low dose radiation treatment [80]. It was found that IR treatment generated about one ICT per cell and roughly 5-10% of cells displayed DRT/DMC following treatment [80]. Thus, 5-10% of all balanced ICTs display DRT/DMC. This high frequency allowed for the use of chromosome engineering to systematically generate defined ICTs that induce DRT/DMC *de novo*.

To engineer such chromosomes, the Thayer lab employed a strategy using the causes recombination/locus of cross-over (x-over) P1 (Cre/*lox*P) recombinase system. This strategy, as outlined in Figure 1.2, generates random ICTs via Cre/*lox*P-mediated, site-specific recombination [77]. Briefly, an adenine phosphoribosyltransferase (APRT) deficient human cell line, HTD114 [121], was transfected with two plasmids containing 34 bp Cre-recognition sequences (*lox*P sites). This generated parental cell clones (P-clones) that contain a *lox*P site in two different chromosomes. Following transient Cre expression, site-specific

recombination between the two *lox*P sites was induced and the APRT selectable marker was reconstituted, generating an ICT. These recombinant clones (R-clones) that contain a defined ICT were then analyzed for DRT/DMC.

Of the 83 R-clones that were generated, five displayed DRT/DMC on at least one translocation derivative [77]. Four of these five cell lines that display DRT/DMC are shown in Figure 1.3. Subsequent studies have shown that when these same translocations are induced by endonuclease digestion (I-Sce1) followed by repair via NHEJ, they also display DRT/DMC indicating that this phenotype is not the result of a specific repair process [77]. Furthermore, some of these clones only display DRT/DMC on one translocation derivative, which suggests that this phenotype is caused by a *cis*-acting mechanism [77]. Importantly, the DRT, DMC and genomic instability induced by Cre//*lox*P recombination phenocopies that of DRT/DMC chromosomes found in cancer cells, making this an excellent model to study how this phenotype develops during cancer [74, 77]. The Thayer lab proposed two models for how ICTs can cause DRT/DMC: 1) the ICT could delete or disrupt a *cis*-acting genetic element that acts to ensure proper DNA replication timing, mitotic chromosome condensation, and chromosome stability (loss-of-function model) or 2) the ICT could generate a dominant interfering element that disrupts DNA replication timing, mitotic chromosome condensation, and chromosome stability (gain-of-function model) [74, 77].

The identification of specific chromosomal loci that are involved in the acquisition of DRT/DMC allowed for the first genetic characterization of this

phenotype. The P175 cell line was chosen for further characterization (Figure 1.4). P175 contains a *lox*P site in one allele of chromosomes 6 and 10 [77]. Following *lox*P recombination, a balanced translocation, t(6;10)(q14-15;q11.2), was generated that displayed DRT/DMC [77]. To identify which *lox*P integration site was responsible for the acquisition of DRT/DMC (the chromosome 10 locus, the chromosome 6 locus, or both), the Thayer lab generated alternative translocation partners to see whether the phenotype segregated with one locus or both [122]. Thus, other chromosomes in the P175 cell line were tagged with *lox*P sites and induced to undergo an ICT with either chromosome 6 or 10, and these new translocations were subsequently assayed for the DRT/DMC phenotype. It was found that three of four new translocations with the chromosome 6 locus displayed the DRT/DMC phenotype, whereas zero of three new translocations with the chromosome 10 locus displayed the phenotype [122]. This suggested that the chromosome 6 locus was required for generating the DRT/DMC phenotype in P175 and the chromosome 10 locus was playing a passive role.

Now that the chromosome 6 locus in P175 was identified as the potential mediator of DRT/DMC, intrachromosomal deletions (ICDs) were made on chromosome 6 [122]. ICDs on chromosome 6 were made using two different methods. One generated deletions anchored at the *lox*P site and extending towards the chromosome 6 centromere (proximal deletions) and the other generated deletions anchored at the *lox*P site and extending towards the q-arm telomere (distal deletions). Deletions of many different sizes were generated in

each direction and it was found that the distal deletions never displayed DRT/DMC, whereas all proximal deletions assayed displayed DRT/DMC [122] (Figure 1.5). The smallest deletion proximal from the *lox*P site on chromosome 6 that induced DRT/DMC was 76 kilobases (kb) [122]. To identify a smaller region of interest within this 76 kb, the Thayer lab successfully used a recombinant adeno-associated virus (rAAV) to generate a smaller, targeted deletion at this region. This smaller deletion (47 kb), interestingly, did not cause DRT/DMC, indicating that the region of difference between the two smallest deletions (29 kb region in Figure 1.6A) was the genomic region responsible for generating DRT/DMC on chromosome 6 [123].

Intriguingly, prior to the identification of the 29 kb region on chromosome 6, a region was identified on the X chromosome that was implicated in the acquisition of a DRT/DMC-like phenotype [79, 81]. York Marahren's lab identified a 21 kb region on the X chromosome that, when deleted, caused delayed replication and genomic instability in *cis* (chromosome condensation in these studies was not examined) [79, 81]. Furthermore, the deletion of this locus on both X chromosomes resulted in a more pronounced replication delay than in cells that contained only a heterozygous deletion, exposing a possible *trans*-effect [81]. Currently, two genetic regions have been identified that cause chromosome-wide delayed replication when deleted, a 21 kb region on the X chromosome and a 29 kb region on chromosome 6.

Gene Inactivation and Asynchronous Replication on Chromosomes

Prior to a more in-depth discussion of the loci that are involved in DRT/DMC, it is necessary to introduce the processes of X-chromosome and autosomal inactivation. The Thayer lab believes that these processes are functionally linked to DRT/DMC.

X Chromosome Inactivation:

X chromosome inactivation (XCI) ensures that X-linked gene expression levels are normalized between female (XX) and male (XY) mammalian cells, despite the difference in the number of X chromosomes. To accomplish this, one X chromosome in female cells undergoes transcriptional inactivation, resulting in only one X chromosome that is transcriptionally competent [124]. The silenced allele is referred to as the inactive X chromosome (Xi) and the transcriptionally competent allele is referred to as the active X chromosome (Xa) [125]. One unique feature of the Xi is that the entire chromosome is late replicating [126]. Interestingly, the same replication origins are utilized on the Xi and Xa, indicating that the origins on each X chromosome are differentially regulated [127, 128]. Therefore, following XCI the Xi is transcriptionally silenced and late replicating while the Xa is transcriptionally active and early replicating [129]. This results in the asynchronous replication of the X chromosomes. However, unlike the DRT

phenotype, XCI is a normal cellular process that does not result in DMC or genomic instability and is necessary for organismal viability [130].

Female embryonic stem cells (ESCs) contain two transcriptionally competent X chromosomes, and upon differentiation, one X chromosome is chosen to become inactivated [131]. The process of XCI involves three distinct activities: Counting, Choice, and Inactivation. Prior to ESC differentiation, the cell takes an inventory of how many X chromosomes are present and inactivates all but one X chromosome [124, 132]. This process of counting follows the 'n-1 rule' to ensure that only one X chromosome is active, even when more than two X chromosomes are present (n refers to the number of X chromosomes in the cell and n-1 equals the number of inactive X chromosomes) [132]. After the X chromosomes have been counted, one X chromosome is chosen at random to become the Xa while the other becomes inactivated [133]. The random choice of which allele will be active and which will be inactive ensures that a mosaicism is established in the adult organism, where the Xi in some cells will be the Xa in others [134]. Finally, the randomly chosen Xi becomes transcriptionally inactivated and late replicating via the establishment of repressive chromatin marks, DNA methylation and a change in nuclear positioning [135].

The processes of counting, choice and inactivation are regulated by a region on the X chromosome called the X inactivation center (XIC) [124, 136]. The XIC is an ~700 kb region on Xq13 that harbors multiple non-coding RNA (ncRNA) genes involved in the establishment of XCI [136-138]. One of these genes is X-inactivation specific transcript (Xist), a 17 kb, spliced, untranslated

ncRNA [126]. During the establishment of XCI, Xist becomes expressed from the X chromosome destined to become the Xi [126]. This monoallelically-expressed transcript spreads in *cis* along the future Xi, coats the chromosome and is thought to recruit chromatin modifying complexes and histone variants [139]. Before Xist spreading and coating, the transcript remains bound to the XIC via tethering of the RNA to a nucleation center within the *Xist* DNA by the transcription factor Yin Yang 1 (YY1) [140]. YY1 is thought to bind the repeat F region of *Xist* DNA on the Xi and interact with the Xist transcript via the Repeat C region [140]. Following Xist RNA accumulation around the XIC, the transcript spreads in *cis* along the Xi and coats the chromosome [140, 141]. The establishment of repressive histone marks, such as histone H3 hypoacetylation and hypermethylation, histone H2A ubiquitinylation, and macroH2A recruitment, follows Xist coating and establishes transcriptional silencing in *cis* [126, 135]. *Xist* appears to be one of the major components in this process, as this gene is both necessary and sufficient for X inactivation [126]. Another XIC component, Tsix, is an antisense transcript to Xist [142]. Similar to most other X-linked genes, Tsix is monoallelically expressed from the Xa (opposite of Xist) [142]. Tsix expression functions to antagonize Xist expression by silencing Xist transcription on the Xa [142]. This interplay of Xist and Tsix expression during ESC differentiation establishes the Xist-expressing chromosome as the Xi and the Tsix-expressing chromosome as the Xa.

Interestingly, some regions on the Xi escape gene inactivation. It has been estimated that 10-15% of genes on the human X chromosome are expressed

biallelically to some degree [143]. These include genes that are present on both the X and the Y chromosomes and genes that lie in regions with low long interspersed nuclear element-1 (LINE-1) concentration [144, 145]. LINE-1 elements are retrotransposons that have integrated into mammalian genomes during the course of evolution [146]. The human genome contains hundreds of thousands of LINE-1 elements, most of which are incapable of retrotransposition due to mutational inactivation or 5´ truncation [146]. Although the typical human autosome contains ~17% LINE-1 sequences, the X chromosome contains almost twice as much [124]. Furthermore, the highest LINE-1 concentration on the X chromosome is around the XIC [147]. This chromosome-specific accumulation of repetitive elements led Mary Lyon to propose that LINE-1s represent “booster elements” that aid in the propagation of Xi silencing [148]. Thus, regions with high LINE-1 concentration have robust silencing on the Xi while the silencing on regions with low LINE-1 concentration is attenuated.

Once gene inactivation and late replication are established on the Xi, a change in nuclear positioning takes place. The coating of Xist RNA on the Xi is followed by its migration to the perinucleolar region or nuclear periphery of the nucleus [149, 150]. The Xi nuclear compartment is very dense with heterochromatin and is void of RNA polymerase II (RNA Pol II) and transcription factors [151]. This compartment can be identified by the presence of an Xist RNA “cloud.” The Xist RNA cloud is the 3-dimensional accumulation of Xist RNA within the nuclear compartment that contains the Xi [135, 149]. Highly repetitive

sequences and inactive genes reside within the Xist RNA cloud while genes that escape XCI are found outside of the cloud [135].

The sufficiency of *Xist* to recapitulate XCI was demonstrated through the use of *Xist*-containing transgenes integrated into autosomes [152]. Initially, transgene studies were used to identify the molecular boundaries of the XIC, but ended up identifying some very interesting characteristics of XCI. Yeast artificial chromosomes (YACs) and cosmids containing *Xist* and surrounding elements were randomly integrated into autosomes in mouse ESCs. Upon ESC differentiation, the integrated autosomes recapitulate some aspects of XCI, including Xist expression, Xist RNA coating and cloud formation, gene inactivation, late replication and some establishment of repressive chromatin marks [153-156]. However, these properties were not observed in all clones following differentiation and single-copy arrays of the transgene were not sufficient to induce these properties, as multi-copy integrations were needed [156]. Interestingly, the ability of *Xist* to recapitulate aspects of XCI on autosomes is not restricted to differentiating ESCs. Multiple labs have shown that *Xist* transgene integration into autosomes is sufficient to establish gene inactivation, Xist RNA cloud formation and repressive chromatin marks in the differentiated cell lines HT1080 and HELA [157-159]. Again, many of the transgene-integrated autosomes had variable inactivation status and others recapitulated only some aspects of XCI, indicating that additional variables are involved [158, 160]. Nevertheless, one unique feature of *Xist* is its ability to dominantly interfere with the normal replication-timing program when integrated into an autosome.

Autosomal Inactivation:

The process of XCI establishes a chromosome-pair non-equivalence, such that the two X chromosomes in female cells function differently. This is evident in the different gene expression patterns, replication timing patterns and chromatin and DNA methylation patterns of the two chromosomes [126]. As it turns out, a similar non-equivalence is also established on autosomes in a process called autosomal inactivation (AI).

Every mammalian chromosome shows differences in gene expression with its homolog. This occurs to an exceptional degree on the X chromosome and to a much lesser extent on autosomes [161]. The gene expression differences between homologous autosomes are established by AI, which includes the processes of genomic imprinting and random monoallelic gene inactivation [161]. Genomic imprinting involves the inactivation of gene expression from either the maternal or paternal allele [161]. Imprinted genes tend to cluster spatially on autosomes and are not very common, only affecting about 1% of autosomal genes [162]. A much more common occurrence is random monoallelic gene inactivation. Like XCI, the choice of which allele will be inactivated is chosen stochastically by each cell and is stably inherited by all subsequent generations. Random monoallelic gene inactivation affects ~5-10% of all autosomal genes and tends to involve large, multi-gene families [163]. The focus of this section will be on AI via random monoallelic gene inactivation, which I will refer to as “random AI.” Interestingly, where ~90% of the genes on the X chromosome are monoallelically expressed and ~10% escape inactivation and

remain biallelically expressed, the opposite is true for autosomes. So, in effect, random AI accomplishes the same thing as XCI, only instead of 10% of genes escaping inactivation, 90% of genes escape inactivation on autosomes.

A wide variety of genes are subject to random AI. Many are involved in chemosensory or immune system functions, such as antigen receptors, pheromone receptors and odorant receptors; however, other genes subject to random AI have more diverse functions [161, 164]. For example, ribosomal DNA (rDNA) and many genes involved in cell adhesion are randomly monoallelically inactivated, indicating that ubiquitously expressed genes, in addition to tissue-specific genes, can be regulated in this manner [165, 166]. Like XCI, random AI appears to be established during early embryogenesis and utilizes many of the same epigenetic factors to establish chromosome-pair non-equivalence [161, 167]. Differential DNA methylation, chromatin modifications and subnuclear localization of the two homologs generally accompany random AI [161, 166, 167]. Furthermore, most monoallelically-expressed loci are at or near regions of ncRNA gene expression, but it is unclear what role, if any, ncRNA function has in the process of random AI [161, 167]. Similar to genes on the X chromosome that are subject to XCI, genes that are subject to random AI tend to reside in areas of high LINE-1 concentration [168].

Most autosomal loci replicate synchronously, meaning that any given locus on an autosome will replicate at the same time in S phase as that of its homolog. However, autosomal loci that are subject to random AI replicate asynchronously [166, 167]. Interestingly, asynchronous replication is established

at these loci around the time of implantation, before gene inactivation and monoallelic expression [164, 166, 167]. Different from whole-chromosome replication asynchrony that occurs on the X chromosome, only ~10% of autosomal loci are subject to replication asynchrony [169]. However, like the X chromosome, this replication asynchrony is coordinated at the whole-chromosome level. Dr. Andrew Chess's lab showed that, although only a small number of autosomal loci are asynchronously replicating, these loci coordinate their replication patterns with one another in *cis* along the entire chromosome [170-172]. For example, on chromosome 1, ~90% of the loci will replicate synchronously and ~10% will replicate asynchronously. Of the ~10% that replicate asynchronously, the early-replicating alleles will all be on the same chromosome 1 homolog and the late-replicating alleles will all be on the other homolog. So, even though these loci are distant from each other and on both sides of the centromere, their replication timing is somehow coordinated along the entire chromosome. This has led to the proposition that a ncRNA, similar in function to Xist, might be responsible for autosomal replication coordination [172]. In addition, autosomal replication asynchrony has been shown to follow the 'n-1 rule' in hyperdiploid cells [170, 172].

Since all loci subject to random AI are asynchronously replicating and since all asynchronously replicating loci coordinate DNA replication along each chromosome, then it would make sense that random monoallelic expression would also be coordinated in *cis* along each chromosome. Curiously, this does not seem to be the case [164]. Studies with clonal human B-cell lines have

established that, on any given autosome, some random monoallelically-expressed genes will be expressed from the maternal allele and some from the paternal allele, indicating that monoallelic expression is not coordinated chromosome-wide [163]. Since asynchronous DNA replication is coordinated along each autosome and monoallelic gene expression is not, this indicates that some random monoallelically-expressed genes are expressed from the late-replicating allele and vice versa. This suggests that, while asynchronous DNA replication and monoallelic expression tend to be concurrent features, they may be controlled by distinct mechanisms. This is supported by the demonstration that tissue-specific, monoallelically-expressed genes are asynchronously replicating even in tissues in which they are not expressed [164, 165]. Furthermore, although random monoallelic gene expression is a stable feature in individual clones once it is established, the existence of monoallelic expression can vary widely between different cells. Thus, in a population of cells, some clones will display monoallelic expression of a specific gene, some clones will display biallelic expression of that gene and some clones will display no expression at all [163, 165, 173]. Despite the differences in expression between different clones, these loci invariably display asynchronous DNA replication timing. As a result, asynchronous replication is a more consistent feature of loci subject to random AI than monoallelic expression, indicating that monoallelic expression *per se* may be a consequence of asynchronous replication.

Random AI and XCI share many unique features, including: 1) the chromosome-wide coordination of asynchronous replication that follows the ‘n-1

rule', 2) the establishment of differential DNA methylation and chromatin marks on each homolog, 3) the differential subnuclear localization of loci subject to inactivation, 4) high LINE-1 concentration of loci subject to inactivation, and 5) gene inactivation occurs very early in development. Some notable differences include: 1) XCI establishes the coordination of monoallelic expression along one chromosome (with the exception of Xist) while random AI does not, 2) XCI inactivates ~90% of X-linked genes, while random AI inactivates ~10% of autosomal genes, and 3) the monoallelic expression on the X chromosome is invariably present in all differentiated cells, whereas random monoallelic expression on autosomes displays cell-cell variability in expression patterns. Despite these differences, it has been suggested that XCI and random AI are controlled by a similar mechanism [170]. XCI is coordinated by the XIC, a *cis*-acting region on the X chromosome that harbors monoallelically-expressed transcripts involved in silencing gene expression. Until recently, no similarly functioning locus had been identified on any autosome. As discussed in the following section, the region on chromosome 6 that is involved in DRT/DMC shares many commonalities with the XIC and appears to function in a similar manner.

Molecular and Functional Analysis of Loci that Control Chromosome-wide Replication Timing and Stability

Earlier in this chapter, it was mentioned that two genetic regions have been identified that result in DRT/DMC when deleted, a 21 kb region on the X chromosome and a 76 kb region on chromosome 6. One interesting feature of the 21 kb region on the X chromosome is that it lies within the XIC. Furthermore, not only does it lie within the XIC, but the 21 kb deletion encompasses the promoter and first three exons of *Xist*, silencing expression [81]. This suggests that not only do *Xist* and the XIC govern the process of XCI, but they also control the chromosome-wide replication timing and stability of the X chromosome.

Since the 76 kb region on chromosome 6 appears to function similarly to the XIC in controlling replication timing and stability, the Thayer lab analyzed this region and found many shared molecular characteristics with the XIC. The 76 kb region on chromosome 6 lies within a large intergenic ncRNA gene, which they named asynchronous replication and autosomal RNA on chromosome 6 (*ASAR6*) (Figure 1.6B-C) [122, 174]. Like Xist, ASAR6 is monoallelically expressed, restricted to the nucleus, and is an RNA pol II product [122, 123, 175]. ASAR6 is not spliced or poly-adenylated and is ~200 kb in length [122, 123]. Although transcription can be detected in a >200 kb region on chromosome 6, it is unlikely that ASAR6 represents one long transcript due to the multiple transcription start sites (TSSs) within the gene [122, 123]. Therefore, *ASAR6* most likely represents multiple, smaller, overlapping ncRNA transcripts.

Furthermore, *ASAR6* exhibits random monoallelic expression, and is expressed in some, but not all, adult tissues [122]. ASAR6 is expressed from the *lox*P containing allele in P175 cells, indicating that deletions in chromosome 6 disrupt the transcribed *ASAR6* gene [122]. Due to the multiple TSSs within *ASAR6*, some of the deletions that cause DRT/DMC on chromosome 6 do not completely eradicate *ASAR6* transcription [122]. Nevertheless, every deletion on chromosome 6 that causes DRT/DMC at least partially disrupts the *ASAR6* gene. Interestingly, a 47 kb deletion on chromosome 6 that does not cause DRT/DMC lies just outside the transcribed region of *ASAR6* (Figure 1.6) [123].

The region surrounding *ASAR6* also contains other transcripts, some of which are protein-coding genes and some of which are ncRNAs (Figure 1.7) [122]. The protein-coding gene mannosidase, endo-alpha (*MANEA*) and the protein-coding gene *KIAA0776* (now called E3 UFM1-protein ligase 1 (*UFL1*)) are biallelically expressed while the protein-coding gene fucosyltransferase 9 (alpha (1,3) fucosyltransferase) (*FUT9*) and the ncRNA four and a half LIM domains 5 (*FHL5*) opposite-strand transcript (*FHL5ost)* are monoallelically expressed in P175 cells [122]. The protein coding gene *FHL5* is not expressed in P175 cells. Interestingly, FUT9 and FHL5ost are both transcribed from the opposite allele as ASAR6 [122]. Therefore, on the chromosome 6 that expresses ASAR6, FUT9 and FHL5ost are both silent and on the chromosome 6 that expresses FUT9 and FHL5ost, ASAR6 is silent.

As is the case with most monoallelically-expressed genes, *ASAR6* is asynchronously replicating [122]. In fact, *ASAR6* lies within an ~1.2 megabase

(mb) region that replicates asynchronously, despite containing both biallelically and monallelically expressed transcripts (Figure 1.7) [122, 123]. As has been described before [170], the asynchronous replication of genes within this ~1.2 mb region is coordinated in *cis* (Figure 1.8) [122]. So the early-replicating alleles of *ASAR6*, *FUT9*, *FHL5ost, MANEA, and KIAA0776* are on the same chromosome and the late-replicating alleles are on the other homolog. Interestingly, the chromosome 6 that expresses ASAR6 is the late-replicating homolog, while the chromosome 6 that expresses FUT9 and FHL5ost is the early-replicating homolog [123]. This is similar to the X chromosome in female mammalian cells, where Xist is expressed from the late-replicating homolog and all of the other monoallelically-expressed genes are expressed from the early-replicating homolog [176].

Although it has been previously reported that all chromosomes coordinate their asynchronous replication in *cis* [170, 172], this does not seem to be the case with chromosome 6. The ~1.2 mb asynchronously replicating domain that harbors *ASAR6* is coordinated in *cis* with some distant asynchronously replicating loci on chromosome 6 and in *trans* with other asynchronously replicating loci (Figure 1.8) [122, 123]. However, in each case, the asynchronous replication is coordinated at the whole-chromosome level.

One last unique feature of the *ASAR6* locus is its ability to control gene silencing in *cis*. Deletions on chromosome 6 that disrupt *ASAR6* and cause DRT/DMC also reactivate the previously silent alleles of *FUT9 and FHL5ost* [122]. Therefore, when *ASAR6* is disrupted, FUT9 and FHL5ost become

biallelically expressed. In the cells used for this study, the only genes on chromosome 6 that displayed monoallelic expression were *ASAR6*, *FUT9*, and *FHL5ost*, so it is unclear if *ASAR6* disruption causes a chromosome-wide loss of monoallelic gene expression or just disruption of nearby monoallelic gene expression. Interestingly, disruption of the Xist transcript can reactivate silenced genes on the Xi [150, 177], so it appears that both *ASAR6* and *Xist* function similarly in the maintenance of linked monoallelic gene expression.

Thus, the *ASAR6* and *Xist* loci have many similarities including: 1) asynchronous replication, 2) antisense, nuclear, RNA pol II-transcribed ncRNA expression that is subject to random monoallelic expression, 3) monoallelic gene expression from the late-replicating allele, 4) gene expression in ESCs, 5) the disruption of either locus results in delayed replication timing and instability of entire chromosomes in *cis*, and 6) the disruption of either locus results in the transcriptional activation of the previously silent alleles of linked monoallelic genes. However, there are also many differences, such as: 1) Xist coats the Xi whereas there is no indication that ASAR6 coats chromosome 6, 2) Xist is expressed in all tissues while *ASAR6* is only expressed in some, 3) Xist is spliced and polyadenylated [175] and ASAR6 is not, and 4) Xist is a single ncRNA transcript while *ASAR6* appears to be a collection of multiple transcripts with multiple promoters. It is still unclear whether the above differences translate into functional differences between *ASAR6* and *Xist*; however, it is becoming clear that these two genes have a similar impact on chromosome function. The functional similarities between the *ASAR6* and *Xist* loci, with regards to

monoallelic gene expression, replication timing, and chromosome stability, indicate that they might represent two examples of important *cis*-acting essential elements that are present on all chromosomes.

The Inactivation/Stability Center: A Model for *Cis* Control of Chromosome-wide Replication Timing and Stability

Existing data indicate that *ASAR6* and *Xist* reside within loci that regulate DNA replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of their respective chromosomes. This is evident by the DRT, DMC, disrupted monoallelic gene expression and chromosome instability phenotypes that are observed when these loci are disrupted. The DRT/DMC phenotype has been detected on chromosome rearrangements involving many different human and mouse chromosomes [74, 77, 80, 81, 122]. Therefore, it seems likely that all mammalian chromosomes contain loci that function to regulate chromosome-wide replication timing of individual chromosomes. Given the similarities in structure and function of the two loci characterized to date, *Xist* and *ASAR6*, I propose that all mammalian chromosomes contain 'inactivation/stability centers' (I/S centers) that function to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of individual chromosomes. Under this scenario every mammalian chromosome contains four distinct types of *cis*-acting elements: origins of replication, centromeres, telomeres, and I/S centers, which all function to ensure proper replication, segregation and stability of individual chromosomes.

In this thesis, I provide data to support the hypothesis that all mammalian chromosomes contain I/S centers that function to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and

stability of individual chromosomes. Specifically, I show that *ASAR6*, like *Xist*, can act dominantly to disrupt replication timing when ectopically integrated into an autosome. Furthermore, ectopic *ASAR6* integration leads to the formation of an ASAR6 RNA "cloud" in the nucleus and chromosome-wide gene silencing of the integrated autosome. This indicates that ASAR6 has the ability to coat chromosomes in *cis* and inactivate gene expression, which are both essential functions for Xist-mediated gene silencing. I also identify a new locus on chromosome 15 that, when disrupted, results in DRT/DMC and genomic instability. Therefore, three loci have now been identified, on chromosomes 6, 15 and X, that are necessary for the proper replication timing, mitotic chromosome condensation, and stability of their respective chromosomes. I finish with some preliminary data in which I test the two models the Thayer lab has proposed for how DRT/DMC is generated (the loss-of-function and gain-of-function models mentioned previously), which opens the avenue for a more in-depth genetic analysis of these loci.

Figures

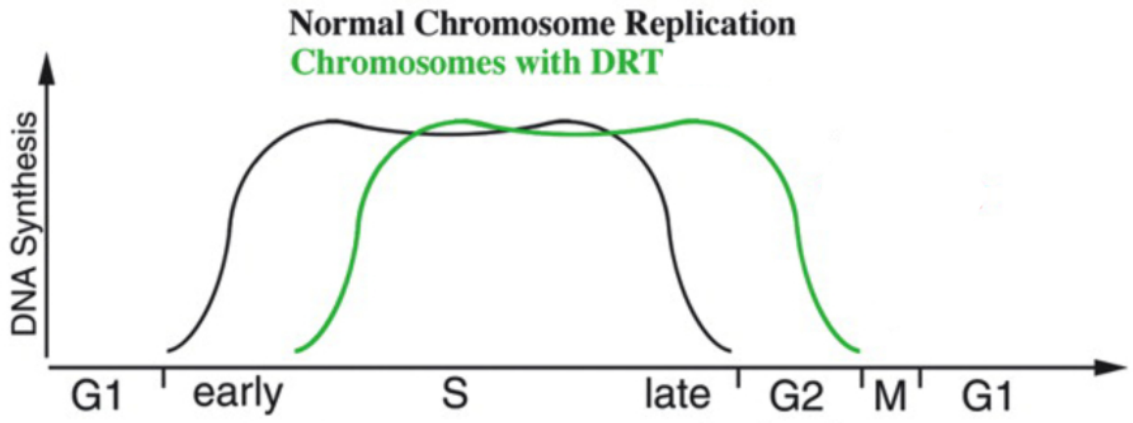


Figure 1.1: Replication Profile of Normal and DRT Chromosomes. The cell cycle phase is represented on the X-axis and relative DNA synthesis is listed on the Y-axis. The black line represents the replication timing profile of a normal chromosome and the green line represents that of a chromosome displaying DRT. Notice how the normal chromosome replicates within the confines of S phase and the DRT chromosome begins replication in early-mid S phase and finishes replication in G2 or M phase. Although they begin and complete DNA replication at different times, the time it takes to replicate each chromosome is comparable. Figure 1.1 is adapted from [77].

Figure 1.2: Schematic diagram of the chromosome engineering strategy. A diagram of the mouse genomic APRT gene, with a unique Hind III site in intron 2, is shown. The 5′ portion of the APRT gene was separated from the 3′ portion at this unique Hind III site. Floxed Neomycin (*Neo)* or Hygromycin (*Hyg)* resistance genes were inserted at the Hind III site in either the 5′ or 3′ portions of the APRT gene, respectively, resulting in the 5′AP-*lox*P and *lox*P-3′RT cassettes. The 5′AP-*lox*P and *lox*P-3′RT cassettes integrated randomly throughout the genome following linearization and electroporation. After Cre transient transfection, reciprocal translocations were generated in a two-step process. First, due to the close proximity of the *lox*P sites flanking the *Neo* and *Hyg* genes, and the fact that they are aligned in the same orientation, the *Neo* and *Hyg* genes were excised as circles via highly efficient (determined to be >90%; data not shown) intra-chromosomal events. Next, Cre directed the remaining *lox*P sites to proceed through a low efficiency (<1 X 10^{-3}) inter-chromosomal reciprocal exchange. This resulted in reconstitution of the APRT gene on one derivative chromosome, and a single *lox*P site on the other derivative, converting cells from APRT-negative (P-clones) to APRT-positive (R-clones) [77, 80]. Figure 1.2 is adapted from [122]. (Figure on next page).

Figure 1.2: Schematic diagram of the chromosome engineering strategy.

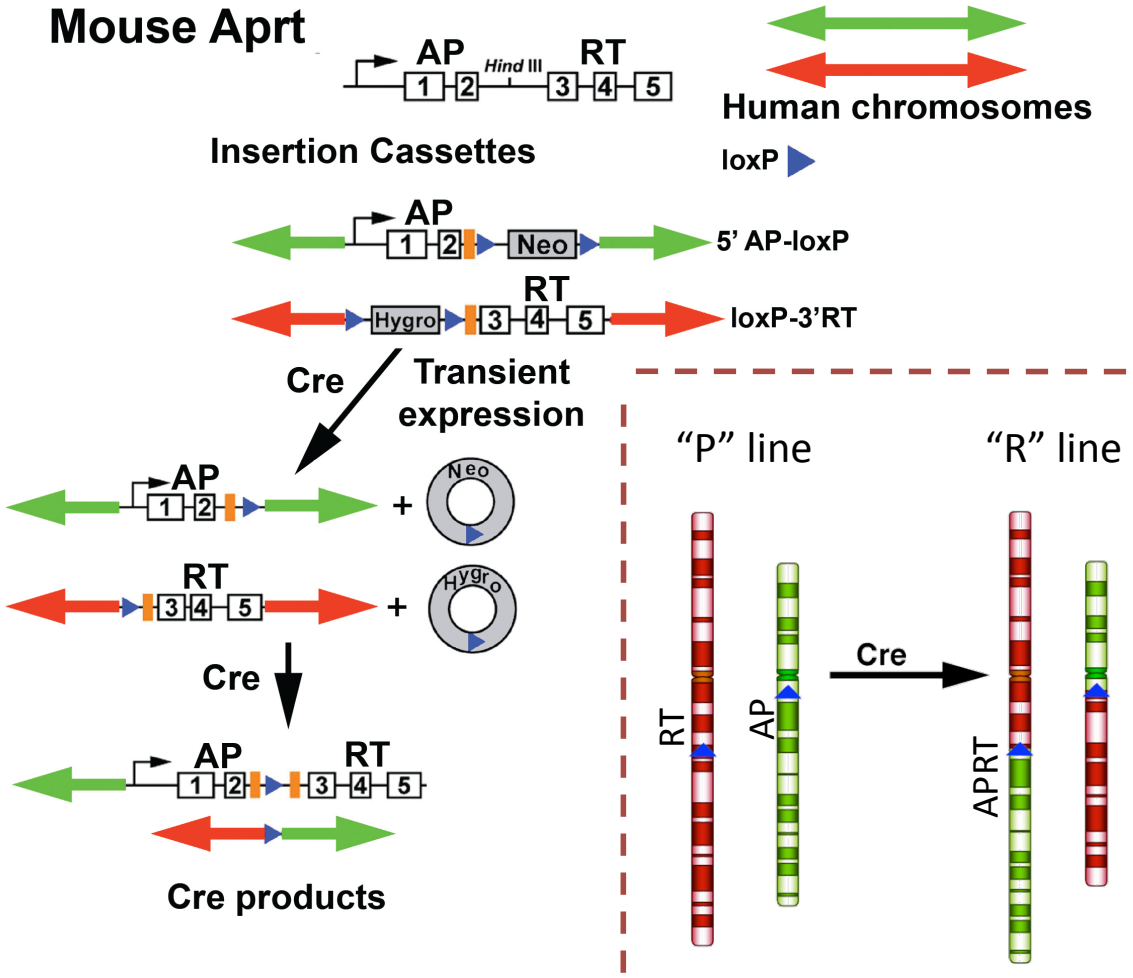
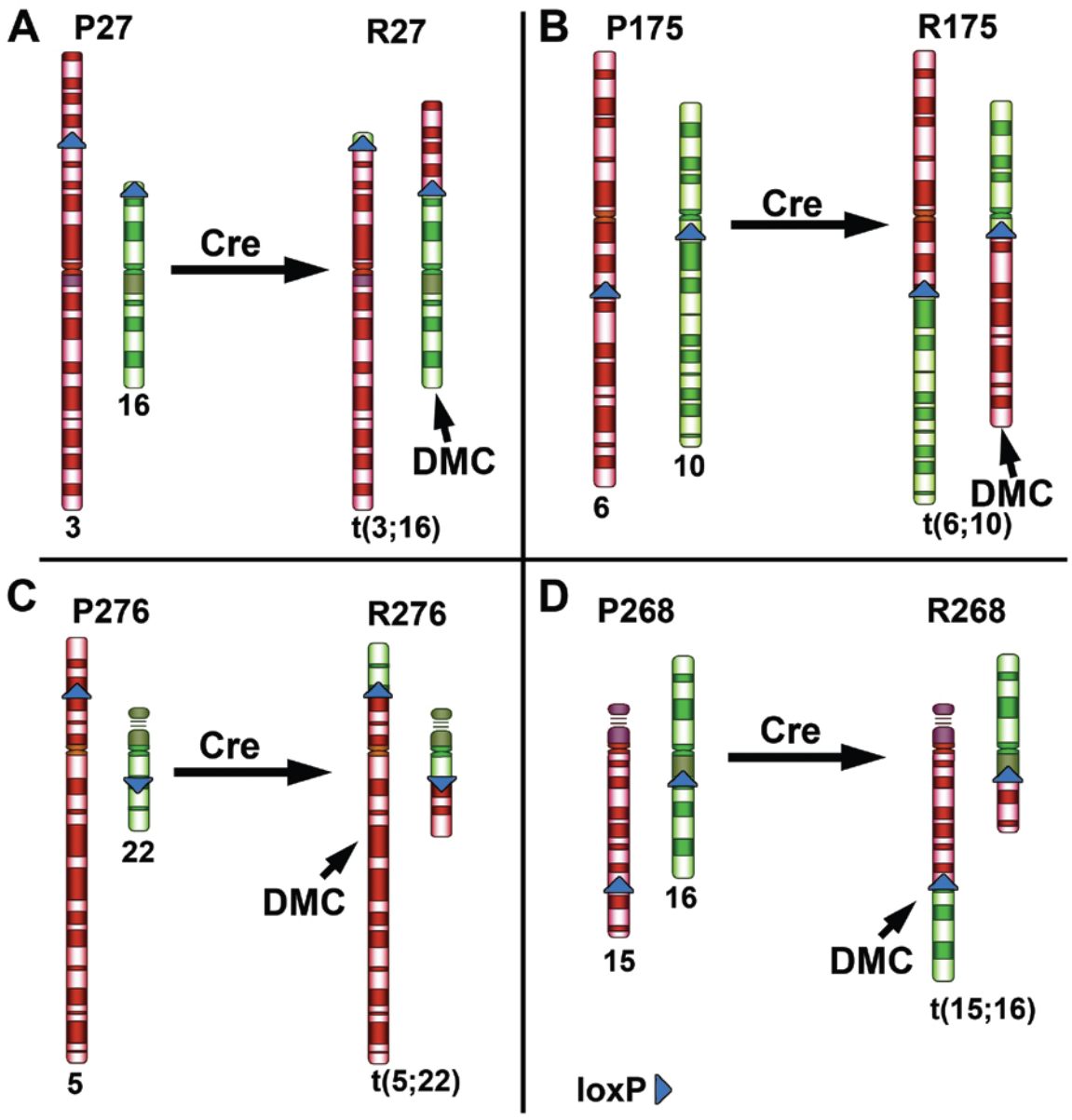


Figure 1.3: DRT/DMC occurs on only one derivative chromosome of certain balanced translocations. A schematic diagram of the (**A**) t(3;16)(p13;p13.3) in R27, (**B**) t(6;10)(q15;q11.2) in R175, (**C**) t(der5p;22)(p14;q11.2) in R276, and (**D**) t(15;16)(q24;q12.1) in R268. Derivative chromosomes that displayed DRT/DMC are indicated with an arrow. Analysis of a fifth balanced translocation, a t(3;13)(q29;q14) (not shown) present in R186, showed that both derivative chromosomes could display DRT/DMC, indicating that the phenotype is not restricted to a single derivative chromosome [77]. Figure 1.3 is adapted from [122]. (Figure on next page).

Figure 1.3: DRT/DMC occurs on only one derivative chromosome of certain balanced translocations.



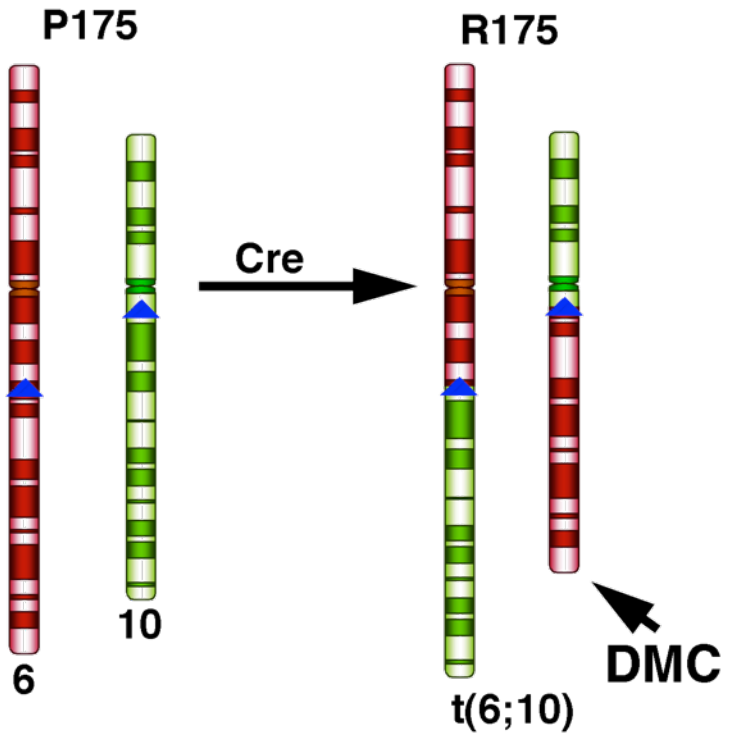


Figure 1.4: Schematic representation of chromosomes 6 and 10 in the P175 and R175 cell lines. The P175 cell line contains a *lox*P site in one allele of chromosome 6 and one allele of chromosome 10. No other chromosomes in this cell line have been manipulated. Following transient Cre expression in P175, R175 was generated which contains a balanced 6;10 translocation. In R175, the chromosome 10 centromere derivative displayed DRT/DMC while the chromosome 6 centromere derivative and all the other chromosomes in the cell did not display DRT/DMC. Figure 1.4 is adapted from [122].

Figure 1.5: Schematic representation of chromosome 6 deletions. P175 cells were induced to undergo an ICD on chromosome 6. Following deletion, clonal cell lines were isolated, each containing a different sized deletion on chromosome 6. Δ175 5´ cells contain a deletion that is anchored at the *lox*P site on chromosome 6 and extending towards the chromosome 6 centromere. Δ175 3´ cells contain a deletion that is anchored at the *lox*P site on chromosome 6 and extending away from the chromosome 6 centromere. Note that chromosome 10 and all of the other chromosomes in Δ175 5´ and Δ175 3´ cells remain unperturbed. Subsequent analysis of multiple Δ175 5´ and Δ175 3´ cell lines revealed that none of the distal deletions result in DRT/DMC and all but one of the proximal deletions result in DRT/DMC, indicating that the region adjacent to the *lox*P site and proximal to the centromere is the region involved in this phenotype. (Figure on next page).

Figure 1.5: Schematic representation of chromosome 6 deletions

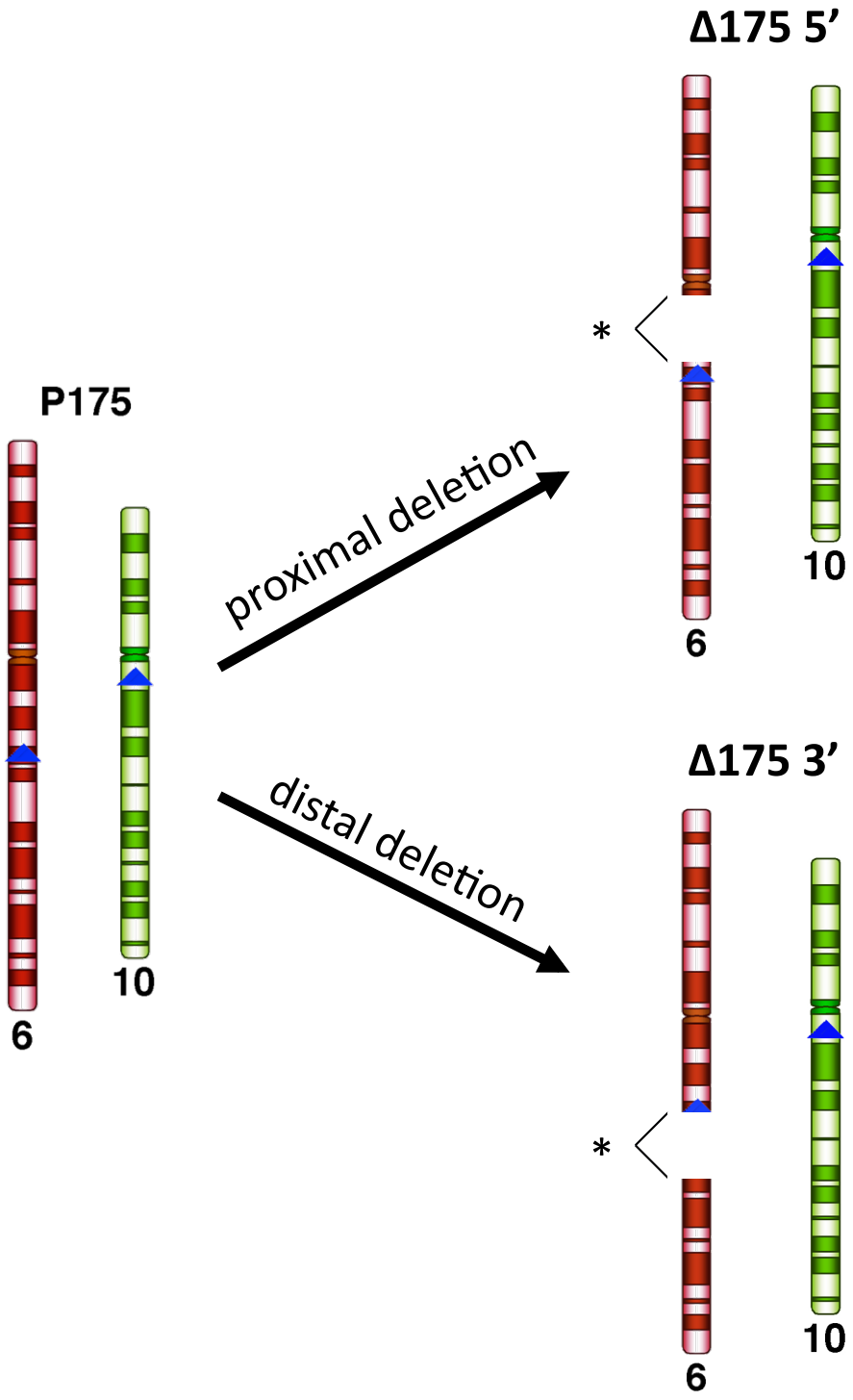
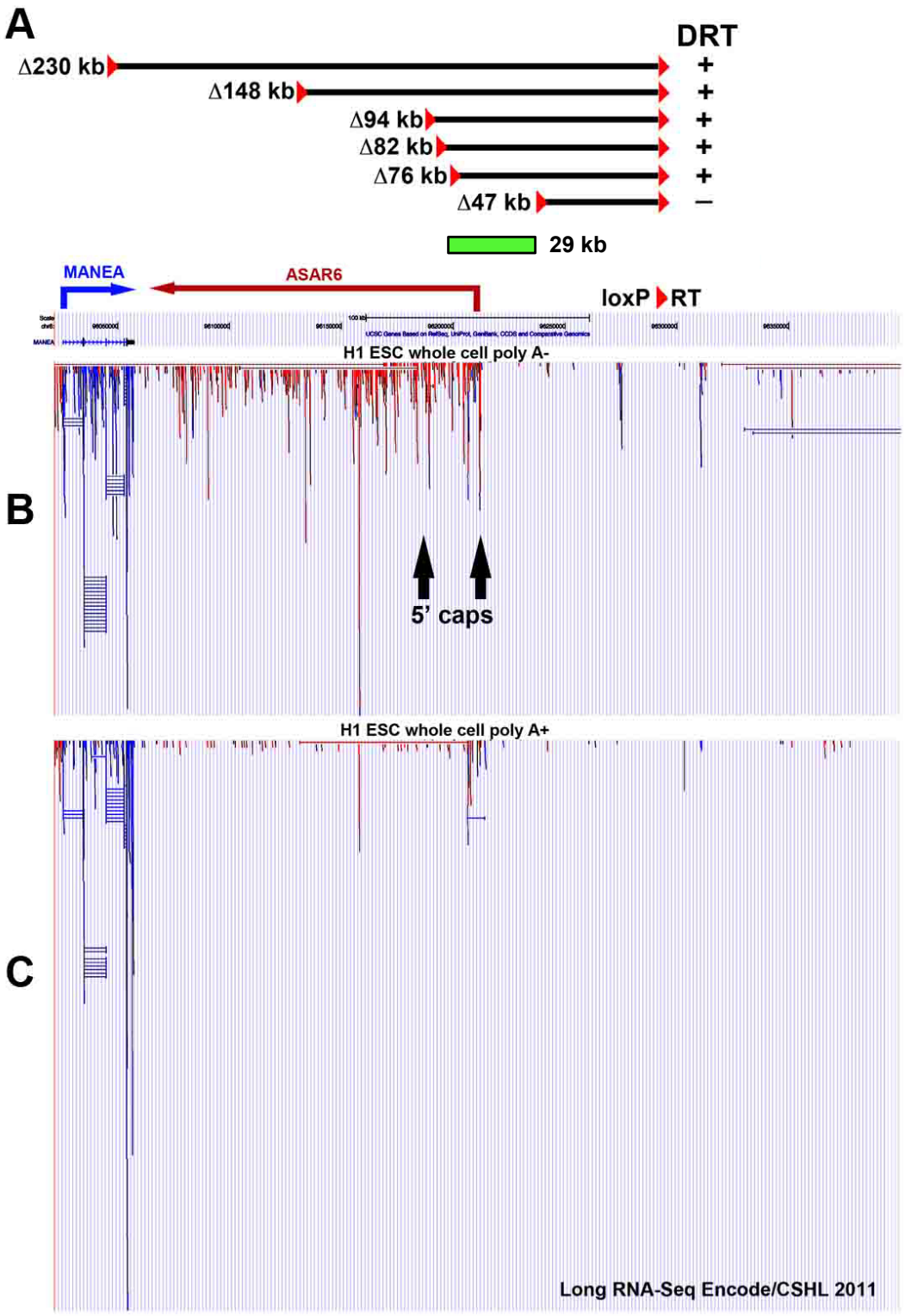


Figure 1.6: Alignment of chromosome 6 deletions along the *ASAR6* locus. The location of two genes on chromosome 6 (*MANEA* and *ASAR6*), the original *lox*P integration site [*lox*P(red triangle)RT] and 6 different deletions in P175 cells [122] are depicted above a screenshot of the University of California, Santa Cruz (UCSC) Genome Browser of this region of chromosome 6. **A**) A set of nested deletions was generated in P175 cells and all except the smallest deletion (Δ47 kb) displayed DRT. The green rectangle above *ASAR6* (29 kb) represents the genetic region that differs between the smallest deletion that caused DRT and the deletion that did not cause DRT. **B** and **C**) UCSC Genome Browser view of the RNA-seq data from whole-cell poly A- (**B**) or poly A+ (**C**) RNA from the human ESC line H1 [178]. The blue tick marks indicate sequence hits from the + direction, and the red tick marks indicate sequence hits from the - direction. Note that ASAR6 RNA is enriched in the poly A- fraction, while MANEA RNA is detected in both poly A- and poly A+ fractions. The locations of 5´ caps from the Encode/RIKEN CAGE [179] track are also shown. Figure 1.6 is adapted from [123]. (Figure on next page).

Figure 1.6: Alignment of chromosome 6 deletions along the *ASAR6* locus.



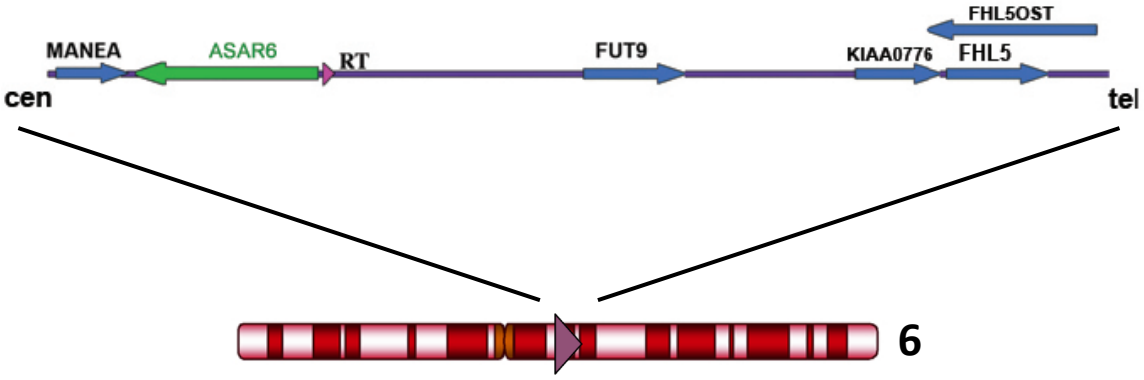


Figure 1.7: Illustration of the ~1.2 Mb region surrounding the *lox*P site on chromosome 6. A zoomed in view of the *lox*P(pink triangle)-3´RT locus on chromosome 6. The genes *MANEA, FUT9, KIAA0776, FHL5* and *FHL5ost* are labeled above a blue arrow showing their relative gene length and transcriptional direction. The *ASAR6* gene is labeled above a green arrow showing its relative gene length and transcriptional direction. The magnified area from *MANEA* to *FHL5ost* represents an ~1.2 mb region of chromosome 6q16. Figure 1.7 is adapted from [122].

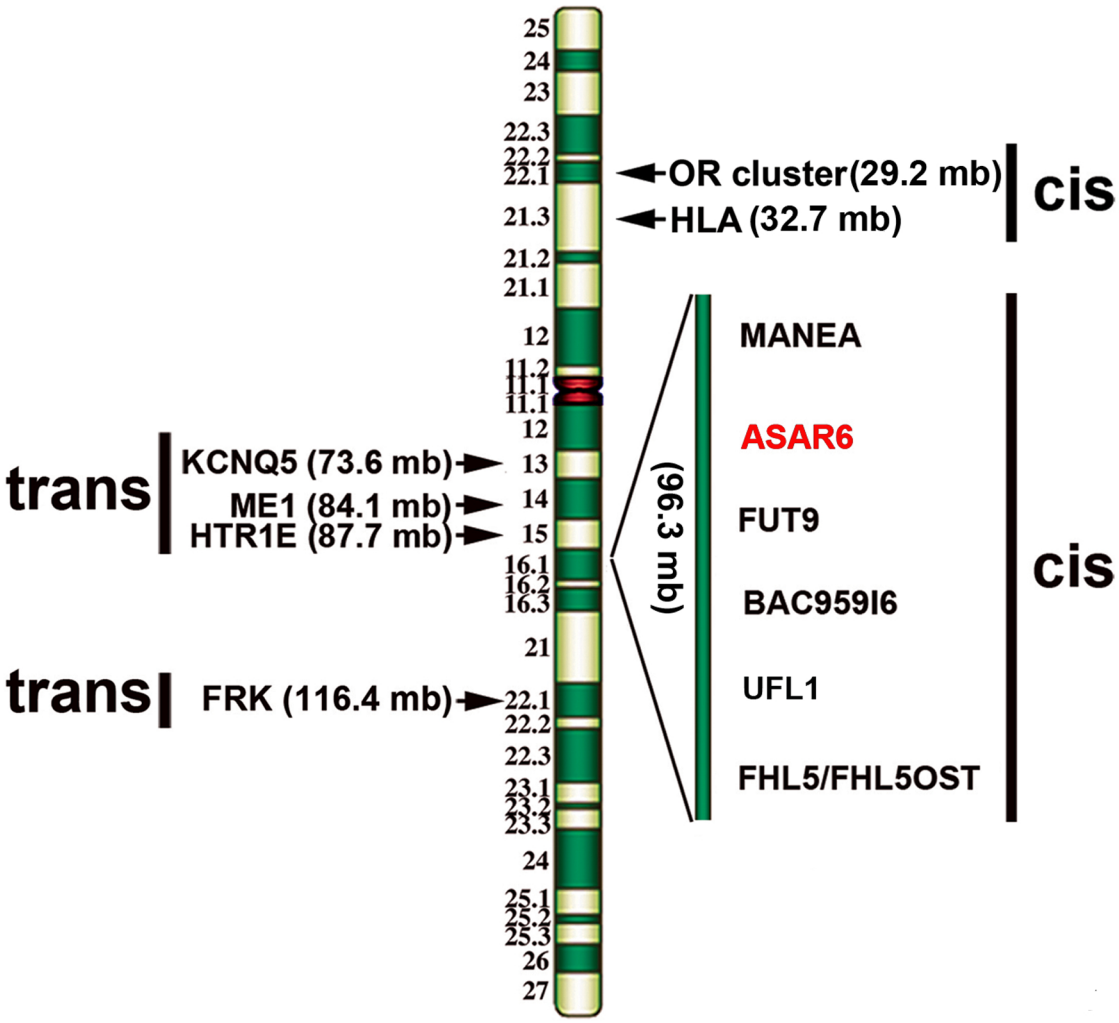


Figure 1.8: Schematic diagram of chromosome 6 showing the location of the loci assayed for asynchronous replication. The ~1.2 mb region of chromosome 6 between *MANEA* and *FHL5/FHL5ost* is expanded on the right. The coordination in asynchronous replication of chromosome 6 monoallelically-expressed genes with *ASAR6* was found to be either in *cis* or in *trans*. Figure 1.8 is adapted from [123].

CHAPTER TWO

Ectopic Integration of an *ASAR6* Transgene Results in Chromosome-wide ASAR6 Coating, Gene Silencing and Delayed Replication

(Part of this chapter was adapted from [123])

Introduction

The Thayer lab has developed a chromosome engineering system that allows for the systematic analysis of human chromosomes with DRT/DMC [75, 77, 80, 122]. This system relies on site-specific recombinases to generate precise chromosomal rearrangements. Using this system, the Thayer lab previously identified four balanced translocations, each displaying DRT/DMC on one of the two derivative chromosomes (Figure 1.3) [77]. Subsequently, it was found that translocations or deletions at a discrete locus on human chromosome 6 resulted in DRT/DMC. The deletions that caused DRT/DMC on chromosome 6 disrupted a large intergenic ncRNA gene named *ASAR6* [122]. Interestingly, smaller deletions that did not disrupt *ASAR6* did not cause DRT/DMC, indicating that *ASAR6* may play a role in the acquisition of this phenotype. The chromosome 6 deletion analysis identified a 29 kb region on chromosome 6 that must be deleted in order for the chromosome to exhibit DRT/DMC (Figure 1.6) [123]. This 29 kb region contains the promoter and 5´ portion of the *ASAR6* gene (see Figure 1.6).

ASAR6 is a monoallelically-expressed gene that displays asynchronous replication between alleles. One important feature of the *ASAR6* locus is its ability to control gene silencing in *cis*. Deletions on chromosome 6 that disrupt *ASAR6* and cause DRT/DMC also reactivate the previously silenced alleles of monoallelically-expressed genes [122]. Therefore, when *ASAR6* is disrupted,

monoallelically-expressed genes on chromosome 6 become biallelically expressed, indicating that *ASAR6* can function to mediate gene silencing in *cis*.

The process of X chromosome inactivation (XCI) ensures that X-linked gene expression levels are normalized between female (XX) and male (XY) mammalian cells, despite the difference in the number of X chromosomes. To accomplish this, one X chromosome in female cells undergoes transcriptional inactivation, resulting in only one X chromosome that is transcriptionally competent in diploid cells [124]. The silenced allele is referred to as the inactive X chromosome (Xi) and the transcriptionally competent allele is referred to as the active X chromosome (Xa) [125]. One unique feature of the Xi is that the entire chromosome is late replicating [126]. Therefore, following XCI the Xi is transcriptionally silenced and late replicating while the Xa is transcriptionally active and early replicating [129]. XCI is regulated by a region on the X chromosome called the X inactivation center (XIC) [124, 136]. The XIC harbors the *Xist* gene, a ncRNA that has been shown to be both necessary and sufficient for X inactivation [126]. During the establishment of XCI, Xist becomes expressed from the X chromosome destined to become the Xi [126]. This monoallelically-expressed transcript spreads in *cis* along the future Xi and coats the chromosome (forming an Xist RNA “cloud” on the Xi within the nucleus). Before Xist coats the chromosome, it remains bound to the Xi via tethering of the RNA transcript to a nucleation center within the *Xist* DNA by the transcription factor Yin Yang 1 (YY1) [140]. YY1 is thought to bind the repeat F region of *Xist* DNA on the X chromosome and interact with the Xist transcript via the Repeat C

region [140]. Following Xist accumulation around the XIC, the transcript spreads in *cis* along the Xi and coats the chromosome, forming the Xist RNA cloud [140, 141]. The spreading of Xist RNA along the Xi is thought to recruit chromatin modifying complexes and histone variants that repress gene expression and induce heterochromatin [139].The Xi forms a territory in the nucleus called the Barr Body, which contains the heterochromatic portions of the Xi and colocalizes with the Xist RNA cloud. The Barr body is depleted of any gene expression as evidenced by the lack of non-genic repetitive sequence (Cot-1) expression [180].

The XIC was originally identified and mapped by analyzing X/autosome translocations. It was found that translocation derivatives containing the XIC exhibited the spread of transcriptional silencing into autosomal DNA, while derivatives that did not contain the XIC showed no spreading of inactivation into the autosomal fragments [181-183]. These data were further refined with the demonstration that transgenes containing the XIC could induce many aspects of XCI on autosomes. YACs and cosmids containing *Xist* and surrounding elements were randomly integrated into autosomes in mouse ESCs. Upon ESC differentiation, these autosomes recapitulate some aspects of XCI, including Xist expression, Xist RNA coating and cloud formation, gene inactivation, late replication and some establishment of repressive chromatin marks [153-156]. However, these properties were not observed in all clones following differentiation and single-copy arrays of the transgene were not sufficient to induce these properties, as multi-copy integrations were needed [156]. Interestingly, the ability of *Xist* to recapitulate aspects of XCI on autosomes is not

restricted to differentiating ESCs. Multiple labs have shown that *Xist* transgene integration into autosomes is sufficient to establish gene inactivation, Xist RNA cloud formation and repressive chromatin marks in the differentiated cell lines HT1080 and HELA [157-159]. Again, many transgene-integrated autosomes had variable inactivation status and others recapitulated only some aspects of XCI, indicating that additional variables are involved [158, 160]. Nevertheless, one unique feature of *Xist* is its ability to dominantly interfere with the normal replication-timing program of autosomes and induce transcriptional silencing.

Xist shares many physical and functional similarities with *ASAR6*. For example, *ASAR6* and *Xist* represent large ncRNA genes that display random mono-allelic expression, asynchronous replication, and control the expression of other mono-allelic genes in *cis* [122, 184]. In addition, deletion of the *Xist* gene in somatic cells results in a delayed replication phenotype that is similar to the DRT phenotype caused by disruption of *ASAR6* [79, 81, 122]. Thus, the chromosomal phenotypes associated with *ASAR6* disruption are remarkably similar to the phenotypes associated with disruption of *Xist* in adult somatic cells [77, 81, 122]. This indicates that these two genes may have similar functions in mediating chromosomal behavior.

In this chapter, I show that, like *Xist*, ectopic integration of *ASAR6* into an autosome causes a chromosome-wide delay in DNA replication. Similar to what is observed with *Xist* transgenes, this effect is only seen when *ASAR6* is integrated in multi-copy arrays. ASAR6 also forms an ASAR6 RNA “cloud” and silences transcription on the integrated autosome, indicating that ASAR6 RNA

coats the integrated chromosome in *cis* and recruits factors that silence gene expression. When this transgene lacks the 29 kb region encompassing the promoter and 5´ portion of *ASAR6*, it can no longer induce these changes. This is the first demonstration of an autosomal transcript exhibiting functions that were thought to be associated only with Xist. This suggests that autosomes may harbor Xist-like transcripts that function to regulate DNA replication timing and monoallelic expression.

Results

Ectopic Integration of a Human *ASAR6* Transgene Delays Replication Timing of Mouse Chromosomes.

One well-characterized activity of the *Xist* gene is its ability to delay DNA replication timing in *cis* when ectopically integrated into chromosomes (reviewed in [152]). This activity is not restricted to ESCs, as ectopic integration of either human or mouse *XIST/Xist* into the chromosomes of differentiated mammalian cell lines can delay replication of entire chromosomes [154, 157, 158, 185]. Therefore, to determine if *ASAR6* also displays this activity, we tested whether ectopic integration of cloned genomic DNA from the human *ASAR6* locus can cause delayed replication timing of mouse chromosomes.

For this analysis, a bacterial artificial chromosome (BAC) was used that contains ~180 kb of genomic DNA spanning the critical region of the *ASAR6* locus required for DRT (RP11-767E7) (Figure 2.1). As a negative control, a BAC from chromosome 13q14 (RP11-236M15) was chosen based on its relatively low LINE-1 content and lack of reported monoallelic expression [162, 163, 168]. Prior to transfection, the *ASAR6* and control BACs were modified by recombineering [186] to contain a Hygromycin B (Hyg) resistance gene, which allowed for positive selection in mammalian cells. Following recombineering, the BACs were analyzed by PCR and endonuclease digestion to ensure that the Hyg resistance gene was targeted correctly and that no other rearrangements took place within the BAC (Figure 2.2). Mouse C2C12 cells were then transfected with the

modified BAC and subjected to selection in media containing Hygromycin B. Individual clones were isolated and analyzed for BAC integration sites and DNA replication timing of the integrated chromosome (Figure 2.3A). DNA replication timing was assayed using a bromodeoxyuridine (BrdU) “terminal label” assay (Figure 2.3B). This protocol allowed us to visualize the latest replicating portions of each chromosome and observe whether there was a difference in the replication timing of any homologous chromosome pair in the cell. Therefore, when pulse-labeled with BrdU in late S-phase, any chromosome that displays DRT will have more BrdU incorporation than its homolog.

First, we analyzed cells that had the control BAC (RP11-236M15+hyg) stably integrated into an autosome. Figure 2.4 shows the replication timing analysis of a clone containing a multi-copy array of the control BAC integrated into one allele of mouse chromosome 10. Cultures were incubated with BrdU and mitotic cells were harvested, processed for BrdU incorporation and subjected to DNA-fluorescence *in situ* hybridization (FISH) using a mouse chromosome 10 BAC (located near the centromere) plus the control BAC as probes. The DNA-FISH signal from the chromosome 10 centromeric region allowed for the identification of the chromosome 10s, and the presence or absence of the control BAC allowed us to distinguish between the integrated and non-integrated chromosomes, respectively. A comparison of the BrdU incorporation pattern between chromosome 10s in multiple cells indicated that the chromosome containing the control BAC was not delayed in replication timing, as indicated by the similar amounts of BrdU incorporation between the integrated and non-

integrated chromosome 10s (Figure 2.4). This suggested that integration of multiple copies of human BAC DNA into mouse chromosomes, by itself, is not sufficient to cause DRT.

Analysis of a clone that has a multi-copy integration (estimated to be ~20 copies) of the *ASAR6* BAC (RP11-767E7+hyg) into mouse chromosome 3 indicated that *ASAR6* integration could disrupt DNA replication timing. Figure 2.5A-D shows that the *ASAR6* integrated chromosome replicated later than its chromosome 3 homolog, indicating that it was delayed in DNA replication. Delayed replication was also detected in a second clone containing a multi-copy integration of the *ASAR6* BAC into a different mouse chromosome (data not shown), suggesting that multi-copy integration of the *ASAR6* BAC into different mouse chromosomes can result in chromosome-wide delayed replication. Interestingly, analysis of a clone that contained a single-copy *ASAR6* BAC integration revealed that the integrated chromosome did not display DNA replication delay (Figure 2.5E-H). Nevertheless, two out of three ectopic integrations of the *ASAR6* BAC caused a chromosome-wide replication delay in *cis*.

We next wanted to narrow down the region on the *ASAR6* BAC that was responsible for the DRT phenotype. As mentioned previously, work in the Thayer lab had identified an ~29 kb critical region involved in the acquisition of the DRT/DMC phenotype [123]. To determine if this critical region was also required for delayed replication following ectopic integration, we deleted this ~29kb region from the *ASAR6* BAC using recombineering strategies (see Figure 2.1 and

Figure 2.6A-B) and introduced the deleted BAC into mouse chromosomes. Figure 2.6C-F shows the results of the replication timing analysis on a clone containing ~20 copies of the deleted BAC integrated into mouse chromosome 1. This analysis indicated that the integrated chromosome 1 did not display delayed replication. Additional integrations of the deleted BAC into three other mouse chromosomes similarly did not display delayed replication (data not shown). In total, we detected delayed replication in two out of three ectopic integrations of the intact *ASAR6* BAC and in zero out of four ectopic integrations of the deleted BAC. While it is not possible to conclude that the BAC with the ~29kb deletion cannot induce delayed replication upon ectopic integration, especially with the limited number of integrations assayed, our observations suggest that the ~29kb critical region of *ASAR6* is necessary for delayed replication using this ectopic integration assay.

ASAR6 Transgene Integration Results in ASAR6 RNA Cloud Formation on the Integrated Chromosome

The chromosome-wide replication delay induced by ectopic *ASAR6* integration mimicked the replication phenotype that results from ectopic *Xist* integration [81, 122]. We next wanted to see whether *ASAR6* integration could induce any other phenotypes associated with *Xist* and XCI. One identifying feature of Xist is its ability to associate with and spread along the entire X chromosome in *cis* [139]. The chromosome-coating activity of Xist RNA can be observed by RNA-FISH, where the Xist RNA signal accumulates around the Xi

and appears as a ”cloud” within the nucleus (Figure 2.7A). This is in contrast to the small, focal points of hybridization observed for most other RNA transcripts (Figure 2.7B). The Xist RNA cloud can also form on autosomes upon ectopic *Xist* integration, indicating that the Xist transcript retains the ability to coat chromosomes, albeit less efficiently, even when introduced into a different chromosomal context [157, 180]. Despite the demonstration that ASAR6 RNA does not coat chromosome 6 when expressed from its native locus in differentiated fibroblasts or primary blood lymphocytes (PBLs) (Figure 2.7B) [122, 123], we assayed whether ectopic integration of an *ASAR6* transgene resulted in the accumulation of ASAR6 RNA along the integrated chromosome. For this experiment, we used a clonal cell line with a multi-copy *ASAR6* transgene integration into chromosome 3 that displayed delayed replication (see Figure 2.5A-D). RNA/DNA-FISH was used to analyze ASAR6 RNA localization and accumulation. We found that multi-copy ectopic integration of *ASAR6* resulted in the formation of an ASAR6 RNA cloud in the nucleus that co-localized with the integrated chromosome, indicating that ASAR6 can coat the chromosome in *cis* (Figure 2.7C-G). To compare the appearance of the ASAR6 and Xist RNA clouds, RNA-FISH was performed with both an Xist probe and an ASAR6 probe. The two RNA-FISH signals appeared very similar in each cell (Figure 2.7H-J), suggesting that ASAR6 RNA localization and accumulation is regulated in a manner similar to Xist when ectopically integrated into an autosome. Interestingly, RNA-FISH on cells with a multi-copy ectopic integration of the *ASAR6* deletion BAC (RP11-767E7Δ29kb in Figure 2.1) revealed that RNA is still

expressed from this transgene (despite lacking the *ASAR6* promoter), confirming that multiple TSSs are present within this gene. However, the ASAR6 transcript that is expressed from the deletion BAC did not form an ASAR6 RNA cloud (Figure 2.8A-D). A comparison of the signal size of the deletion BAC transcript to that of the intact ASAR6 and Xist signals indicated that it is indeed much smaller in size, similar to that of most other RNA transcripts (Figure 2.8E-F). This indicates that ectopic *ASAR6* integration results in the *cis* association of the ASAR6 transcript with the integrated chromosome and that the 29 kb promoter region is necessary for this effect.

ASAR6 RNA Coating Correlates with Chromosome-wide Gene Inactivation

Another identifying feature of *Xist* is its ability to initiate and direct the silencing of linked gene expression. At the onset of XCI, Xist is expressed from the Xi, coats the chromosome, and recruits repressive histone and DNA methylating complexes, which result in chromosome-wide silencing of gene expression [126, 135]. *Xist* appears to be one of the major components in this process, as this gene has been shown to be both necessary and sufficient for Xi gene silencing [126]. With the demonstration that ASAR6 RNA can coat chromosomes following ectopic integration, we next wanted to determine if the chromosome coating and ASAR6 RNA cloud formation is associated with transcriptional silencing of the integrated chromosome. C2C12 cells are myoblasts that were derived from an inbred mouse strain and contain few, if any, nucleotide polymorphisms that could be used in PCR-based assays to detect

allele-specific expression [187]. Furthermore, assaying gene expression of multiple individual genes is not necessarily indicative of whole-chromosome transcriptional status. Therefore, a method developed by Hall et al. was used to assay global nuclear transcription [157]. This method involves labeling Cot-1 DNA to use as a probe in non-denatured nuclei, thereby detecting expressed repetitive RNA sequences. Cot-1 is comprised of repetitive sequences, such as LINE-1s and small interspersed nuclear elements (SINEs), and is typically used in DNA-FISH as a competitor to block non-specific probe binding to these sequences. Most repetitive sequences in the genome are not autonomously transcribed; however, they reside within the introns of nearly all genes and are therefore present in essentially all pre-messenger RNAs (pre-mRNAs). Thus, by using Cot-1 DNA as a probe to detect repetitive RNA, it is possible to detect global transcription in the nucleus. Furthermore, since chromosomes reside within distinct chromosome territories in the interphase nucleus [188], combining Cot-1 RNA-FISH with other markers allows for the identification of specific chromosome territories that lack transcription.

Analysis of global nuclear transcription in cells that had an ASAR6 RNA cloud indicated that Cot-1 RNA was absent from nuclear regions occupied by ASAR6 RNA (Figure 2.9A-D). Furthermore, quantification of the signal intensity for ASAR6 and Cot-1 RNAs indicated that Cot-1 RNA was depleted in regions that had the ASAR6 signal (Figure 2.9E-I). This suggests that ASAR6 RNA co-localizes with a repressive nuclear compartment. Additionally, since ASAR6 remains associated in *cis* with the integrated chromosome (in this case mouse

chromosome 3), I conclude that the integrated chromosome 3 lies within this repressive nuclear compartment. It is impossible to tell from these data which came first: ASAR6 coating or transcriptional silencing; however, it is known from XCI studies that Xist coating precedes transcriptional inactivation [124]. Furthermore, when the deleted *ASAR6* transgene was analyzed in a similar manner, we found that the deleted RNA did not co-localize with a transcriptionally silenced compartment (Figure 2.10). This suggests that ASAR6 cloud formation is necessary to induce transcriptional silencing of an ectopically integrated chromosome in *cis* and that ASAR6 RNA expression without cloud formation is not sufficient to silence gene expression.

Discussion

Mammalian cells replicate their genomes every cell cycle during a defined replication-timing program. It is clear that the determinants of replication timing are not encoded within the sequence of the origins of replication, but rather the timing of origin firing is dictated by chromosomal location [10, 189]. Recent studies indicate that at least half of the genome is subject to changes in the temporal sequence of DNA replication during development [15]. The current thinking is that replication timing is directly linked to complex higher-order features of chromosome architecture [13, 190]. However, both the mechanisms and the significance of this temporal replication program remain poorly defined.

Previous work indicated that the *ASAR6* and *Xist* genes share many characteristics, including: 1) they both display random monoallelic expression, 2) they both display asynchronous replication that is coordinated with other linked monoallelic genes, 3) disruption of either gene results in delayed replication timing and instability of entire chromosomes in *cis*, 4) disruption of either gene results in the transcriptional activation of the previously silent alleles of linked mono-allelic genes, and 5) both are expressed from the late-replicating allele [79, 81, 122, 123, 176, 184]. Furthermore, another well-characterized activity of the *Xist* gene is the ability to delay replication timing of entire chromosomes upon ectopic integration of cloned genomic DNA [reviewed in [152]]. In this report, it was found that ectopic integration of cloned human genomic DNA containing *ASAR6* had the ability to delay replication of mouse chromosomes. Additionally,

the ability of an *ASAR6* transgene to delay replication at ectopic locations occurred only when multiple copies of the transgene were integrated, which was also observed with *Xist* genomic transgenes [154-157].

One interesting aspect of *ASAR6* and *Xist* is that the loss-of-function and gain-of-function both give the same replication timing phenotype. Deletions at the native loci cause DRT, and ectopic integrations on different chromosomes also cause DRT [79, 81, 122, 152]. At first this seems counterintuitive, but these data may suggest that these loci are multi-functional. For example, on the X chromosome, it is known that a deletion at the *Xist* locus causes chromosome-wide DRT, indicating that the *Xist* locus functions to promote early replication of the X chromosome [79, 81]. However, it is also known that expression of Xist is required for chromosome-wide silencing and late replication [138]. Interestingly, a deletion at the *Xist* locus on the Xa (the X chromosome that does not express Xist) causes DRT, suggesting that the Xist transcript is not involved in promoting early replication [79]. Therefore, it is possible that one function of the *Xist* locus is to produce the Xist transcript, which silences gene expression and delays DNA replication, while another function of this locus is to promote early DNA replication through a transcript-independent mechanism. Thus, the default state of the X chromosome would be to replicate early (referred to as normal replication timing) and this could be dominantly interfered with via the Xist transcript. This idea would be the same for the *ASAR6* locus. The ASAR6 locus may function in a transcript-dependent manner to silence expression of one allele of random monoallelically-expressed genes on chromosome 6 during the

process of random AI and promote late-replication of these specific loci. The other function of this locus would be transcript-independent and promote early chromosome replication via an unknown mechanism.

It was also found that deleting 29 kb from the *ASAR6* transgene negated the effects it had on replication timing. This region spans the presumed promoter of *ASAR6* and also encompasses the 5´ portion of the gene. One limitation of this study is that we compared the phenotypes of cells that had an ectopic integration of the intact *ASAR6* BAC or the deleted BAC into different chromosomal loci. Since the site of integration of these transgenes could not be controlled, it is possible that position-effects of the local chromatin environment were influencing their activity. A possible future study would be to randomly insert a single *lox*P site into a mouse autosome and then integrate different *ASAR6* transgenes (which contain a single *lox*P site) in a Cre-dependent manner. Although this experiment would only allow for single-copy integration of a transgene, Chow et al. have successfully used a similar strategy to assay the function of *Xist*-containing transgenes that are expressed from an inducible promoter [160]. Therefore, it would be possible to compare the activity of multiple different transgenes without having to worry about the effects of different chromatin environments.

Previous studies have indicated that ASAR6 does not coat chromosome 6 in differentiated lymphocytes or fibroblasts [122, 123]. In this study, we found that ectopic integration of an *ASAR6* transgene resulted in ASAR6 RNA cloud formation in the nucleus and accumulation of the transcript around the integrated

chromosome. How can this finding be reconciled with data indicating that ASAR6 does not coat chromosomes? The multi-copy *ASAR6* integration clearly results in overexpression of the transcript. However, the RNA-FISH analysis was only designed to analyze nascent transcripts that are present at the site of transcription and not total RNA levels in the nucleus. Therefore, the accumulation of the ASAR6 RNA transcript at these loci cannot be explained simply by overexpression and suggests that the RNA transcript is being actively retained around the integrated chromosome. These data indicate that ASAR6 has the ability to coat chromosomes, much like the Xist transcript, only it does not do so in at least two differentiated cell types. But why would ASAR6 retain this activity? It is possible that ASAR6 does coat chromosome 6 at some point in development or only in specific cell types. Additional studies will be needed to confirm this possibility.

Interestingly, the RNA transcript transcribed from the *ASAR6* deletion BAC (RP11-767E7-Δ29kb) does not coat the integrated chromosome. This suggests that there is some information encoded in this 29 kb region that is necessary for the coating activity of ASAR6. Interestingly, a similar phenomenon was also observed with *Xist* transgene deletions. Jeon et al. showed that, while ectopic integration of a full-length *Xist* transgene resulted in Xist accumulation and coating, a small deletion in the 5´ end of the gene completely counteracted this activity [140]. It was subsequently found that the deleted region contained three YY1 binding sites, and without these YY1 binding sites within the genomic DNA the Xist transcript was unable to associate with the integrated chromosome in *cis*

[140]. Since YY1 acts as a bridge by binding *Xist* DNA and RNA, it was suggested that deleting the YY1 DNA binding sites resulted in the migration of the Xist transcript to other regions in the nucleus [140, 141]. Therefore, one possibility for why the deleted ASAR6 does not coat chromosomes could be that the DNA-RNA binding protein responsible for ASAR6 coating can no longer bind the DNA, the RNA or both. The protein that facilitates ASAR6 coating is currently not known, however, an analysis of the 29 kb deletion indicates that a YY1 binding site does reside in this region (AACATGGCG at chr6:96334002-96334011 bp [NCBI Build 36/hg18]). Whatever the reason, this 29 kb region is clearly essential for the coating function of ASAR6 and future experiments narrowing this vital region down to a smaller size will be beneficial in identifying its function.

In addition to ASAR6 coating, it was demonstrated that *ASAR6* transgene integration could silence gene expression on the integrated chromosome. This transcriptional silencing was dependent on ASAR6 coating, as indicated by the lack of silencing when ASAR6 RNA did not form a cloud in the nucleus. The easiest explanation for ASAR6-induced silencing is that this transcript has the ability to recruit chromatin- and/or DNA-modifying enzymes in a manner similar to Xist, however further studies will be needed to confirm this possibility. The demonstration that *ASAR6* has the ability to silence gene transcription in *cis* correlates with previous observations that deletion of *ASAR6* resulted in the activation of previously silent alleles of linked monoallelically-expressed genes. Therefore, I propose that the *ASAR6* locus normally functions to silence gene

expression of random monoallelically-expressed genes on chromosome 6. This process would begin early in development, around the time that random AI is occurring, and would result in ASAR6 RNA coating chromosome 6 and initiating silencing of random monoallelically-expressed genes. Following differentiation, once the silencing of one allele is established, ASAR6 would no longer associate in *cis* with chromosome 6. So, unlike XCI, where Xist RNA associates with the Xi throughout all stages of development, ASAR6 RNA associates with chromosome 6 at the very early stages of development and not in adult tissues. Of course, this is not the only function of the *ASAR6* locus. This locus, through some unknown mechanism, also controls chromosome-wide DNA replication timing by promoting early DNA replication. This function is supported by data indicating that chromosomal deletions and translocations at this locus result in DRT/DMC [77, 122].

The DRT/DMC phenotype has been detected on chromosome rearrangements involving many different mammalian chromosomes [74, 77, 80, 81, 122]. Therefore, it seems likely that all mammalian chromosomes contain loci that function to regulate chromosome-wide replication timing of individual chromosomes. Given the similarities in structure and function of the two loci characterized to date, *Xist* and *ASAR6*, I propose that all mammalian chromosomes contain 'inactivation/stability centers' that function to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of individual chromosomes.

Materials and Methods

Recombineering

SW102 cells [191] were electroporated with 4 micrograms (μg) of purified BAC DNA (RP11-236M15 or RP11-767E7) at 1.35 kiloVolts (kV) and 600 ohms with a capacitance of 10 microFarad (μF) using a 0.1 cm gap cuvette. Cells were added to 1 mL Lysogeny Broth (LB), recovered for 1 hour at 30°Celsius (C), and plated on LB/agar plates containing 25 μg/ milliliter (mL) Chloramphenicol (Cam) and 12.5 μg/mL Tetracycline (Tet) for 48 hours at 30°C. Clones were screened for the presence of an intact BAC by restriction enzyme digest and polymerase chain reaction (PCR) and positive clones were pooled (SW102 + BAC cells). To insert a Hygromycin B resistance gene (Hyg^R) into our BAC, we used a counter-selection modification strategy. SW102 cells are normally resistant to Streptomycin (Str). We first introduced DNA into the BAC that conferred Str sensitivity and Kanamycin (Kan) resistance to the cells and then replaced it with DNA containing Hyg^R and an Ampicillan (Amp) resistance gene. Upon Amp resistance, the cells then revert back to Str resistance and Kan sensitivity. 25 mL of LB + 25 μg/mL Cam was inoculated with 500 μL of overnight starter culture containing SW102 + BAC cells and shaken at 30°C for 3 hours. Cells were induced by heat shock at 42°C for 15 min and put on ice. Cells were washed twice with ice-cold distilled H_2O (dH_2O) and resuspended in ~400 μL dH_2O. The PCR product used in the first recombineering step was generated by amplifying

rpsL-neo template DNA (GeneBridges, Dresden, Germany) with the rpsL-neo For (5´- CTTATCGATGATAAGCTGTCAAACATGAGAATTGATCCGGAACCCTTAA TGGCCTGGTGATGATGGCGGGATCG-3´) and rpsL-neo Rev (5´-CCGATGCAA GTGTGTCGCTGTCGACGGTGACCCTATAGTCGAGGGACCTATCAGAAGAAC TCGTCAAGAAGGCG-3´) primers. This PCR product contains 50 base pairs (bp) of homology to the pBACe3.6 vector on each end. Prior to electroporation, the PCR product was digested with DpnI, phenol/chloroform extracted, ethanol precipitated and resuspended in 40 μL dH_2O. 50 μL of induced cells were combined with 2 μL PCR product and electroporated as mentioned above. Cells were plated on LB/agar plates containing 25 μg/mL Cam and 15 μg/mL Kan. Cam + Kan resistant cells were selected for Str sensitivity and correct targeting was confirmed by restriction enzyme digestion and PCR. To replace the rpsL-neo DNA with our Hyg^R gene, we used the protocol outlined above with some modifications. The PCR product used for recombination was generated by amplifying loxP-hygro-amp in pCR2.1 with Hyg For (5´-CCGATGCAAGTGTGTC GCTGTCGACGGTGACCCTATAGTCGAGGGACCTACAGGAAACAGCTATGAC CATG-3´) and Hyg Rev (5´-CTTATCGATGATAAGCTGTCAAACATGAGAATTG ATCCGGAACCCTTAATTGTAAAACGACGGCCAGT-3´) primers. Before induction, cells were grown up in LB + 25 μg/mL Cam and 15 μg/mL Kan and following electroporation cells were plated on LB/agar plates containing 25 μg/mL Cam, 50 μg/mL Str and 25 μg/mL Amp. Correct targeting was confirmed by restriction enzyme digestion and PCR. This made the RP11-236M15+hyg and RP11-767E7+hyg BACs.

To make the 29kb deletion in RP11-767E7, we used a galactokinase (galK) selection method that has been described previously [191]. Briefly, the galK cassette was amplified with the primers galK For (5´-AAGTGTGCACATATG TGTTAGATGAAATATTGAGAAGGAACTTGAGTAAACCCTGTTGACAATTAAT CATCGGCA-3´) and galK Rev (5´-TCATAATATGCATGGTAGGAAGTCTCCAG GAACTGACCCGTATAACAGGATTCAGCACTGTCCTGCTCCTT-3´). We electroporated PCR product into SW102+RP11-767E7+Hyg cells using the protocol described above. Following electroporation, cells recovered for an hour at 30°C and were washed twice with minimal media (M9) and plated on M63 plates. Cultures were left to grow at 32°C for 6-7 days. Correct targeting was confirmed by restriction enzyme digestion and PCR. This deletes chr6:96203250-96232818 (NCBI GRCh37/hg19).

PCR and Restriction Digest to Screen for Targeted Recombinants

BAC DNA was purified from bacteria and subjected to PCR in a 25-50 µl volume using 50-100 nanograms (ng) of DNA, 1x Standard Taq Buffer (New England Biolabs, Inc.), 200 microMolar (µM) each deoxynucleotide triphosphate, 0.2 µM of each primer, and 3 units (U) of Taq DNA Polymerase (New England Biolabs, Inc.) under the following reaction conditions: 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 45 sec, and 72°C for 1 min, with a final extension time of 10 min at 72°C. PCR products were separated on 1% agarose gels, stained with ethidium bromide, and photographed under ultraviolet light illumination. The new junctions formed by Hyg insertion in RP11-236M15+hyg

and RP11-767E7+hyg BACs were amplified using the following primers (used in Figure 2.2A):

Junc. #1 For: 5′-GCATAACCAAGCCTATGCCTAC-3′

Junc. #1 Rev: 5′- GTTATCTACACGACGGGGAGTC-3′

Junc. #2 For: 5′- ATTGACATGTCGTCGTAACCTG-3′

Junc. #2 Rev: 5′- AATTCGCCAATGACAAGACG-3′

+ cont. For: 5′- TCGAGATTTTCAGGAGCTAAGG-3′

+ cont. Rev: 5′- TCATGGAAAACGGTGTAACAAG-3′

The primers used in Figure 2.6B for detection of 29kb deletion:

New junc. For: 5′-AAAAATGTCCTGAAACAGAAAAGAA-3′

New junc. Rev: 5′-GCTGTCGCTGAACAATATGAAG-3′

Deleted region For: 5′-AAAGGCTAGTTTAGTAATTCAGACA-3′

Deleted region Rev: 5′-CTGGTCCTCATCCAGAGCTTAC-3′

Restriction endonuclease digestion was performed according to manufacturer's protocols (New England Biolabs, Inc.) and run on a 1% agarose gel overnight at low voltage, stained with ethidium bromide, and photographed under ultraviolet light illumination.

Transfection of BAC DNA Into Cells

BAC DNA was purified using the PrepEase® BAC purification kit (Usb). Mouse C2C12 cells were plated in 6-well dishes to attain 90-95% confluence the following day. On the day of transfection 100 μL of OPTIMEM (Gibco) was mixed with 1 μg uncut BAC DNA and another 100 μL OPTIMEM was mixed with

4 μL Lipofectamine 2000 (Invitrogen). Tubes were incubated at room temperature for 5 minutes, mixed together and incubated another 30 min. 2 mL of DMEM plus 10% fetal bovine serum (Hyclone) was added to each well in addition to the 200 μL of DNA:lipofectamine complexes. Cells were incubated at 37°C and, 48-72 hours post-transfection, each well was split into a 15 cm dish containing DMEM + 10% serum + 100 μg/mL Hygromycin B (Calbiochem). Hygromycin resistant clones were isolated and screened for the presence of BAC DNA by PCR and DNA-FISH.

DNA-FISH

Trypsinized cells were centrifuged at 1,000 rpm for 10 minutes in a swinging bucket rotor. The cell pellet was re-suspended in 75 milliMolar (mM) potassium chloride for 15-30 minutes at 37°C, re-centrifuged at 1,000 rpm for 10 minutes and fixed in 3:1 methanol:acetic acid. Fixed cells were added drop-wise to microscope slides to generate mitotic chromosome spreads using standard methods [192]. Slides with mitotic spreads were baked at 85°C for 20 minutes and then treated with 0.1 mg/ml RNAase for 1 hour at 37°C. After RNAase treatment, the slides were washed in 2xSSC (1xSSC is 150 mM NaCl and 15 mM sodium citrate) with 3 changes for 3 minutes each and dehydrated in 70%, 90%, and 100% ethanol for 3 minutes each. The slides were denatured in 70% formamide in 2xSSC at 70°C for 3 min followed by dehydration in ice cold 70%, 90% and 100% ethanol.

BAC and Fosmid probes: Mitotic chromosome spreads were prepared as described above. BAC and Fosmid (F-factor Cosmid cloning vectors) DNAs were nick-translated using standard protocols to incorporate biotin-11-dUTP or digoxigenin-dUTP (Invitrogen). BAC and Fosmid DNAs were directly labeled with Cy3-dUTP, FITC-dUTP, Spectrum Orange-dUTP or Spectrum Green-dUTP (Vysis, Abbott Laboratories) using nick-translation or random priming using standard protocols. Final probe concentrations varied from 40-60 ng/*u*l. BAC or Fosmid probes were denatured at 75^{O}C for 10 minutes and prehybridized at 37^{O}C for 30 minutes. Probes were applied to slides and incubated overnight at 37^{O}C. Post-hybridization washes consisted of three 3-minute rinses in 50% formamide/2XSSC, three 3-minute rinses in 2XSSC, and finally three 3-minute rinses in PN buffer (0.1M Na2HPO4 + 0.0M NaH2PO4, ph 8.0, +2.5% Nonidet NP-40), all at 45^{O}C. Slides were then counterstained with either propidium iodide (2.5ug/ml) or DAPI (15ug/ml) and viewed under UV fluorescence (Olympus).

RNA/DNA-FISH

Cells were plated on microscope slides treated with concanavalin A (Sigma) at ~50% confluence and incubated for 4 hours in complete media in a 37^{o}C humidified CO_2 incubator. Slides were rinsed 1 time with sterile RNase free PBS. Slides were incubated for 30 seconds in CSK buffer (100mM NaCl, 300mM Sucrose, 3mM $MgCl_{2,}$10mM Pipes, ph 6.8), 5 minutes in CSK buffer plus 0.1% Triton X-100, and then for an addition 30 seconds in CSK buffer at room temperature. Cells were fixed in 4% paraformaldehyde in PBS for 10 minutes at

room temperature. Slides were rinsed in 70% ETOH and stored in 70% ETOH at 4° C until use. Just prior to use, slides were dehydrated through an ETOH series (70%, 90% and 100%) and allowed to air dry. Denatured probes were prehybridized with Cot-1 DNA at 37° C for 30min. Slides were hybridized at 37°C for 14-16 hours. Slides were washed as follows: 3 times in 50% formamide/2xSSC at 42° C for 5 minutes, 3 times in 2xSSC at 42° C for 5 minutes, 3 times in 4xSSC/0.1% Tween 20 at room temperature for 3 minutes. Slides were then fixed in 4% paraformaldehyde in PBS for 5 minutes at room temperature, and briefly rinsed in 2xSSC at room temperature. The slides were then dehydrated in 70%, 90% and 100% ETOH, and then processed for DNA-FISH, including the RNAase treatment step, as described above. Slides were then counterstained with either propidium iodide (2.5ug/ml) or DAPI (15ug/ml) and viewed under UV fluorescence (Olympus). Z-stack images were generated using a Cytovision workstation.

Cot-1 RNA-FISH

Cot-1 RNA-FISH was carried out exactly as outlined in the above RNA-FISH protocol except that the nick translation reaction was only allowed to go for 90 minutes as opposed to the regular overnight incubation. Furthermore, the prehybridization step following probe denaturation was omitted. Denatured probe was directly added to slides.

Replication Timing Assay

The BrdU replication timing assay was performed on exponentially dividing cultures as follows: asynchronously growing C2C12 cells were exposed to 20 ug/ml of BrdU (Sigma) for 2, 2.5, or 3 hours [193]. Mitotic cells were harvested in the absence of colcemid, treated with 75 mM KCl for 15-30 minutes at 37°C, fixed in 3:1 methanol:acetic acid and dropped on wet, ice-cold slides. The chromosomes were denatured in 70% formamide in 2xSSC at 70°C for 3 minutes and processed for DNA-FISH, as described above. The incorporated BrdU was then detected using a FITC-labeled anti-BrdU antibody (Becton Dickinson). Slides were stained with propidium iodide (0.3 mg/ml), cover slipped, and viewed under UV fluorescence.

All images were captured with an Olympus BX Fluorescent Microscope using a 100X objective, automatic filter-wheel and Cytovision workstation. Individual chromosomes were identified by G-banding or hybridization with BACs. Utilizing the Cytovision workstation, each chromosome was isolated from the metaphase spread and a line drawn along the middle of the entire length of the chromosome. The Cytovision software was used to calculate the pixel area and intensity along each chromosome for each fluorochrome occupied by the DAPI and BrdU (FITC) signals. The total amount of fluorescent signal was calculated by multiplying the average pixel intensity by the area occupied by those pixels.

Figures

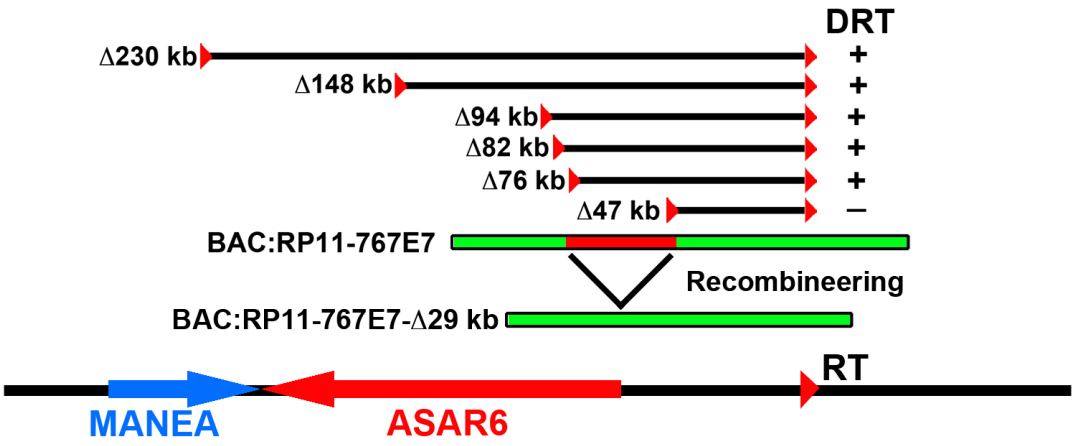


Figure 2.1: Schematic representation of the *ASAR6* BAC used for integration into mouse chromosomes. The approximate locations of the *lox*P integration sites (red triangles) in P175 cells, deletions [in kilobases (Δkb)], and the location of the *ASAR6* BAC (RP11-767E7, green and red) are indicated. The *ASAR6* BAC spans half of the *ASAR6* gene and past the *lox*P-3´RT cassette. The *ASAR6* BAC was modified by recombineering to contain a deletion of the ~29 kb (red) critical region identified by the deletion analysis (BAC: RP11-767E7-Δ29kb BAC) [123]. Figure 2.1 is adapted from [123].

Figure 2.2: Confirmation of Hyg insertion into BACs. A) Junction PCR for the detection of correct Hyg plasmid targeting. Non-recombined BAC DNA (RP11-767E7) and recombined BAC DNA (RP11-767E7+hyg) were subjected to PCR with primers that amplify the two new junctions formed by targeted Hyg insertion (Junc. 1 and Junc. 2). A positive control was included that amplifies a region on the BAC vector backbone that was not affected by recombineering (+ cont). Note that only the recombined BAC contains the new junctions formed by Hyg insertion. The band in the RP11-767E7 lane for the Junc. 2 primer set is non-specific amplification. **B**) Restriction endonuclease digestion of non-recombined RP11-767E7 and RP11-236M15 or recombined RP11-767E7+hyg and RP11-236M15+hyg DNA. DNA was digested either with KpnI or XhoI. Digestion of recombined BACs with KpnI gives a band at 2.5 kb while non-recombined BACs do not have this band (lower blue arrow). Digestion of recombined BACs with XhoI gives a band at 8 kb while non-recombined BACs do not have this band (upper blue arrow). Marker (M) sizes are shown in kb to the left of the DNA ladder. Note that aside from the two new bands, the digestion pattern is identical between the recombined and non-recombined BACs, indicating that the recombined BACs have not undergone any further rearrangements. (Figure on next page).

Figure 2.2: Confirmation of Hyg insertion into BACs.

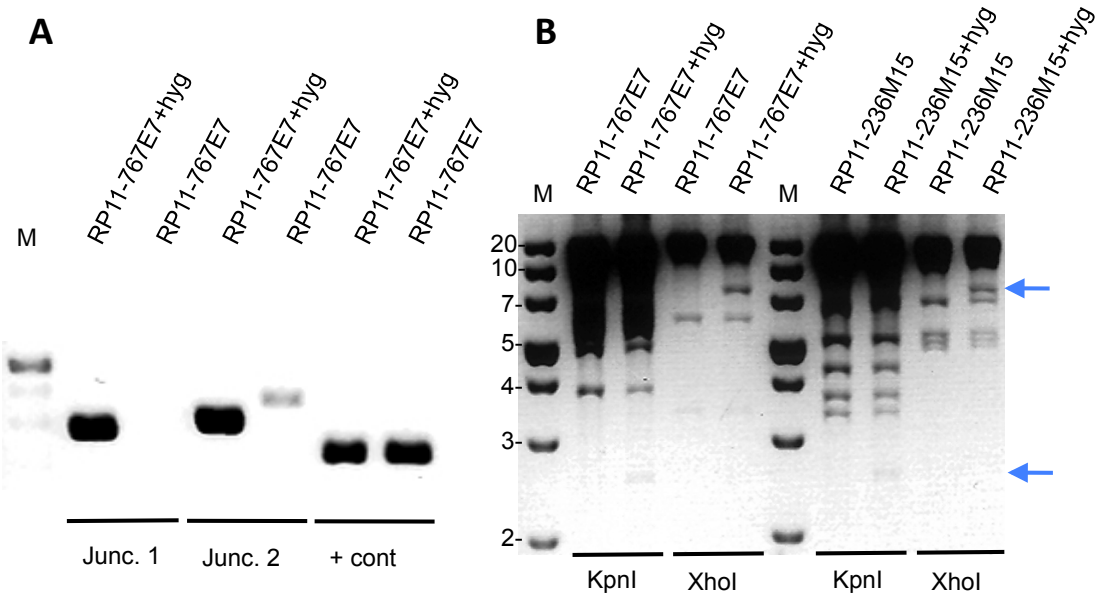


Figure 2.3: Diagram of ectopic integration experiment and BrdU terminal label assay. A) Before ectopic integration of BAC DNA into cells, the BAC (red ring) was first modified to contain a gene mediating resistance to Hygromycin B (hygro). Modified BAC DNA was purified and transfected into C2C12 cells by lipofection and DNA replication timing was analyzed. An example is shown of two homologous mouse chromosomes (blue), one containing an integration of the BAC DNA (red band) and one that has not been modified. **B)** BrdU terminal labeling procedure. Normal chromosomes (black line) replicate within the confines of S phase, while DRT/DMC chromosomes (red line) replicate later in the cell cycle. To analyze the replication timing difference, BrdU is added to the media of asynchronously growing cells for a specific amount of time (green arrow) and cultures are harvested for mitotic spreads and processed for DNA-FISH and BrdU incorporation. All cells that are in mitosis at the time of harvest will have incorporated BrdU in late S phase and the differences in replication timing can be visualized by differences in BrdU staining. Panel **B** was adapted from [122]. (Figure on next page).

Figure 2.3: Diagram of ectopic integration experiment and BrdU terminal label assay.

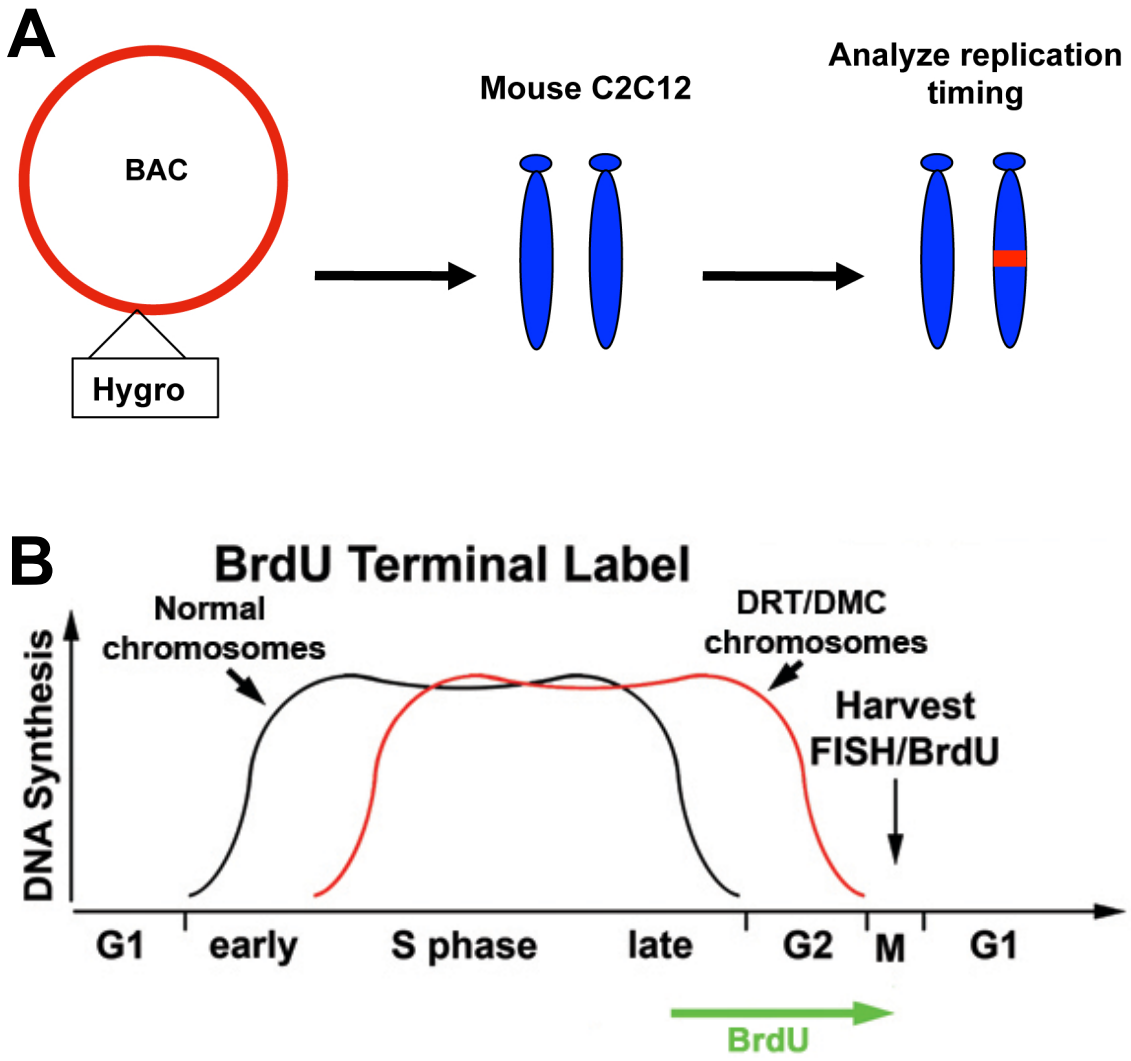


Figure 2.4: Multi-copy ectopic integration of the control BAC does not delay DNA replication. Cells containing a multi-copy array of the control BAC integrated into mouse chromosome 10 were incubated with BrdU for 2 hours, harvested for mitotic cells, and processed for DNA-FISH using the control BAC (RP11-236M15) plus a mouse chromosome 10 BAC (RP24-61D2) as probes (both in red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The chromosome 10s in panel **A** are indicated by arrows. **B**) The chromosome 10s from panel **A** were cut out and aligned with each color displayed separately or in combination. The top red signal is from the chromosome 10 BAC and the lower red signal from chromosome A is from the control BAC. **C**) Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the chromosome 10s from panel **B** are shown. Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome, with the centromere at the beginning of the X-axis and the telomere at the end of the X-axis. Note that both chromosomes contain comparable BrdU incorporation. **D**) Quantification of the BrdU incorporation in seven different cells. The blue bars represent the chromosome containing the control BAC and the red bars represent the normal chromosome 10s. Some of the cells in this clone contained three chromosome 10s. The values represent the total number of pixels (area x intensity) x 1000. (Figure on next page).

Figure 2.4: Multi-copy ectopic integration of the control BAC does not delay DNA replication.

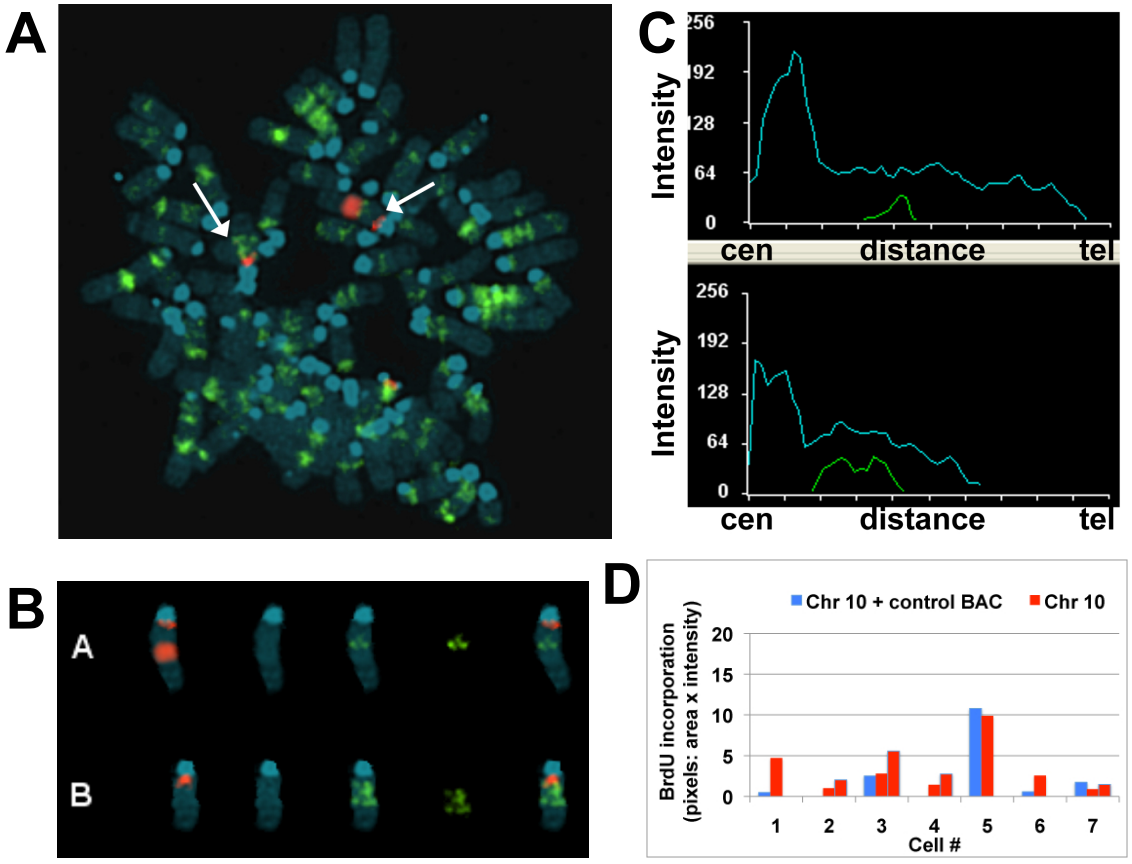


Figure 2.5: Multi-copy ectopic integration of *ASAR6* results in delayed replication of mouse chromosomes. A-D) Cells containing a multi-copy array of the *ASAR6* BAC integrated into mouse chromosome 3 were incubated with BrdU for 2.5 hours, harvested for mitotic cells, and processed for DNA-FISH using the *ASAR6* BAC (RP11-767E7) plus a mouse chromosome 3 BAC (RP23-430A13) as probes (both in red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The chromosome 3s in panel **A** are indicated by arrows. Panels **A-D** were adapted from [123]. **B**) The chromosome 3s from panel **A** were cut out and aligned with each color displayed separately or in combination. The top red signal on both chromosomes is from the chromosome 3 BAC and the lower red signal from one of the chromosomes is from the *ASAR6* BAC. **C**) Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the chromosome 3s from panel **B** are shown. Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome, with the centromere at the beginning of the X-axis and the telomere at the end of the X-axis. Note that the chromosome 3 with the multi-copy *ASAR6* BAC integration contains much more BrdU incorporation, indicating that it displays delayed replication. **D**) Quantification of the BrdU incorporation in six different cells. The blue bars represent the chromosome containing the *ASAR6* BAC and the red bars represent the normal chromosome 3s. Most of the cells in this clone contained three chromosome 3s. The values represent the total number of pixels (area x intensity) x1000. **E-H**) Cells containing a single-copy of the *ASAR6* BAC integrated into mouse chromosome

16 were incubated with BrdU for 2.5 hours, harvested for mitotic cells, and processed for DNA-FISH using the *ASAR6* BAC as a probe (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The chromosome 16s in panel **E** are indicated by arrows. **F**) The chromosome 16s from panel **E** were cut out and aligned with each color displayed separately or in combination. Each chromosome 16 was identified by G-banding so no mouse chromosome 16 probe was needed. **G**) Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the chromosome 16s from panel **F** are shown. Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome, with the centromere at the beginning of the X-axis and the telomere at the end of the X-axis. Note that both chromosome 16s contain comparable BrdU incorporation. **H**) Quantification of the BrdU incorporation in six different cells. The red bars represent the chromosome containing the *ASAR6* BAC and the blue bars represent the normal chromosome 16. The values represent the total number of pixels (area x intensity) x1000. (Figure on next page).

Figure 2.5: Multi-copy ectopic integration of *ASAR6* results in delayed replication of mouse chromosomes.

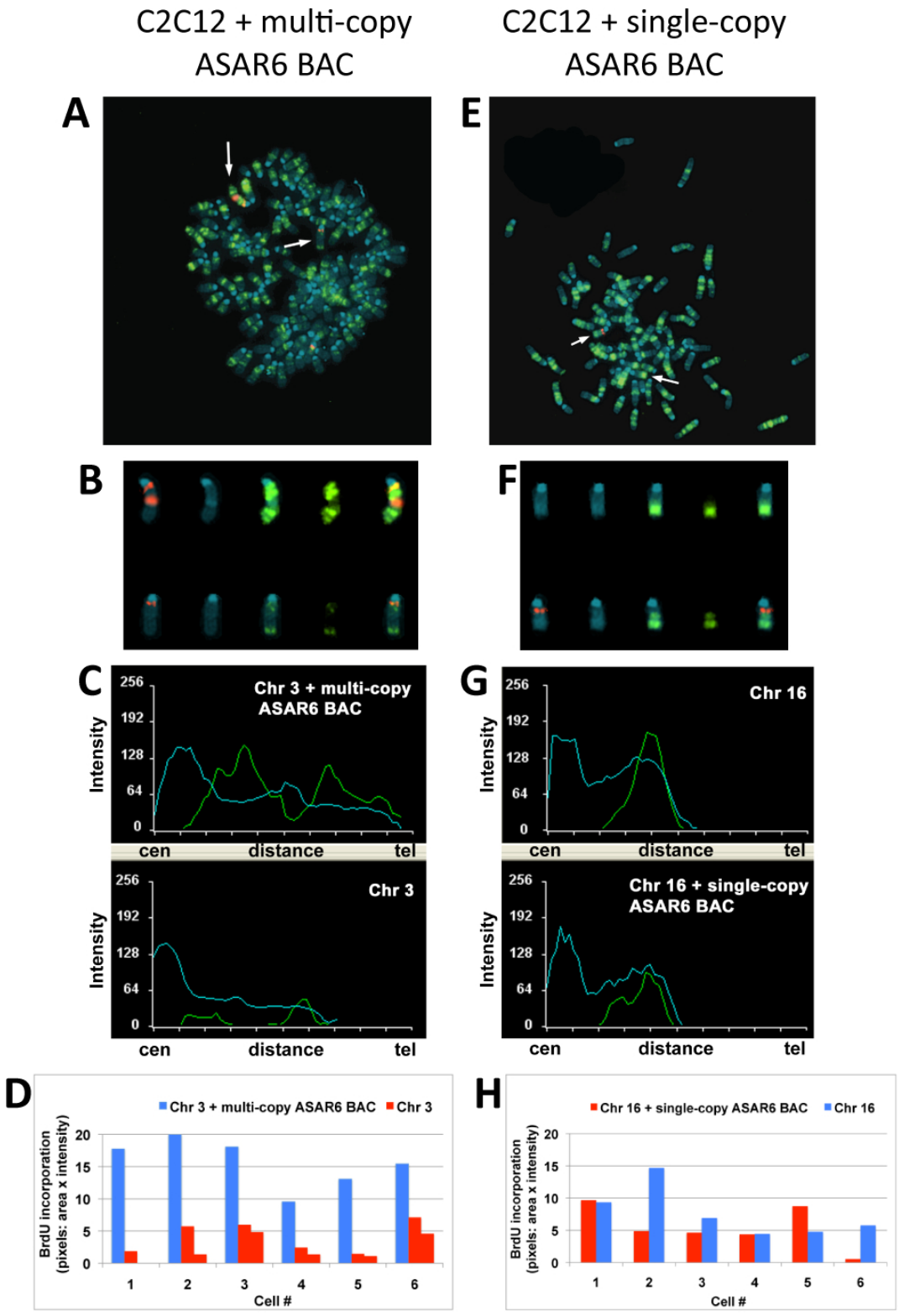


Figure 2.6: Generation and ectopic integration of the *ASAR6* deletion BAC. **A**) Restriction endonuclease digestion of the deletion BAC (RP11-767E7Δ29kb) and the non-deleted RP11-767E7+RpsL-neo. RP11-767E7+RpsL-neo is an intermediate BAC that was constructed before the RP11-767E7+hyg BAC (see methods). DNA was digested either with PmeI (first two lanes) or FspI (last two lanes). Digestion of the deleted BAC with PmeI gives a band at 4 kb while the non-deleted BAC does not have this band. Digestion of the deleted BAC with FspI gives a band at 6 kb and 2.5 kb while the non-deleted BAC does not have these bands. Marker (M) sizes are shown in kb to the left of the DNA ladder. Note that aside from the new bands, the digestion pattern is identical between the deleted and non-deleted BACs, indicating that the deleted BAC has not undergone any further rearrangements or truncations. **B**) Junction PCR for the detection of correctly targeted deletion. Non-deleted BAC DNA (RP11-767E7+RpsL-neo) and deletion BAC DNA (RP11-767E7Δ29kb) were subject to PCR with primers that amplify the new junction formed by deletion of 29 kb from the BAC (New junc.). A negative control was included that amplifies a region within the 29 kb deletion (Deleted region). Note that only the deletion BAC contains the new junction formed by truncation. **C-F**) Cells containing a multi-copy array of the *ASAR6* BAC containing an ~29kb deletion (see Figure 2.1, RP11-767E7-Δ29 kb) integrated into mouse chromosome 1 were incubated with BrdU for 3.0 hours, harvested for mitotic cells, and processed for DNA-FISH using the *ASAR6* BAC (RP11-767E7) plus a mouse chromosome 1 BAC (RP23-34K7) as probes (both in red), and for BrdU incorporation using an antibody

against BrdU (green). The DNA was stained with DAPI (blue). The chromosome 1s are indicated with arrows. **D**) The chromosome 1s from panel **C**, were cut out and aligned with each color displayed separately or in combination. The red signal near the centromere on both chromosomes is from the chromosome 1 probe and the red signal near the telomere on one of the chromosomes is from the *ASAR6* BAC **E**) Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the chromosome 1s from panel **D** are shown. Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome, with the centromere at the beginning of the X-axis and the telomere at the end of the X-axis. Note that the chromosome 1 containing the *ASAR6* deletion BAC shows similar BrdU incorporation as the normal chromosome 1. **F**) Quantification of the BrdU incorporation in six different cells. The blue bars represent the chromosome containing the *ASAR6* deletion BAC and the red bars represent the normal chromosome 1s. The values represent the total number of pixels (area x intensity) x1000. Panels **C-F** were adapted from [123]. (Figure on next page).

Figure 2.6: Generation and ectopic integration of the *ASAR6* deletion BAC.

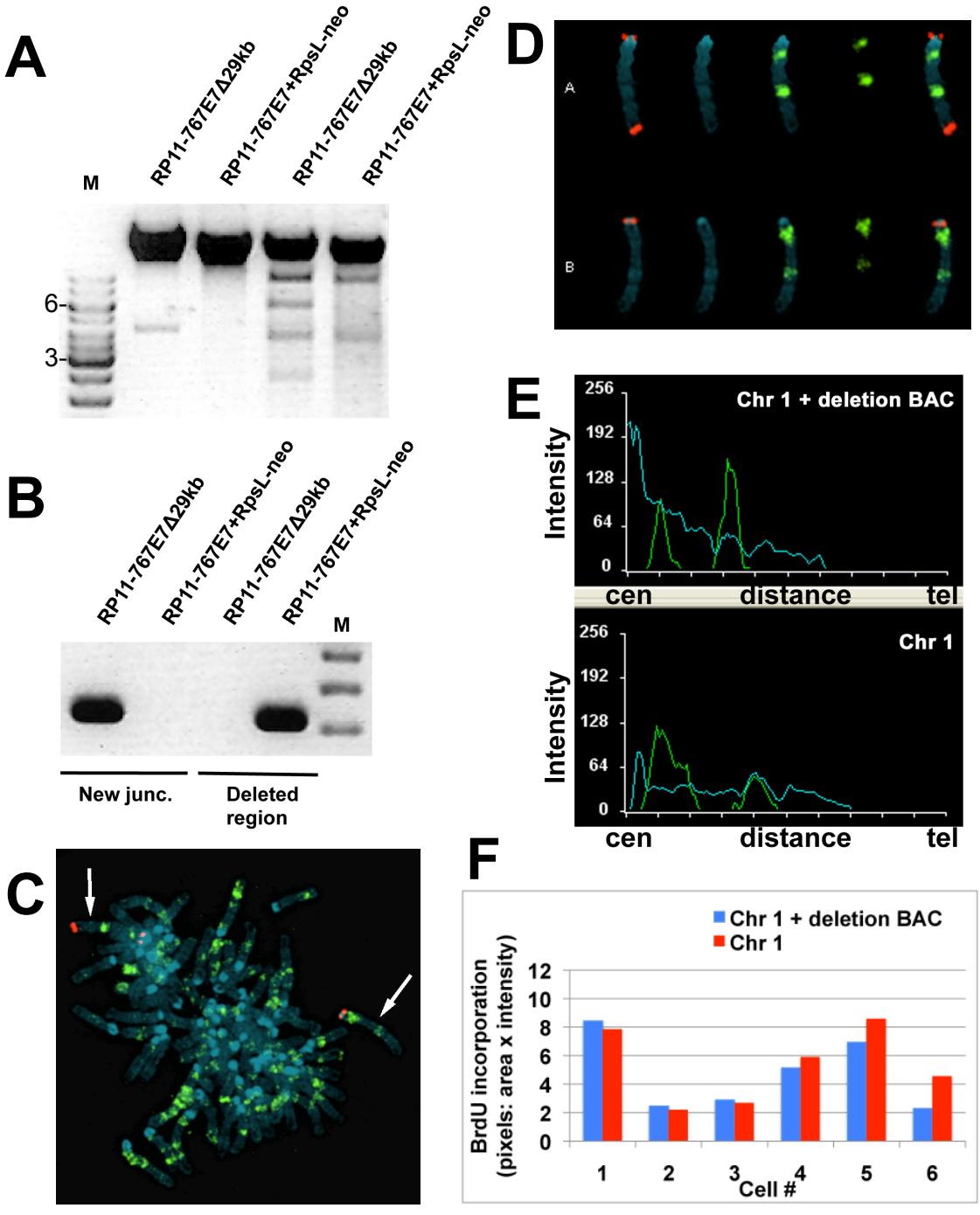


Figure 2.7: Ectopic integration of *ASAR6* results in RNA accumulation around the integrated chromosome. A-B) RNA/DNA-FISH for expression of Xist or ASAR6. C2C12 cells (**A**) and peripheral blood lymphocytes (**B**) were subjected to RNA-FISH (green) using a cocktail of Xist fragments (provided by Dr. York Marahrens) (**A**) and a Fosmid probe that lies within *ASAR6* (G248P86031A6 in **B**). Slides were subsequently re-fixed and processed for DNA-FISH (red) using BAC RP23-95N14 located within the XIC (**A**) and a BAC near the *ASAR6* locus (RP11-959I6 in **B**). Nuclear DNA was detected with DAPI (blue). Panel **B** was adapted from [123]. Note the differences in size of the RNA (green) signal in panels **A** and **B**. **C-G**) RNA/DNA-FISH for expression of ASAR6 in C2C12 cells containing a multi-copy integration of the *ASAR6* transgene into mouse chromosome 3. Cells were subjected to RNA-FISH (green) using Fosmid G248P83419A4 to detect ASAR6 RNA and subsequently re-fixed and processed for DNA-FISH (red) using BAC RP11-48G17 to detect the transgene. **E**) Magnified image from panel **D** with the ASAR6 RNA and DNA signals merged. **F-G**) Same image as panel **E** but with either the ASAR6 RNA signal alone (**F**) or the DNA signal alone (**G**). **H-J**) RNA/RNA-FISH with the same probe used in panel **A** to detect Xist RNA (red) and the same probe in panel **C** to detect ASAR6 RNA (green). The X chromosome is polyploid in these cells, which is why there are multiple Xist RNA signals. Note the similarities in size between the two RNA signals. (Figure on next page).

Figure 2.7: Ectopic integration of *ASAR6* results in RNA accumulation around the integrated chromosome.

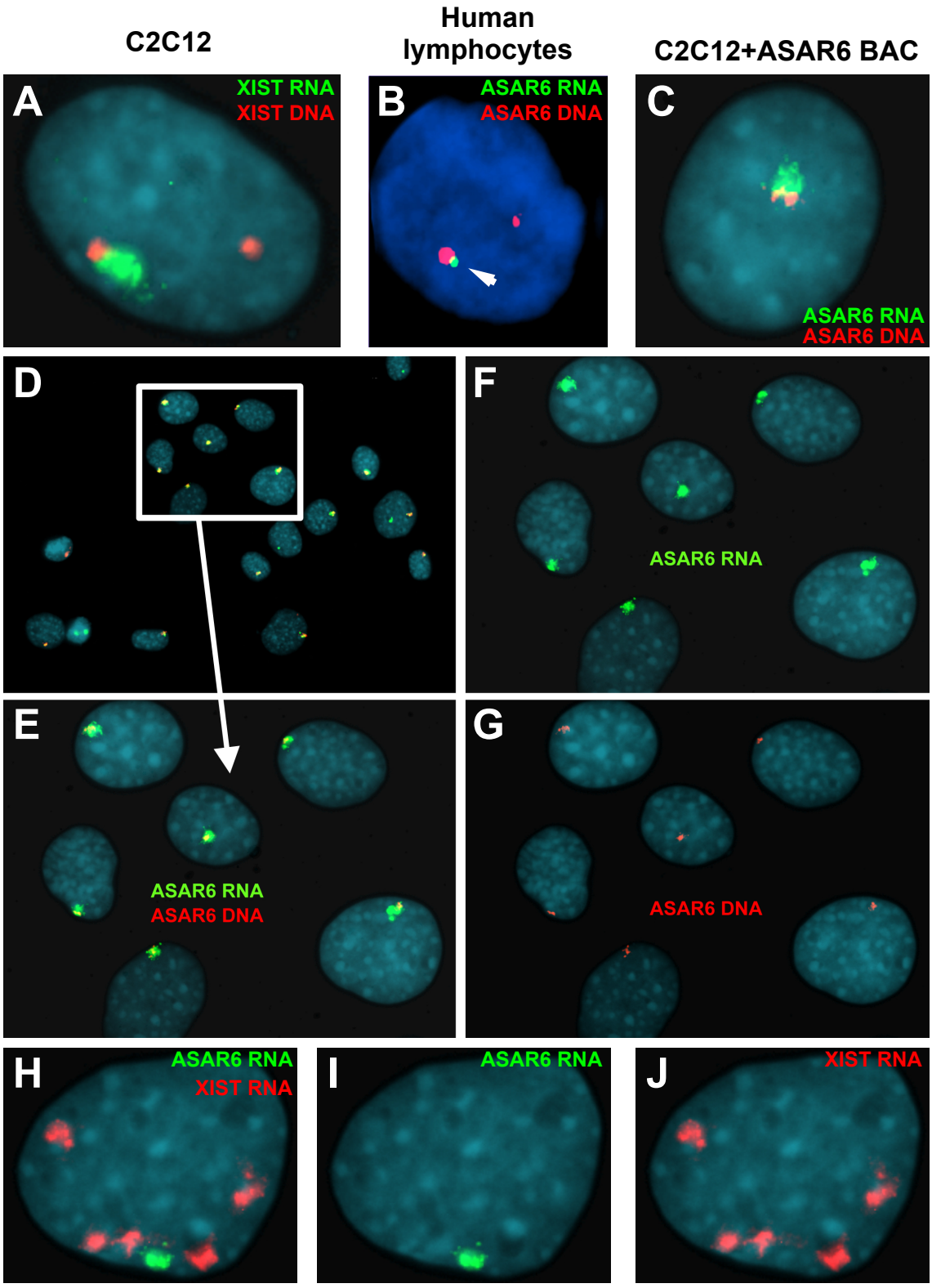


Figure 2.8: Deletion of 29kb surrounding the *ASAR6* promoter inhibits RNA accumulation and cloud formation. A-D) RNA/DNA-FISH for expression of ASAR6 in C2C12 cells containing a multi-copy integration of the deleted *ASAR6* BAC (RP11-767E7-Δ29kb in Figure 2.1) in mouse chromosome 1. Cells were subjected to RNA-FISH (green) using Fosmid G248P83419A4 to detect ASAR6 RNA and subsequently re-fixed and processed for DNA-FISH (purple) using BAC RP11-48G17 to detect the transgene. **B**) Magnified image from panel **A** with the ASAR6 RNA and DNA signals merged. **C-D**) Same image as panel **B** but with either the *ASAR6* DNA signal alone (**C**) or the RNA signal alone (**D**). Note how the ASAR6 RNA signal is very small in size and does not accumulate at a specific site in the nucleus. We observed these small, focal signals in most cells; however, occasionally some cells would have a larger, "cloudlike" signal (arrow in panel **A**). **E**) Xist and ASAR6 expression in cells containing an ectopic integration of the deleted *ASAR6* BAC. RNA/RNA-FISH with same probe used in Figure 2.7A to detect Xist RNA (red) and the same probe in panel **A** to detect ASAR6 RNA (green). Note the differences in size between the two RNA signals. **F**) Size of the Xist RNA signal in C2C12 cells (first bar), the ASAR6 RNA signal in C2C12 cells containing a multi-copy integration of the intact *ASAR6* BAC in mouse chromosome 3 (second bar), and the ASAR6 signal in C2C12 cells containing a multi-copy integration of the deleted *ASAR6* BAC in mouse chromosome 1 (third bar). Fifty RNA signals were scored for each cell line and the average RNA signal area, in pixels, was calculated on the Y-axis. Note the difference in signal

size of the deleted ASAR6 RNA signal compared to the intact ASAR6 RNA and Xist RNA signal. (Figure on next page).

Figure 2.8: Deletion of 29kb surrounding the *ASAR6* promoter inhibits RNA accumulation and cloud formation.

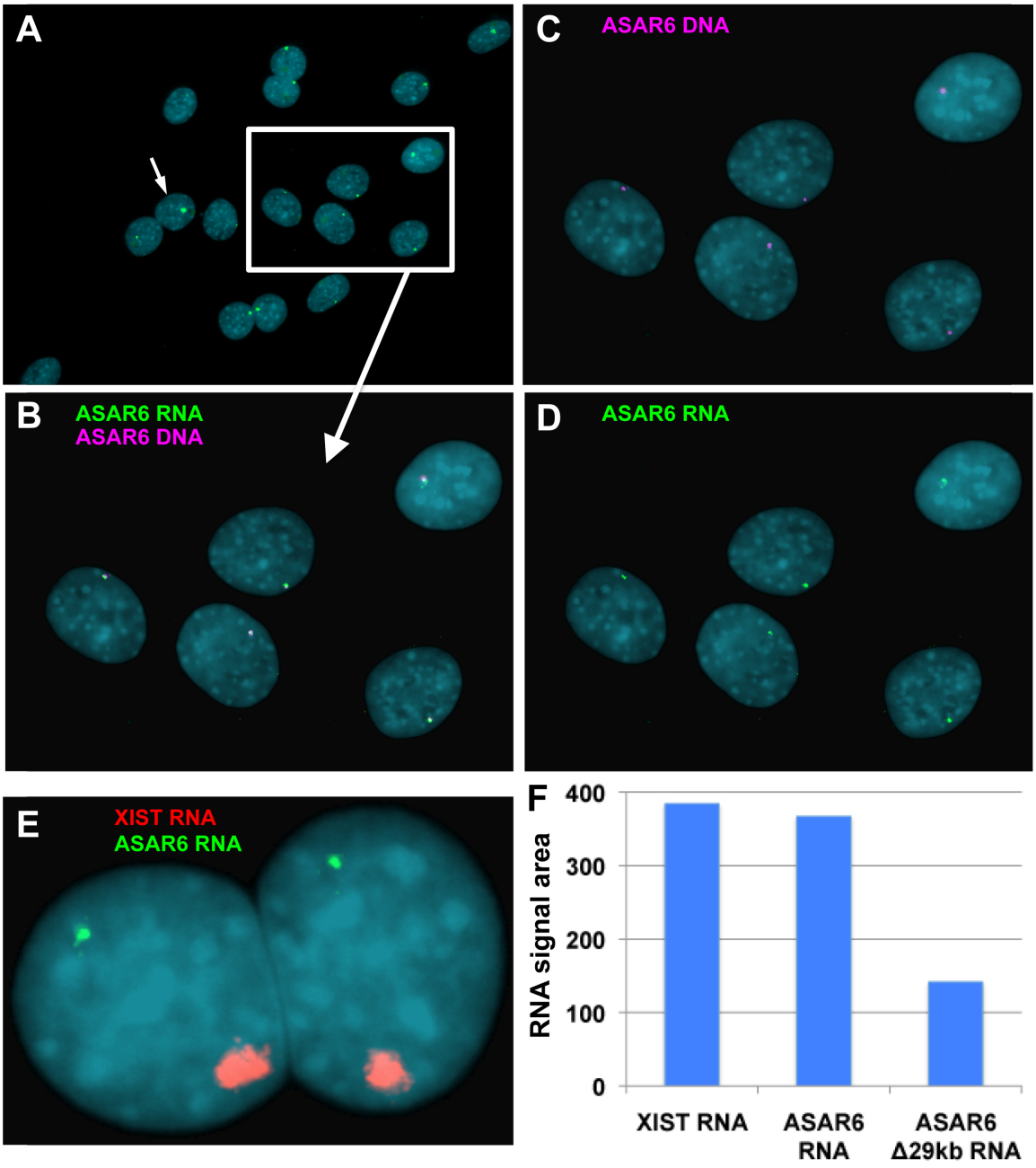


Figure 2.9: Ectopic integration of *ASAR6* induces transcriptional silencing of the integrated chromosome. A-I) RNA-FISH for expression of ASAR6 and Cot-1-containing RNA in C2C12 cells containing a multi-copy integration of the intact *ASAR6* transgene in mouse chromosome 3. Cells were subjected to RNA-FISH using Fosmid G248P83419A4 to detect ASAR6 RNA (red) and Cot-1 DNA (Invitrogen) to detect pre-mRNAs (green). The DNA was stained with DAPI (blue). **A-D**) Example of cells stained for ASAR6 (red), Cot-1 (green) and/or DAPI (blue). Notice that the areas containing ASAR6 RNA are depleted of Cot-1 RNA (two arrowheads). The chromosome 3 containing the *ASAR6* transgene is duplicated in these two cells so there are two ASAR6 RNA signals instead of one. **E-I**) Example of a cell stained for ASAR6, Cot-1 and/or DAPI. A line was drawn through the area of the cell that contained an ASAR6 signal (a-b). **I**) Quantification of the intensity of the RNA signal for ASAR6 (red), Cot-1 (green) and DAPI (blue) along the white line from point a to point b. Signal intensity is on the Y-axis and distance along the line from a to b is labeled on the X-axis. Notice that the spike in intensity for ASAR6 coincides with a drop in intensity for Cot-1 RNA. (Figure on next page).

Figure 2.9: Ectopic integration of *ASAR6* induces transcriptional silencing of the integrated chromosome.

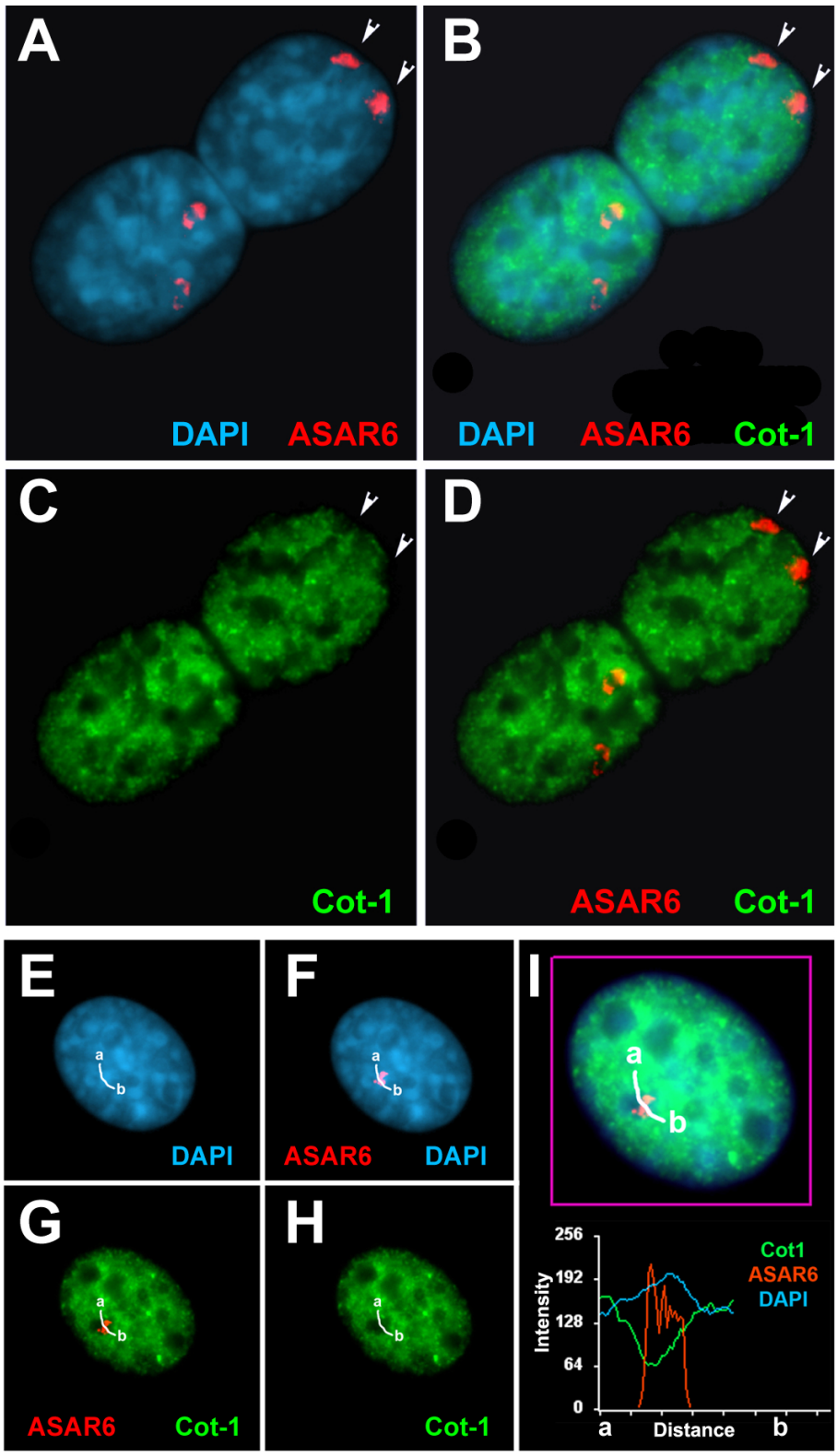
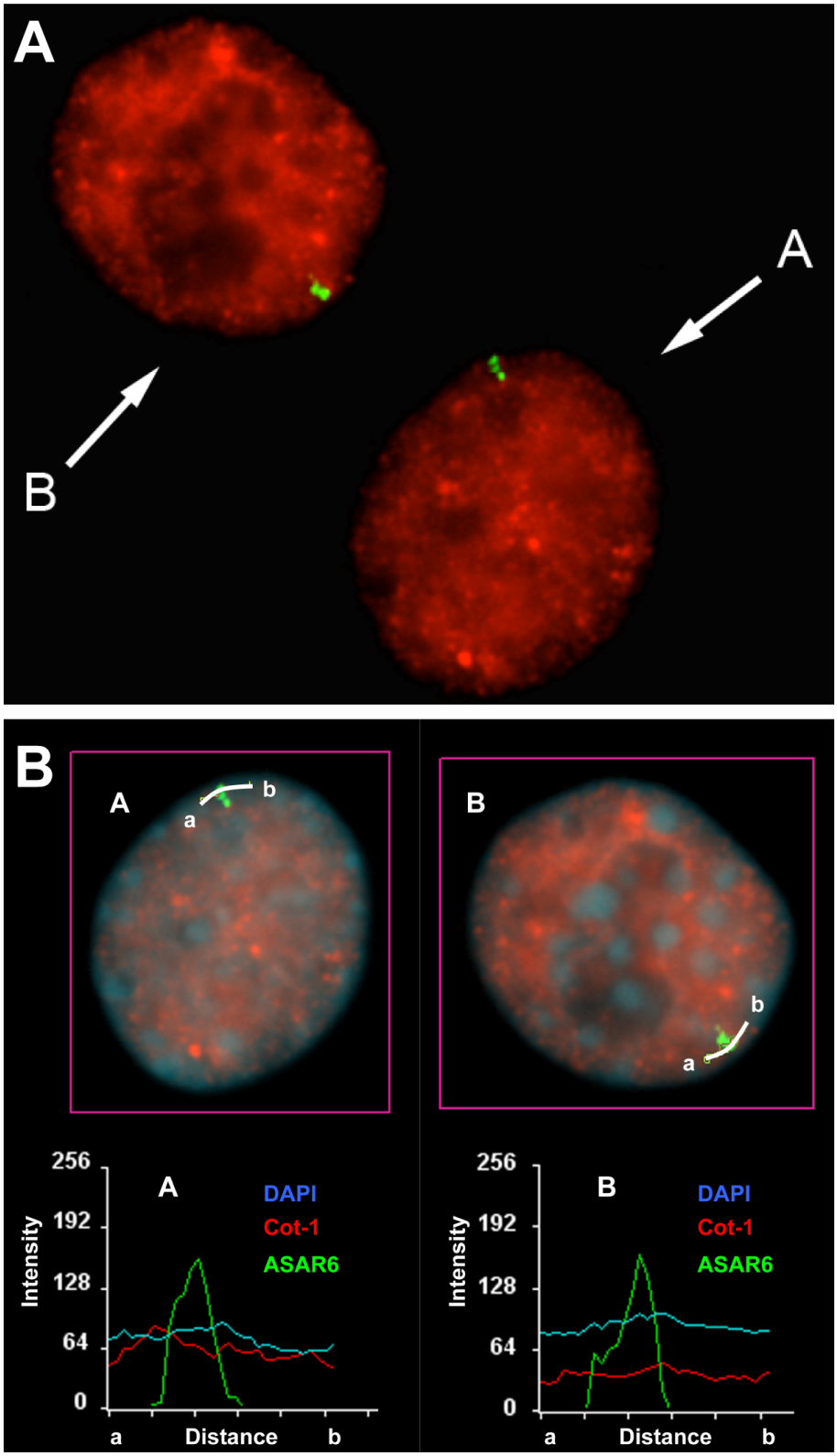


Figure 2.10: ASAR6 cloud formation is necessary for transcriptional silencing. A) RNA-FISH for expression of ASAR6 and Cot-1-containing RNA in C2C12 cells containing a multi-copy integration of the deleted *ASAR6* BAC (RP11-767E7-Δ29kb) in mouse chromosome 1. Cells were subjected to RNA-FISH using Fosmid G248P83419A4 to detect ASAR6 RNA (green) and Cot-1 DNA (Invitrogen) to detect pre-mRNAs (red). The two cells (A and B) are indicated with arrows. **B**) Analysis of cells A and B from panel **A**. A line was drawn through the area of each cell that contained an ASAR6 signal (a-b). The RNA signal intensity was quantified for each cell along the white line from point a to point b and graphed below. Signal intensity is on the Y-axis and distance along the line from a to b is labeled on the X-axis. Notice that the spike in intensity for ASAR6 does not coincide with a drop in intensity for Cot-1 RNA. (Figure on next page).

Figure 2.10: ASAR6 cloud formation is necessary for transcriptional silencing.



CHAPTER THREE

Identification and Characterization of a New Inactivation/Stability Center on Chromosome 15q24

Introduction

The Thayer lab has developed a chromosome engineering system that allows for the systematic analysis of human chromosomes with DRT/DMC [75, 77, 80, 122]. This system relies on site-specific recombinases to generate precise chromosomal rearrangements. Using this system, the lab previously identified four balanced translocations, each displaying DRT/DMC on one of the two derivative chromosomes (Figure 1.3) [77]. Subsequently, it was found that translocations or deletions at a discrete locus on human chromosome 6 resulted in DRT/DMC. The deletions that caused DRT/DMC on chromosome 6 disrupted a large intergenic ncRNA gene named *ASAR6* [122]. Interestingly, a smaller deletion that did not disrupt *ASAR6* did not cause DRT/DMC, indicating that *ASAR6* may be playing a role in the acquisition of this phenotype [123]. Additionally, rearrangements at the *ASAR6* locus resulted in chromosome structure instability and abnormal karyotypes [77, 122, 123].

ASAR6 is a monoallelically-expressed gene that displays asynchronous replication between alleles. One unique feature of the *ASAR6* locus is its ability to control gene silencing in *cis*. Deletions in chromosome 6 that disrupt *ASAR6* and cause DRT/DMC also reactivate the previously silenced alleles of monoallelically-expressed genes [122]. Therefore, when *ASAR6* is disrupted, monoallelically-expressed genes on chromosome 6 become biallelically expressed, indicating that *ASAR6* can function to mediate gene silencing in *cis*. In addition, ectopic integration of an *ASAR6*-containing transgene results in the

chromosome-wide silencing of gene expression on the integrated chromosome, further establishing a link between *ASAR6* expression and gene silencing (Chapter Two in this thesis).

Interestingly, an ~21 kb deletion in the XIC on the X chromosome results in DRT. This deletion disrupts a monoallelically-expressed transcript called Xist. The *Xist* gene shares many physical and functional similarities with *ASAR6*, including: 1) they both display random monoallelic gene expression, 2) they both display asynchronous replication that is coordinated with other linked monoallelic genes, 3) disruption of either gene results in delayed replication timing and instability of entire chromosomes in *cis*, 4) disruption of either gene results in the transcriptional activation of the previously silent alleles of linked mono-allelic genes, and 5) both are expressed from the late-replicating allele [79, 81, 122, 123, 176, 184].

Existing data indicate that *ASAR6* and *Xist* reside within loci that regulate DNA replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of their respective chromosomes. This is evident by the DRT, DMC, disrupted monoallelic gene expression and chromosome instability that are observed when these loci are deleted or disrupted. The DRT/DMC phenotype has been detected on chromosome rearrangements involving many different human and mouse chromosomes [74, 77, 80, 81, 122]. Therefore, it seems likely that all mammalian chromosomes contain loci that function to regulate chromosome-wide replication timing of individual chromosomes. Given the similarities in structure and function of the two loci characterized to date, *Xist*

and *ASAR6*, I propose that all mammalian chromosomes contain 'inactivation/stability centers' (I/S centers) that function to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of individual chromosomes.

If the above hypothesis is correct, then every chromosome should contain at least one of these genetic elements. In this chapter, I describe my efforts to generate intrachromosomal deletions at four loci known to be involved in translocation-induced DRT/DMC in an attempt to identify additional I/S centers. I found that deletions and inversions at a locus on chromosome 15q24 result in DRT/DMC, identifying it as a novel I/S center. Gene expression analysis at this locus reveals biallelic expression of a protein-coding gene called SCAPER and a non-annotated ncRNA. Yet, despite being biallelically-expressed, this region is asynchronously replicated and high in LINE-1 repeat concentration, which is similar to the other I/S centers that have been identified to date.

Results

Identification of Candidate I/S Centers on Other Chromosomes

In my attempt to identify new genomic regions that function similar to the *ASAR6* and *Xist* loci, I focused on three P-lines that display translocation-induced DRT/DMC upon *lox*P recombination (P175 in Figure 1.4, P186 in Figure 3.1A, and P268 in Figure 3.1B) [77]. Because a translocation at these *lox*P sites induces DRT/DMC, I reasoned that at least one of the *lox*P sites must reside within or near an I/S center. Therefore, generating deletions at these *lox*P loci in the P-lines should induce DRT/DMC. Previous work in P175 cells indicated that nearly all deletions proximal from the *lox*P site on chromosome 6 resulted in DRT/DMC and all deletions distal from the *lox*P site on chromosome 6 did not result in DRT/DMC [122, 123]. Furthermore, unanalyzed deletions have been made on chromosome 3 (Smith, L.E., Thayer, M, unpublished data). Since deletions have been made on 2 of the 6 chromosomes in the three P-lines, I have made deletions in the remaining four chromosomes (Figure 3.2).

Before deletions were made, I attempted to identify the exact molecular insertion site of the *lox*P containing plasmid on each of the four chromosomes. The *lox*P cassette integration site in chromosome 13 was previously identified at position 46,138,582 base pairs (NCBI Build GRCh37/hg19) (Thayer, M, unpublished data) (Figure 3.3A). Interestingly, by Southern blot and PCR, it was determined that multiple *lox*P cassettes integrated into chromosome 13 in an inverted manner. This indicated that multiple *lox*P sites exist on this chromosome

in opposite orientations, which meant that deletions could be made in both orientations on chromosome 13. Inverse PCR was used to determine the plasmid integration sites in the three other chromosomes [194]. For this analysis, I used Southern blot hybridizations to characterize the plasmid-genome junctions of either the 5´AP-*lox*P cassette or the *lox*P-3´RT cassette, size selected restriction fragments, circularized the fragments with ligase, and used inverse PCR with nested primers directed to the vector sequences from the *lox*P cassettes (data not shown). Direct sequencing of the PCR products indicated that the 5´AP-*lox*P cassette integrated in chromosome 16 at position 54,407,235 base pairs (NCBI Build GRCh37/hg19) (Figure 3.3B) in the P268 cell line and the *lox*P-3´RT cassette integrated in chromosome 15 at position 76,858,743 base pairs (NCBI Build GRCh37/hg19) (Figure 3.3C) in the P268 cell line. The orientations of the *lox*P sites are such that deletions can be made proximal from the *lox*P site on chromosome 16 and distal from the *lox*P site on chromosome 15. Unfortunately, the *lox*P integration site on chromosome 10 in P175 cells could not be identified despite multiple attempts. Nevertheless, deletions were still made in chromosome 10, but the exact size of the deletions could not be determined.

Deletions in Chromosomes 10 and 13 do not Cause DRT/DMC

Deletions were made in chromosomes 10, 13 and 16 using the strategy outlined in Figure 3.4. Briefly, P175, P186 or P268 cells were infected with a lentivirus containing a Blasticidin resistance gene (Blast), a *lox*P site and exons 3, 4 and 5 of mouse APRT. Cells resistant to Blasticidin were pooled, transfected

with Cre recombinase, and selected for the reconstitution of APRT. Cells that contained intrachromosomal deletions on the targeted chromosomes were selected based on their sensitivity to Blasticidin. Individual clones that contained a deletion on the targeted chromosome were isolated and expanded for further analysis. Linear amplification-mediated PCR (LAM-PCR) [195] was used to clone and sequence the lentiviral integration sites in a subset of these deletion clones. LAM-PCR was performed using primers directed at the 5′ LTR, and was carried out at the Fred Hutchinson Cancer Research Center's Clonal Analysis Core facility (http://www.fhcrc.org/science/shared_resources/cceh-clonal/index.html). Fifteen of the >50 lentiviral integrations in chromosomes 10, 13, and 16 were identified using this method. In addition, I have confirmed all 15 integration sites using PCR with primers directed at the lentiviral 5′LTR-genome junction sites (data not shown). Each deletion was also confirmed using multiple independent assays, including: Southern blot hybridizations, LOH analysis and DNA-FISH using BACs or Fosmids located within the deleted region (data not shown).

Multiple deletions in chromosome 10 were generated in P175 cells (∆175∆3´ cells, Table 3.1). DNA replication timing was assayed using a BrdU "terminal label" assay (Figure 2.3B). A comparison of the BrdU incorporation patterns between each chromosome 10 in ∆175∆3´-24a cells indicated that chromosome 10 deletions in P175 cells did not cause DRT (Figure 3.5A). Further analysis of three other clones that have deletions in chromosome 10 indicated that none displayed DRT/DMC (data not shown). This indicates that the *lox*P integration site in chromosome 10 plays no role in DRT/DMC and correlates with

data showing that new chromosomal translocations involving chromosome 10 in P175 cells did not cause DRT/DMC [122].

Multiple deletions in chromosome 13 were generated in P186 cells (Δ186Δ3´ cells, Table 3.2). A comparison of the BrdU incorporation patterns between each chromosome 13 in Δ186Δ3´-6b cells indicated that chromosome 13 deletions in P186 cells did not cause DRT (Figure 3.5B). Analysis of three other clones that have a deletion in chromosome 13 indicated that none displayed DRT/DMC (data not shown). This indicates that the *lox*P locus on chromosome 13 is not involved in the DRT/DMC phenotype.

Multiple deletions in chromosome 16 were generated in P268 cells (Δ268Δ3´ cells, Table 3.3). Despite confirming 3 lentiviral integrations by LAM-PCR, LOH could not be detected in any of the clones. Due to these conflicting data, none of the clones were analyzed for DRT/DMC. Further screening will be needed to confirm that a deletion is present.

Deletions and Inversions in Chromosome 15q24 Cause DRT/DMC

Deletions distal to the centromere from the *lox*P integration site were made in chromosome 15 using the strategy outlined in Figure 3.6. The strategy is similar to that used to make deletions in chromosomes 10, 13, and 16 and identical to the strategy used to make deletions in chromosome 6 [122]. Individual clones that contained a deletion were isolated and expanded for further analysis. For each clone, the exact lentiviral integration site was identified by LAM-PCR. Each deletion was characterized using multiple independent assays,

including: Southern blot hybridizations, junction PCR designed to span the *lox*P-3´RT-genome junction to determine if loss-of-junction (LOJ) had occurred, LOH analysis and DNA-FISH using a BAC or Fosmid located within the deleted region (Figure 3.7 and data not shown). Using this strategy, I generated 11 different deletions distal from the *lox*P site in chromosome 15 ranging in size from ~2 kb to ~23 Mb (Table 3.4 and Figure 3.8).

In addition to the distal deletions in chromosome 15, I also made deletions in the opposite orientation (proximal from the *lox*P site). To do this, I used a second site-specific recombinase system, flippase/flippase recognition target (FLP/Frt) [196], to engineer deletions anchored at the *lox*P-3´RT cassette and extending towards the centromere on chromosome 15 (Figure 3.9). Again, all deletions were confirmed by Southern blot hybridizations, junction PCR designed to span the *lox*P-3´RT-genome junction to determine if LOJ has occurred, LOH analysis and DNA-FISH using BACs or Fosmids located within the deleted region (data not shown). Using this strategy, I have generated three different deletions proximal from the *lox*P integration site in chromosome 15, and these deletions ranged in size from <126 Kb to >18 Mb (Table 3.4).

The deletion strategies not only allowed isolation of different-sized deletions, but also separate clonal isolates that have the same deletion. Therefore, it was possible to analyze the replication timing of clones that had different-sized deletions and clones that had the same deletion but grew separately from one another. Similar Southern banding patterns and the

presence of identical lentiviral integration sites allowed for the identification of clones that had identical deletions (Table 3.5).

In addition to making deletions on chromosome 15, I also generated intrachromosomal inversions. To isolate cells that contained an inversion, I used the same strategy used to make deletions distal from the *lox*P site (Figure 3.6), only instead of isolating APRT+, Blast sensitive colonies, I isolated APRT+, Blast resistant colonies. Therefore, instead of isolating clones that had lost the Blast gene, clones were isolated that had undergone *lox*P recombination without a deletion. Blast resistant colonies could result from one of two processes, either an interchromosomal translocation or an intrachromosomal inversion. Lentiviral integration sites were identified by LAM-PCR and confirmed by PCRs identifying the new junctions formed by inversion. DNA-FISH was also used to confirm that a translocation involving chromosome 15 had not taken place (data not shown). This strategy allowed for the isolation of inversions in both orientations. Using this strategy, I isolated three different intrachromosomal inversions at this locus on chromosome 15, one extending proximal from the original *lox*P site and two extending distal (Table 3.6).

Before the replication timing of the deletions and inversions on chromosome 15 was analyzed, the replication timing of the chromosome 15s in the parental P268 cells was analyzed as a control. P268 is triploid for chromosome 15, which added another layer of complexity to the replication timing analysis. Nevertheless, the DNA replication timing of the three intact chromosome 15s in P268 showed synchronous replication timing and the

chromosomes displayed a banded pattern of BrdU incorporation that is consistent with the known replication timing for chromosome 15 (Figure 3.10). Thus, since the replication timing in the parent line is normal, any replication timing aberrations that are observed must be caused by the chromosome manipulations. Next, DNA replication timing was assayed using a BrdU "terminal label" assay (Figure 2.3B). Analysis of cells with an ~161 kb deletion indicated that some chromosome 15 deletions caused DRT (Figure 3.11A-D). Furthermore, DRT/DMC was observed in multiple other clones as well (Figure 3.12). In all, eight rearrangement clones (6 deletions and 2 inversions) were identified that displayed DRT/DMC (Table 3.4 and Table 3.6), suggesting that this locus on chromosome 15q24 is a novel I/S center.

Different than what was observed with the chromosome 6 deletions, where all deletions in one orientation resulted in DRT and all deletions in the other orientation did not result in DRT [122], the deletion analysis on chromosome 15 was not as precise. Examples of deletions in both orientations that did not display DRT/DMC were observed (Figure 3.13 and Table 3.4). Surprisingly, even some clones that had identical deletions gave opposite phenotypes. For instance, clone ∆268-4g had been shown to display DRT (Figure 3.11), while clone ∆268-4e had been shown to not display DRT (Figure 3.13), even though these two clones had been identified as having the same deletion (Table 3.5). Therefore, despite having the same genetic deletion, one clone displayed the phenotype and one did not. There are many possibilities that could explain this lack of penetrance: 1) It is possible that *trans*-acting factors are

differentially regulated in these clones and this is having an effect on the phenotype, 2) I/S centers could be epigenetic features and it is possible that a new I/S center can form following the deletion of an existing one, and 3) The extreme genetic instability phenotype associated with chromosome 15 deletions (shown below) may be causing intrachromosomal rearrangements that influence the severity of the DRT/DMC phenotype.

Disruption of the Chromosome 15 Locus in P268 Cells Results in Genomic Instability

Previous data from the Thayer lab showed that DRT/DMC results in CSIN and CIN [75, 77, 123]. During the replication analysis of deletions and inversions in chromosome 15, it became apparent that these clones had extremely abnormal karyotypes. Deletions and inversions in chromosome 15 resulted in multiple rearrangements involving chromosome 15 and aneuploidy (Figure 3.14 and Figure 3.11E). This instability often caused complications in the DNA replication analysis, as some clones had no intact chromosome 15s and could not be analyzed (Table 3.4 and Table 3.6). Importantly, of the clones that could be analyzed, instability was generally only observed in those that displayed DRT/DMC. For example, 50% of the cells in clone Δ268-4f (which displays DRT/DMC) had at least one chromosomal translocation involving chromosome 15, whereas only 3% of the cells in clone Δ268-4o (which has the same sized deletion as Δ268-4f but did not display DRT/DMC) had a chromosome 15 translocation (data not shown). Therefore, this suggests that genomic instability

is a result of DRT/DMC and not the particular genetic rearrangement. There is also a distinct possibility that the clones that were too rearranged to be analyzed did display DRT/DMC at one point, which can explain the extreme karyotypic abnormalities. This analysis indicates that DRT/DMC on chromosome 15 in P268 cells can result in genomic instability.

Deletions and Translocations at 15q24 are Present in Tumor Cell Lines and Correlate with Chromosome 15 Mutagenesis

Deletion and/or translocation at the *lox*P-3´RT locus caused delayed replication, delayed condensation and genomic instability in P268 cells. The *lox*P-3´RT cassette in P268 cells integrated in the middle of an ~500 Kb gene called S phase cyclin A-associated protein in the ER (*SCAPER*). We next wanted to see whether disruption of the 15q24 (*SCAPER*) locus occurs in cancer development. The Wellcome Trust Sanger Institute has performed LOH analysis on ~800 tumor cell lines and made this information freely available for analysis [197]. A quick overview of these data indicated that all or part of *SCAPER* experienced LOH in 171 of 768 tumor cell lines (roughly 22%). Ten cell lines were also found that had a homozygous deletion in *SCAPER*. As expected, many of these tumor lines displayed LOH of large regions or the entirety of chromosome 15, but some had LOH concentrated at a small region within *SCAPER*. We acquired five of the cell lines used in the Sanger analysis that had a more localized disruption of the *SCAPER* gene and did a DNA replication timing analysis in conjunction with DNA-FISH. HCC1143, which has a breakpoint

junction right in the middle of *SCAPER* and LOH of most of the gene, was found to display replication delay on a chromosome 15 derivative (Figure 3.15). Replication delay was also observed in HCC1395, which has a small area of LOH encompassing *SCAPER* (Figure 3.16). As was evident from the Sanger SNP analysis, both of these cell lines had numerous rearrangements involving chromosome 15 and also displayed aneuploidy (Figure 3.15 and Figure 3.16). Although they had multiple chromosome 15 rearrangements, no overt replication delay on chromosome 15 was observed in HCC1569, K-562, or MCF7 cells. However, given the transient nature of this phenotype, it is possible that these cells did display DRT/DMC at some point in time during their history (see Chapter One). From these data, I conclude that *SCAPER* disruption is present in some tumor cell lines and correlates with DRT/DMC and numerous rearrangements affecting chromosome 15.

The Chromosome 15q24 Locus Contains a High Concentration of LINE-1 Repeats

Now that three loci (on chromosomes 6, 15, and X) had been implicated in the acquisition of DRT/DMC and genomic instability, a comparison could be made to identify unique molecular signatures that were common to all three [81, 122]. One striking feature of the chromosome 6 (*ASAR6*) and X chromosome (*Xist*) loci is that they harbor an above average concentration of LINE-1 repeats. Where the genome average of LINE-1 repeats is ~17%, the *ASAR6* and *Xist* loci contain ~45% and ~50%, respectively ([198] and unpublished analysis).

Expanding this analysis to 300 Kb surrounding the *lox*P cassette integration site in *SCAPER* indicates that this locus contains ~54% LINE-1 repeats (Figure 3.3C). In contrast, the locus on chromosome 13 that was previously identified to not be involved in DRT/DMC contains ~15% LINE-1 sequence, similar to the genome average (Figure 3.3A). Therefore, all three I/S centers identified to date have a high LINE-1 density.

Asynchronous Replication on Chromosome 15

Another hallmark of the *ASAR6* and *Xist* loci is that they exhibit asynchronous DNA replication. To examine the DNA replication patterns of the *SCAPER* locus, we used a FISH-based assay. When hybridizing a probe to a particular site in S phase cells, some cells will display two probe signals, indicating that neither allele has replicated yet (SS pattern) and some cells will display four probe signals, indicating that both alleles have replicated (DD pattern). Some cells also display three probe signals, indicating that only one allele has replicated (SD pattern). If a locus is asynchronously replicating then it displays the SD pattern in ~30-50% of the cells, whereas synchronously replicating loci display the SD pattern in only ~10-20% of cells [170, 171, 199]. Using the "single-dot-double-dot" assay, the replication pattern of *SCAPER* was analyzed in human skin fibroblasts (HSFs). It was found that two different probes within *SCAPER* show the SD pattern in 35% and 47% of cells (Figure 3.17). Therefore, the *SCAPER* locus is asynchronously replicated.

Previous studies have shown that genes that are asynchronously replicated coordinate their DNA replication such that all early-replicating alleles reside on one homolog and all late-replicating alleles reside on the other homolog (*cis* coordination) [170, 171]. To determine if the asynchronous replication of the *SCAPER* locus was coordinated with other asynchronously replicating genes, three other asynchronously replicating genes on chromosome 15 were identified based on their predicted monoallelic-expression status and replication analysis by the Thayer lab [162, 163] (Figure 3.17C). We then tested if these asynchronously replicating loci also displayed coordination in their asynchronous replication. The level of coordination was examined using a two-color DNA-FISH assay and scoring cells that simultaneously displayed the SD signal for both loci [170, 171]. For this analysis, a BAC probe representing *SCAPER* was used in combination with probes representing *MYO1E, PTPN9* and *PEAK1*. It was found that the asynchronous replication of *SCAPER* was coordinated in cis with all three probes, indicating that asynchronously replicating genes on chromosome 15 are coordinated in *cis* (Figure 3.17C).

One limitation of the “single dot-double dot” assay is that the asynchronous replication of loci greater than 50 Mb apart are difficult to score, as a signal coming from the paternal allele of one locus may be closer to the maternal allele of the other locus. Therefore, to assay the random asynchronous replication of chromosome 15 loci at the whole chromosome level, a second replication-timing assay known as Replication Timing-Specific Hybridization, or ReTiSH was used [166]. In the ReTiSH assay, cells are labeled with BrdU for

different times and then harvested during metaphase (Figure 3.18A). Regions of chromosomes that incorporate BrdU are visualized by a modification of Chromosome Orientation-Fluorescence In Situ Hybridization (CO-FISH), where the replicated regions (BrdU-labeled) are converted to single stranded DNA and then hybridized directly with specific probes [200]. Since metaphase chromosomes are analyzed for hybridization signals located on the same chromosome in metaphase spreads, the physical distance between the two loci is not a limitation of the ReTiSH assay [166]. One distant region on chromosome 15 that is asynchronously replicated is the nucleolar-organizing region (NOR) on the P-arm. This region contains tandem copies of ribosomal DNA (rDNA) that are monoallelically expressed and asynchronously replicated [166]. To assay coordination of asynchronous replication in a second human cell type, primary blood lymphocytes (PBLs) were used for the ReTiSH analysis. PBLs were exposed to BrdU for either 5 or 14 hours, mitotic cells were harvested, processed for BrdU incorporation and subjected to ReTISH using an 18s rDNA probe, a chromosome-15 centromeric probe and a probe within *SCAPER*. It was found that the late-replicating alleles of *SCAPER* and rDNA were on the same chromosome, and therefore asynchronous replication of *SCAPER* and rDNA is coordinated in *cis* (Figure 3.18B-E and Figure 3.17C). Thus, *SCAPER* replication is coordinated with other asynchronously-replicating loci in *cis*, even loci that cross the centromere.

Complex Expression Pattern at the *SCAPER* Locus

Asynchronous DNA replication is correlated with monoallelic gene expression [167]. One hallmark of the *ASAR6* and *Xist* loci is that they exhibit monoallelic gene expression of nearby genes [122, 125]. The asynchronous replication of *SCAPER* and other nearby genes suggested that these genes might also be monoallelically expressed. Therefore, I assayed the expression pattern of *SCAPER* on chromosome 15 using reverse transcription-PCR (RT-PCR). I analyzed expression from HTD114 (the parent cells of P268), P268 and two mouse mono-chromosomal hybrids (mouse L cells, each containing one or the other chromosome 15 from HTD114 cells). *SCAPER* expression was detected in all cell lines; however, it was found to be biallelic using multiple different assays (Figure 3.19A-B). For example, RT-PCR using primers that span heterozygous SNPs and other primers that span an intron (assayed in mono-chromosomal hybrids) indicated that the spliced SCAPER transcript was biallelically expressed (Figure 3.19B). Therefore, it appears that the biallelic expression in this region comes from expression of the spliced, protein-coding gene *SCAPER*. However, it is also possible that there is transcription on the opposite strand in this region, so strand-specific RT-PCR was used to generate cDNA from either strand in HTD114 cells. Analysis of the strand-specific cDNA revealed transcription on both strands from within the *SCAPER* gene (Figure 3.19C) and heterozygous SNP analysis indicated that both strands are biallelically expressed (data not shown). It is known that monoallelic expression can very widely between different cell types and even different clones of the

same cell type [163, 201]; therefore, *SCAPER* expression was assayed in clonal lymphoblastoid cell lines. Analysis of four clones revealed that *SCAPER* was biallelically expressed in each clone (data not shown).

Due to the complex expression pattern observed within SCAPER (see above), I next determined if this region is transcribed by RNA Pol II by treating cells with α-amanitin, which is a selective inhibitor of this polymerase [202], and assaying expression of SCAPER RNA using a semi-quantitative RT-PCR assay. The results of this analysis are shown in Figure 3.20, and indicated that transcription in this region was indeed sensitive to α-amanitin. Similarly, RNA expressed from the protein-coding gene P300 was also sensitive to α-amanitin treatment. In contrast, expression of 45S RNA (an RNA Polymerase I product) and a tRNA gene (an RNA Polymerase III product) were not inhibited by α-amanitin.

In an effort to characterize gene expression on this chromosome for future experiments, a gene expression analysis of an ~2 Mb region surrounding *SCAPER* was performed. As summarized in Table 3.7, while nearly every gene in this region was transcribed, there was no indication of monoallelic expression. This is surprising considering this region is asynchronously replicating and 3 genes, PTPN9, ISL2 and PEAK1, have been shown to be or predicted to be monoallelically expressed [162, 163, 203]. I next asked whether we could detect monoallelic gene expression at distant sites on chromosome 15 by analyzing the expression of genes that were reported to be monoallelically expressed [163], asynchronously replicating [204], or within the imprinted Prader-Willi (PW) locus.

Most of these genes were either not expressed or biallelically expressed; however, as expected, RNAs within the PW locus and the *FAM174b* gene were found to be monoallelically expressed (Table 3.8). Expression analysis from the mono-chromosomal hybrids indicated that the imprinted gene in the PW locus was expressed from the same chromosome that contained the *lox*P-3´RT cassette in P268 and that *FAM174b* was expressed from the opposite allele (data not shown).

From these analyses, I conclude that the *SCAPER* locus in P268 cells is transcribed by RNA Pol II and is characterized by the biallelic expression of both sense and antisense transcripts. In addition, this region is biallelically transcribed in multiple different cell types, despite being asynchronously replicated in PBLs and HSFs. In fact, no gene within a 2 Mb vicinity of this region is monoallelically expressed in P268 cells. Therefore, while the *SCAPER* locus displays both sense and antisense transcription by RNA Pol II and is asynchronously replicated, it does not display monoallelic gene expression.

Discussion

The DRT/DMC phenotype has been detected on chromosome rearrangements involving many different human and mouse chromosomes [77, 81, 205-207]. Therefore, it seems likely that all mammalian chromosomes contain loci that function to regulate chromosome-wide replication timing of individual chromosomes. Given the similarities in structure and function of the two loci characterized to date, *Xist* and *ASAR6*, I have proposed that all mammalian chromosomes contain 'I/S centers' that function to maintain proper replication timing, mitotic chromosome condensation, mono-allelic gene expression and stability of individual chromosomes. Under this scenario every mammalian chromosome contains four distinct types of *cis*-acting elements, origins of replication, centromeres, telomeres, and I/S centers, all functioning to ensure proper replication, segregation and stability of individual chromosomes.

To uncover I/S centers on other chromosomes, I have made deletions at loci that have been shown to be involved in translocation-induced DRT/DMC. During this analysis, I identified a region on chromosome 15 that, upon disruption, results in DRT/DMC. It should be noted that for the six chromosomes that the Thayer lab has made deletions in (Figure 3.2), deletions have only been made in both orientations for chromosomes 6, 13 and 15. Therefore, it is possible that deletions in the other orientation on chromosomes 3, 10 and 16 will display DRT/DMC as well, so these loci should not be ruled out as potential regulators of this phenotype. Furthermore, deletions on chromosomes 3 and 16 have still not

been analyzed for DRT/DMC. Thus, the purpose of this study was to identify novel I/S centers and not to rule out specific chromosomal loci.

While it is evident that many deletions in *SCAPER* on chromosome 15 resulted in DRT/DMC, it is noteworthy that some did not. This was particularly surprising given the fact that the chromosome 6 deletion data in *ASAR6* did not show this variability [122]. There are several explanations for this lack of penetrance: One is that the ploidy imbalance of chromosome 15 is affecting the phenotype in some way. One trend that we noticed is that deletion clones that remain mostly triploid for chromosome 15 tend to display DRT/DMC more often than deletion clones that are mostly tetraploid for chromosome 15. This is an intriguing possibility because deletion of both *Xist* alleles was shown to have a much more dramatic impact on DNA replication timing than deletion of just one allele, indicating there are *trans*-effects that affect the severity of the DRT/DMC phenotype on the X chromosome [79, 81]. Under this scenario, it would be hypothesized that the additional copy of chromosome 15 could rescue the phenotype through a *trans*-acting mechanism.

Another possibility is that the generation of DRT/DMC might be via a *cis*-acting epigenetic mechanism. The best evidence for this possibility is that I generated clones that have the exact same deletion in chromosome 15, yet some displayed DRT/DMC and some did not. Assuming these clones were genetically identical, the only possibility would be that an epigenetic mechanism was somehow differentially regulated in these clones. If this region harbors an epigenetic mark that regulates chromosome-wide replication timing, it is possible

that a new I/S center could arise *de novo* elsewhere on the chromosome in some cells following excision of the initial one. A similar phenomenon has been observed for another *cis*-acting chromosomal element, the centromere. Neocentromere formation has been documented in multiple organisms following excision of a centromere or the generation of an acentric translocation derivative [208]. Since the centromere is essentially an epigenetic structure [209], it appears capable of forming on any DNA sequence [210]. Another cis-acting chromosomal element, the DNA replication origin, has essentially no sequence specificity in higher eukaryotes [7]. Although the replication origin is technically a genetic element, the pre-RC is an epigenetic feature that forms on all origins and is essential for their function. New replication origins can also form upon the deletion of existing ones [7]. So it is formally possible that I/S centers represent epigenetic features of every chromosome and that new ones can arise following the loss of an existing one.

A third possibility is that the instability associated with the chromosome 15 deletions is somehow impacting the DRT/DMC phenotype. Previously, the Thayer lab has found that once the DRT/DMC chromosome had experienced secondary translocations, it often ceased to display DRT/DMC [74, 77]. Therefore, it is possible that some of the clones that did not display DRT/DMC had undergone chromosomal rearrangements that, in effect, rescued the phenotype. In our DNA replication timing analysis, only chromosome 15s that had not undergone an interchromosomal translocation were analyzed. However, from our DNA-FISH analysis it was difficult to determine if any intrachromosomal

events have occurred. Therefore, it is possible that some clones have had their delayed chromosomes rescued by intrachromosomal events. Since the chromosome 15-deletions are much more unstable than the chromosome 6-deletions, this possibility could explain why the deletion analysis on chromosome 6 was not as complicated. This could also explain the differences in phenotype between clones that have the same deletion. Since these clones grew independently of one another, they may have accumulated different rearrangements that could potentially impact replication timing differentially.

Aside from identifying a novel I/S center, another contribution of this work is the demonstration that intrachromosomal inversions can cause DRT/DMC. It is now known that a translocation, deletion, and inversion at specific loci are all sufficient to cause DRT/DMC. Inversions are interesting because, unlike translocations and deletions, the derivative chromosome retains all the same genetic information in *cis*. The inversion chromosomes may prove valuable for future studies aimed at identifying the molecular mechanism responsible for DRT/DMC.

The Thayer lab has now characterized several common features of the *Xist*, *ASAR6* and *SCAPER* loci: 1) disruption of these loci results in delayed replication timing of entire chromosomes in *cis*, 2) disruption of all three loci results in dramatic instability of the affected chromosome, 3) all three loci have a high LINE-1 content, and 4) each locus displays asynchronous replication that is coordinated with other linked asynchronously replicated loci. Surprisingly, *SCAPER* was biallelically expressed in all cell types examined. However, given

the variability observed with many monoallelic genes [163], it is possible that *SCAPER* is monoallelically expressed in some tissues and biallelically expressed in HTD114 cells and lymphoblastoid cells.

Random monoallelic expression is regulated differentially in different cell types [163, 165]. On the other hand, asynchronous DNA replication is a much more consistent chromosomal feature. Therefore, it is possible that asynchronous DNA replication is the more critical feature of these loci and monoallelic gene expression is more of a by-product of differential replication timing. However, we know that the monoallelically-expressed *Xist* gene, and possibly the *ASAR6* gene, is essential for allelic inactivation of the silent alleles of linked monoallelically-expressed genes. As mentioned in Chapter Two, it is possible that these loci have dual functions: one being transcript-independent and controlling DNA replication timing and chromosome stability and the other one being transcript-dependent and controlling allelic inactivation. It is possible that the *SCAPER* locus only has the transcript-independent function and does not control allelic inactivation in *cis*. Unfortunately, we have not assayed whether chromosome 15 deletions cause reactivation of previously silenced alleles of monoallelically-expressed genes, as is the case for deletions in chromosomes 6 and X. This experiment may shed light on the function of this region. Regardless, this study has identified a new I/S center on chromosome 15 and supports my hypothesis that all chromosomes contain at least one of these centers.

Materials and Methods

Cell Culture

Low passage primary human skin fibroblasts were obtained from ATCC and cultured in DMEM plus 10% fetal bovine serum (Hyclone). Primary blood lymphocytes were isolated after venipuncture into a Vacutainer CPT (Becton Dickinson, Franklin Lakes, NJ) per the manufacturer's recommendations and grown in 5 mL RPMI 1640 (Life Technologies) supplemented with 10% fetal bovine serum (Hyclone) and 1% phytohemagglutinin (Life Technologies). P268 cells are a human APRT deficient cell line derived from the HT-1080 fibrosarcoma [211], and were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (Hyclone). P268 derivatives were grown as above with the addition of 500 mg/ml Geneticin (Gibco), 200 mg/ml Hygromycin B (Calbiochem), and/or 10 ug/ml Blasticidin S HCl (Invitrogen). The deletion-line derivatives were grown in DMEM supplemented with 10% dialyzed fetal bovine serum (Hyclone), 10 mg/ml azaserine (Sigma) and 10 mg/ml adenine (Sigma) to facilitate selection of APRT-expressing cells. All cells were grown in a humidified incubator at 37°C in a 5% carbon dioxide atmosphere.

DNA-FISH

Trypsinized cells were centrifuged at 1,000 rpm for 10 minutes in a swinging bucket rotor. The cell pellet was re-suspended in 75 mM potassium chloride for 15-30 minutes at 37°C, re-centrifuged at 1,000 rpm for 10 minutes

and fixed in 3:1 methanol:acetic acid. Fixed cells were added drop-wise to microscope slides to generate mitotic chromosome spreads using standard methods [192]. Slides with mitotic spreads were baked at 85°C for 20 minutes and then treated with 0.1 mg/ml RNAase for 1 hour at 37°C. After RNAase treatment, the slides were washed in 2xSSC (1xSSC is 150 mM NaCl and 15 mM sodium citrate) with 3 changes for 3 minutes each and dehydrated in 70%, 90%, and 100% ethanol for 3 minutes each. The slides were denatured in 70% formamide in 2xSSC at 70°C for 3 min and whole chromosome paints were used according to the manufacturer's recommendations and hybridization solutions (American Laboratory Technologies and Vysis). Detection of digoxigenin-dUTP probes utilized a three-step incubation of slides with sheep FITC-conjugated anti-digoxigenin antibodies (Roche) followed by rabbit FITC-conjugated anti-sheep antibodies (Roche) followed by goat FITC-conjugated anti-rabbit antibodies (Jackson Laboratories). Slides were stained with DAPI (12.5 mg/ml) or propidium iodide (0.3 mg/ml), cover slipped, and viewed under UV fluorescence with appropriate filters (Olympus).

Centromeric, BAC, and Fosmid probes: Mitotic chromosome spreads were prepared as described above. Slides were treated with RNase at 100ug/ml for 1h at 37^{O}C and washed in 2xSSC and dehydrated in 70%, 90% and 100% ethanol. Chromosomal DNA was denatured at 75^{O}C for 3 minutes in 70% formamaide/2XSSC, followed by dehydration in ice cold 70%, 90% and 100% ethanol. BAC and Fosmid DNAs were nick-translated using standard protocols to incorporate biotin-11-dUTP or digoxigenin-dUTP (Invitrogen). BAC and Fosmid

DNAs were directly labeled with Cy3-dUTP, FITC-dUTP, Spectrum Orange-dUTP or Spectrum Green_dUTP (Vysis, Abbott Laboratories) using nick-translation or random priming using standard protocols. Final probe concentrations varied from 40-60 ng/*u*l. Centromeric probe cocktails (Vysis) plus BAC or Fosmid DNAs were denatured at 75^{O}C for 10 minutes and prehybridized at 37^{O}C for 30 minutes. Probes were applied to slides and incubated overnight at 37^{O}C. Post-hybridization washes consisted of three 3-minute rinses in 50% formamide/2XSSC, three 3-minute rinses in 2XSSC, and finally three 3-minute rinses in PN buffer (0.1M Na2HPO4 + 0.0M NaH2PO4, ph 8.0, +2.5% Nonidet NP-40), all at 45^{O}C. Signal detection was carried out as described [212]. Amplification of biotinylated probe signal utilized alternating incubations of slides with anti-avidin (Vector) and FITC-Extravidin (Sigma). Slides were then counterstained with either propidium iodide (2.5ug/ml) or DAPI (15ug/ml) and viewed under UV fluorescence (Olympus).

Replication Timing Assay

The BrdU replication timing assay was performed on exponentially dividing cultures as follows: asynchronously growing cells were exposed to 20 ug/ml of BrdU (Sigma) for 2-9 hours [213]. Mitotic cells were harvested in the absence of colcemid, treated with 75 mM KCl for 15-30 minutes at 37°C, fixed in 3:1 methanol:acetic acid and dropped on wet ice cold slides. The chromosomes were denatured in 70% formamide in 2xSSC at 70°C for 3 minutes and processed for DNA FISH, as described above. The incorporated BrdU was then

detected using a FITC-labeled anti-BrdU antibody (Becton Dickinson). Slides were stained with propidium iodide (0.3 mg/ml), cover slipped, and viewed under UV fluorescence.

All images were captured with an Olympus BX Fluorescent Microscope using a 100X objective, automatic filter-wheel and Cytovision workstation. Individual chromosomes were identified with either chromosome-specific paints or centromeric probes in combination with BACs from the deleted regions. Utilizing the Cytovision workstation, each chromosome was isolated from the metaphase spread and a line drawn along the middle of the entire length of the chromosome. The Cytovision software was used to calculate the pixel area and intensity along each chromosome for each fluorochrome occupied by the DAPI and BrdU (FITC) signals. The total amount of fluorescent signal was calculated by multiplying the average pixel intensity by the area occupied by those pixels.

ReTiSH

We used the ReTiSH assay essentially as described [166]. Briefly, unsynchronized, exponentially growing cells were treated with 30μM BrdU (Sigma) for 5 and 14 hours. Colcemid (Sigma) was added to a final concentration of 0.1 μg/mL for 1 h at 37°C. Cells were trypsinized, centrifuged at 1,000 rpm, and resuspended in prewarmed hypotonic KCl solution (0.075 M) for 40 min at 37°C. Cells were pelleted by centrifugation and fixed with methanol-glacial acetic acid (3:1). Fixed cells were drop gently onto wet, cold slides and allowed to air-dry. Slides were treated with 100μg/ml RNAse A at 37°C for 10 min. Slides were

rinsed briefly in d_2H_20 followed by fixation in 4% formaldehyde at room temperature for 10 minutes. Slides were incubated with pepsin (1 mg/mL in 2N HCl) for 10 min at 37°C, and then rinsed again with d_2H_20 and stained with 0.5 µg/µL Hoechst 33258 (Sigma) for 15 minutes. Slides were flooded with 200µl 2xSSC, coverslipped and exposed to 365-nm UV light for 30 min using a UV Stratalinker 2400 transilluminator (Stratagene). Slides were rinsed with d_2H_20 and drained. Slides were incubated with 100µl of 3U/µl of ExoIII (Fermentas) in ExoIII buffer for 15 min at 37°C. The slides were then processed directly for DNA FISH as described above, except with the absence of a denaturation step.

Semi-quantitative RT-PCR

Total RNA was extracted using Trizol (Invitrogen) reagent. Total RNA was subjected to reverse transcriptase reactions using Superscript III (Invitrogen) according to the manufacturers instructions. PCRs were carried out with a first cycle of 2 minutes at 95°C, 45 seconds at 60°C and 1 minute at 72°C followed by 35-42 cycles of 30 seconds at 95°C, 45 seconds at 60°C and 1 minute at 72 °C. The conditions were chosen so that none of the PCRs reached a plateau at the end of the amplification protocol, i.e. they were in the exponential phase of amplification. Each set of reactions always included a genomic DNA positive control, and a no sample and a no reverse transcriptase negative controls. The PCR products were resolved on 1% agarose gels and stained with ethidium bromide. The gels were photographed under UV illumination, and the resulting image was inverted using Photoshop (Adobe).

Strand-specific RT-PCR

Total RNA was extracted from HTD114 cells using Trizol (Invitrogen) reagent. Total RNA was subjected to reverse transcriptase reactions using Superscript III (Invitrogen) according to the manufacturers instructions. A primer was included in the mix that amplified SNP rs12916573 on the antisense strand:

5´- TAATACGACTCACTATAGGGGCATACTCTTTGGCCTTTTGAG-3´

SNP rs12916573 on the sense strand:

5´- TAATACGACTCACTATAGGGTGTCAGCCCTAAAGTTGATCTG-3´

SNP rs2468125 on the antisense strand:

5´- TAATACGACTCACTATAGGGAACACGAGGGAATAAAAACTAGC-3´

and SNP rs2468125 on the sense strand:

5´- TAATACGACTCACTATAGGGGGACCTCAAATATCCTTTGTTATATGG-3´

Each primer was included in tubes with or without reverse transcriptase. A tube without any primer was also included as a control. Random-primed cDNA was made using the same protocol but with random hexamers instead of gene-specific primers. Each cDNA primer has a T7 (5´-TAATACGACTCACTATAGGG-3´) tag on the 5´ end to ensure specific amplification. Each cDNA reaction was amplified by 35 rounds of PCR at a 62ºC annealing temp. with:

T7 and 5´-TGTCAGCCCTAAAGTTGATCTG-3´ for antisense rs12916573

T7 and 5´- GCATACTCTTTGGCCTTTTGAG-3´ for sense rs12916573

T7 and 5´- TAAACAACTTTCCCAAAACAAAAAG-3´ for antisense rs2468125

T7 and 5´- AACACGAGGGAATAAAAACTAGC-3´ for sense rs2468125

The reaction was nested with the same conditions for 30 more cycles with:

T7 and 5´- TGTACCATCATCTCCCCAAGTAG-3´ for antisense rs12916573

T7 and 5´- AAGAAGCCATTGCCTAAAATCC-3´ for sense rs12916573

T7 and 5´- AAAGACAGCTGAATTCTCACAGC-3´ for antisense rs2468125

T7 and 5´- AGATTCCTCCCCTAAAACCAAG-3´ for sense rs2468125

PCR and Southern Blotting to Screen for Deletions

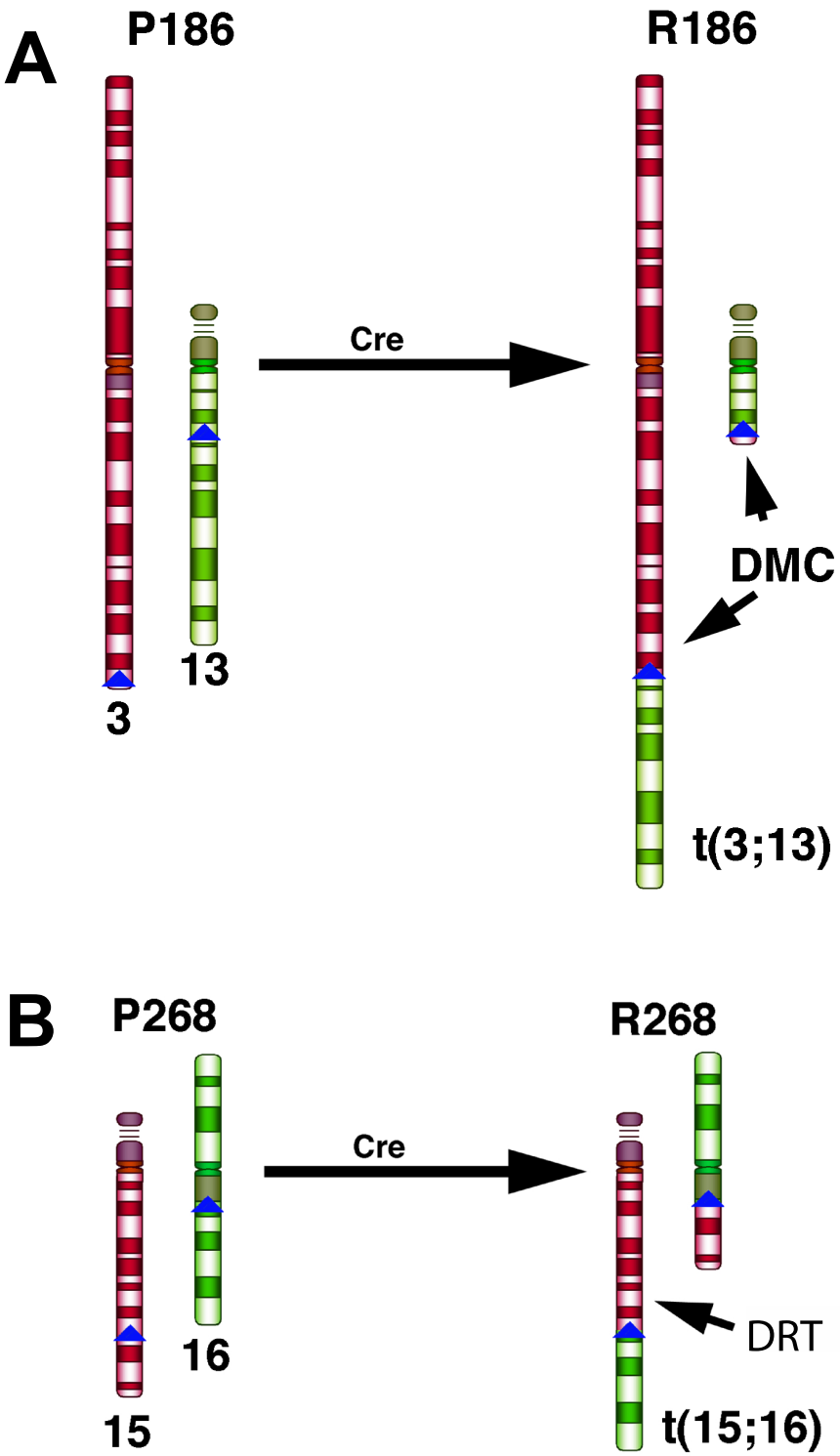
DNA was isolated from cells and subjected to PCR in a 25-50 µl volume using 50-100 nanograms (ng) of DNA, 1x Standard Taq Buffer (New England Biolabs, Inc.), 200 μM each deoxynucleotide triphosphate, 0.2 μM of each primer, and 3 U of Taq DNA Polymerase (New England Biolabs, Inc.) under the following reaction conditions: 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 45 sec, and 72°C for 1 min, with a final extension time of 10 min at 72°C. PCR products were separated on 1% agarose gels, stained with ethidium bromide, and photographed under ultraviolet light illumination.

BclI endonuclease digestion was performed according to the manufacturer's protocols (Fermentas) and digested DNA was run on a 0.7% agarose gel overnight at low voltage and subjected to Southern blotting using standard protocols. Blots were exposed to film for 10-30 hours at -80ºC.

Figures and Tables

Figure 3.1: Schematic representation of chromosomes in the P186/R186 and P268/R268 cell lines. A) The P186 cell line contains a *lox*P site in one allele of chromosome 3 and one allele of chromosome 13. Following transient Cre expression in P186, R186 was generated which contains a balanced 3;13 translocation. In R186, both translocation derivatives display DRT/DMC while all the other chromosomes in the cell do not display DRT/DMC. **B**) The P268 cell line contains a *lox*P site in one allele of chromosome 15 and one allele of chromosome 16. No other chromosomes in this cell line have been manipulated. Following transient Cre expression in P268, R268 was generated which contains a balanced 15;16 translocation. In R268, the chromosome-15-centromere derivative displays DRT/DMC while the chromosome-16-centromere derivative and all the other chromosomes in the cell do not display DRT/DMC. Figure 3.1 is adapted from [122]. (Figure on next page).

Figure 3.1: Schematic representation of chromosomes in the P186/R186 and P268/R268 cell lines.



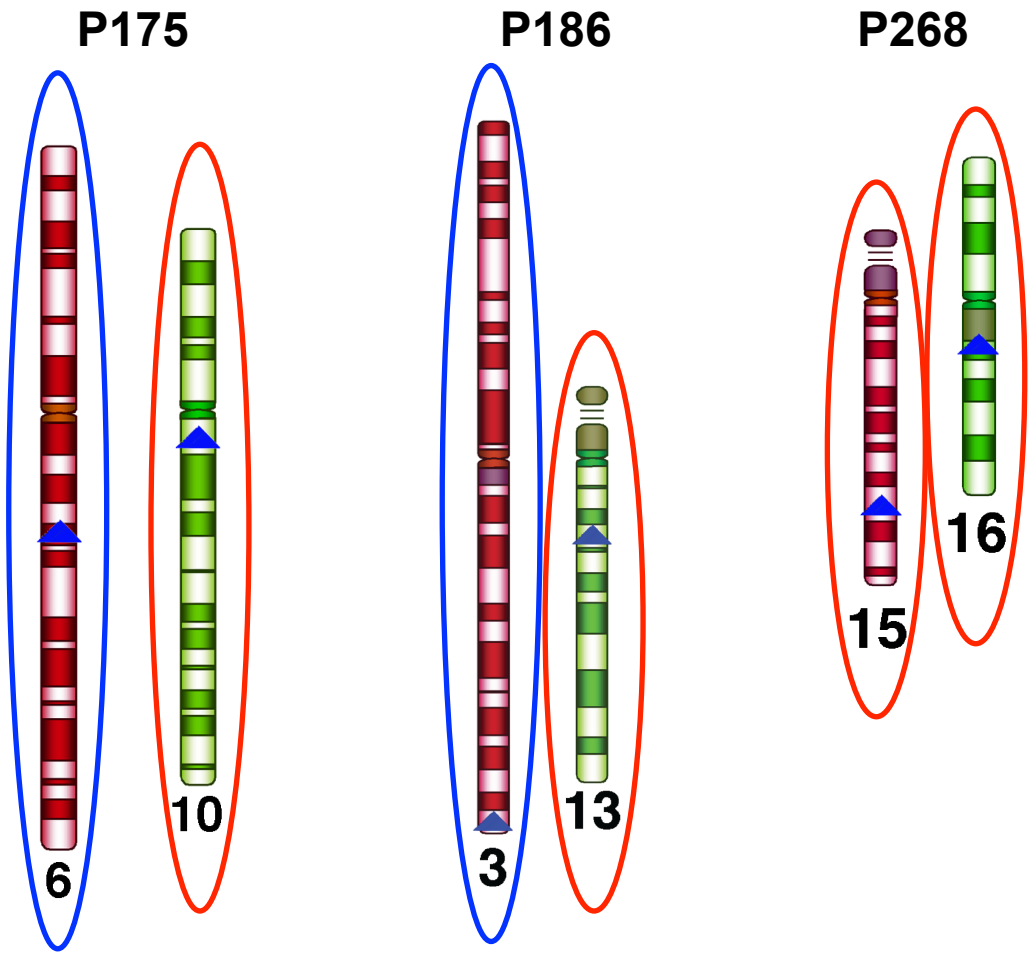


Figure 3.2: Illustration of chromosomes that will undergo deletion analysis. The chromosome number is indicated below each chromosome in the P175, P186 and P268 cell lines. Deletions have previously been made in chromosomes 6 and 3 (circled in blue). In this chapter, intrachromosomal deletions will be generated in chromosomes 10, 13, 15 and 16 (circled in red).

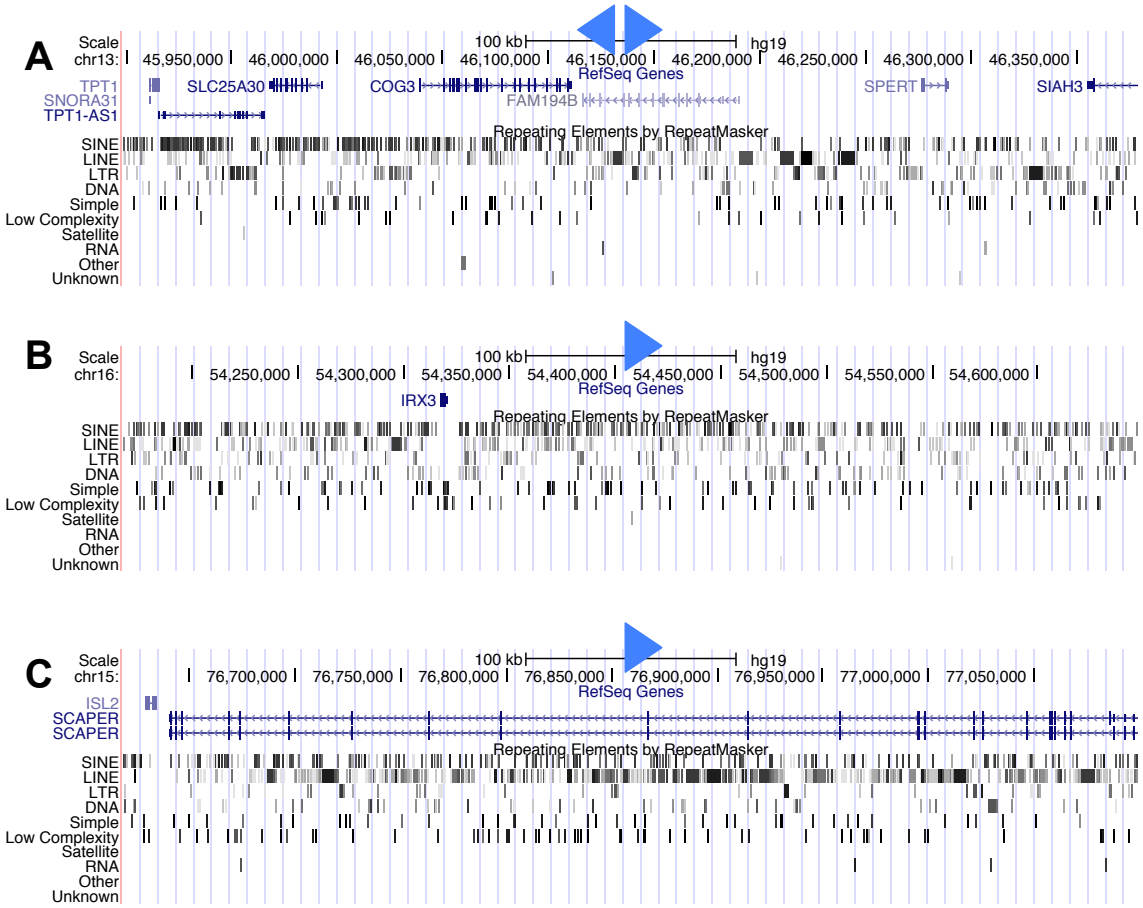


Figure 3.3: Location of the original *lox*P cassette integration sites in chromosomes 13, 16 and 15. Shown is a screenshot of the UCSC Genome Browser for the *lox*P-integrated regions of chromosome 13 in P186 (**A**), chromosome 16 in P268 (**B**), and chromosome 15 in P268 (**C**). The exact *lox*P integration site is indicated by a blue triangle (note that chromosome 13 has multiple *lox*P sites in an inverted orientation) and the locations of annotated Refseq genes [214] and repetitive sequence elements [215] are shown in each screenshot. Base positions are from NCBI Build GRCh37/hg19.

Figure 3.4: Strategy used to make deletions in chromosome 10. In the top left corner is a depiction of chromosomes 6 (red) and 10 (green) in the P175 cell line. For visualization purposes, chromosome 10 has been enlarged and flipped on its side. **Step 1**) The original 5´AP-*lox*P integration site is indicated in a magnified view of chromosome 10 in P175. P175 cells were infected with a lentivirus containing a Blasticidin resistance gene (Blast), a *lox*P site (blue triangle) and exons 3, 4 and 5 of mouse APRT. **Step 2**) Infected clones were pooled and transfected with Cre recombinase. **Step 3**) Following *lox*P recombination, APRT was reconstituted. If an intrachromosomal deletion occurred, Blast will be removed from the genome along with the deleted region, resulting in sensitivity to Blasticidin. Therefore, clones that were APRT positive and Blast sensitive were isolated and expanded for further screening. This strategy was also used to make deletions in chromosome 13 in P186 cells and chromosome 16 in P268 cells. (Figure on next page).

Figure 3.4: Strategy used to make deletions in chromosome 10.

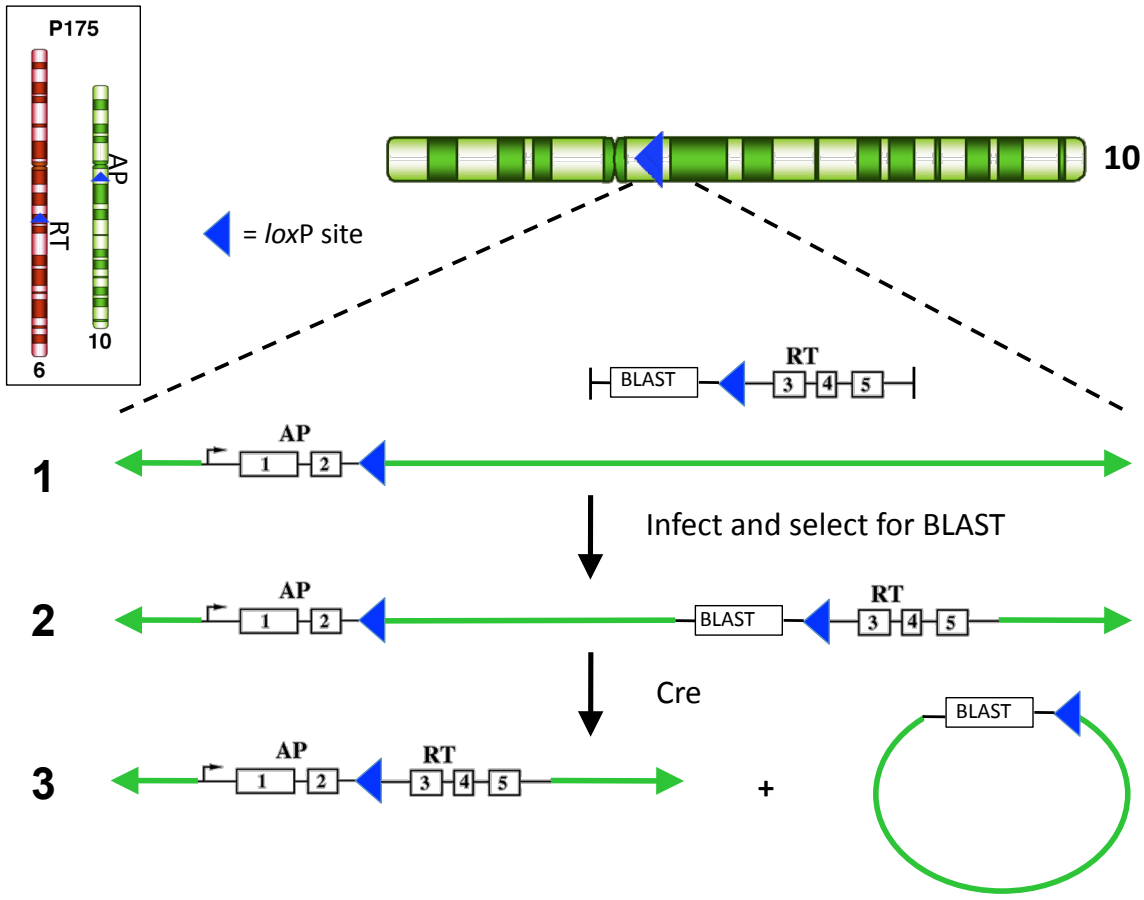


Table 3.1: Chromosome 10 deletion clones. The 21 clones with a deletion distal from the *lox*P site on chromosome 10 (Δ175Δ3´ clones) are listed on the left of the table. Multiple clones exist for those with an asterisk (*) next to their name. The exact viral integration sites were identified for some of the clones based off LAM-PCR data. Because the original 5´AP-*lox*P cassette integration site is unknown in these clones, an exact deletion size could not be determined. Relative deletion size was estimated based on LOH analysis. The heterozygous SNP closest to the centromere on the q arm of chromosome 10 (rs4275553) was assayed for LOH, and those clones that did not lose heterozygosity were put in the "smallest" category. To the right is a summary of the screening methods that have been performed (Y) or not performed (N) for each clone as well as which have been analyzed for DRT/DMC. Bp positions are from the NCBI Build GRCh37/hg19. (Table on next page).

Table 3.1: Chromosome 10 deletion clones.

Chromosome 10 Lentiviral Integration Sites in P175		Deletion Size (relative)		Screened by		Analyzed
Clone	Integration Site (bp)	Proximal	Distal	Southern	BAC-FISH	for DRT/DMC
5'AP-*lox*P	unknown	N/A	N/A	N/A	N/A	N/A
Δ175Δ3'-4c	42,543,719	N/A	smallest (no LOH)	Y	N	Y
Δ175Δ3'-6a	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-7a	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-17c	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-17d	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-21a	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-22a	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-24b	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-24c	unknown	N/A	smallest (no LOH)	Y	N	N
Δ175Δ3'-2b	43,123,687	N/A	small (LOH at rs4275553)	Y	N	N
Δ175Δ3'-17a	unknown	N/A	small (LOH at rs4275553)	Y	N	N
Δ175Δ3'-19a	unknown	N/A	small (LOH at rs4275553)	Y	N	N
Δ175Δ3'-24d*	unknown	N/A	small (LOH at rs4275553)	Y	N	N
Δ175Δ3'-15a	unknown	N/A	mid-small (LOH at rs4948975)	Y	N	N
Δ175Δ3'-17b	unknown	N/A	mid-small (LOH at rs4948975)	Y	N	N
Δ175Δ3'-24a	unknown	N/A	mid (LOH at rs10857619)	Y	Y	Y
Δ175Δ3'-6c	55,485,838	N/A	mid-large (LOH at rs11003227)	Y	Y	Y
Δ175Δ3'-20b	unknown	N/A	mid-large (LOH at rs11003227)	Y	N	N
Δ175Δ3'-4b	56,292,927	N/A	large (LOH at rs7896552)	Y	Y	Y
Δ175Δ3'-20a	unknown	N/A	large (LOH at rs7896552)	Y	N	N
Δ175Δ3'-15b	unknown	N/A	largest (LOH at rs1336194)	Y	N	N

Figure 3.5: Deletions in chromosomes 10 and 13 do not cause DRT/DMC. **A**) ∆175∆3´-24a cells were incubated with BrdU for 5 hours, harvested for mitotic cells, processed for DNA-FISH using a chromosome-10 telomeric probe (red signal near the telomere) and a BAC probe (RP11-35B22) that maps to the deleted region (red signal near the centromere), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The two chromosome 10s in the top left panel are indicated by arrows. The chromosomes from the top left panel were cut out and aligned with each color displayed separately or in combination below. The chromosome 10 with only one red signal (A) is the chromosome that contains the deletion. On the right, pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the two chromosome 10s (A and B). Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome. The similar amount of BrdU incorporation in the two chromosomes indicated that these cells did not display DRT **B**) ∆186∆3´-6b cells were incubated with BrdU for 4.5 hours, harvested for mitotic cells, processed for DNA-FISH using a BAC probe (RP11-274P12) near the chromosome 13 centromere (red signal near the centromere) and a BAC probe (CTD-3195C8) that maps to the deleted region (red signal slightly lower on the q arm), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The two chromosome 13s in the top left panel are indicated by arrows. The chromosomes from the top left panel were cut out and aligned with each color displayed separately or in combination below. The chromosome 13 with only one red signal (A) is the

chromosome that contains the deletion. On the right, pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the two chromosome 13s (A and B). Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome. The similar amount of BrdU incorporation in the two chromosomes indicated that these cells did not display DRT (Figure on next page).

Figure 3.5: Deletions in chromosomes 10 and 13 do not cause DRT/DMC.

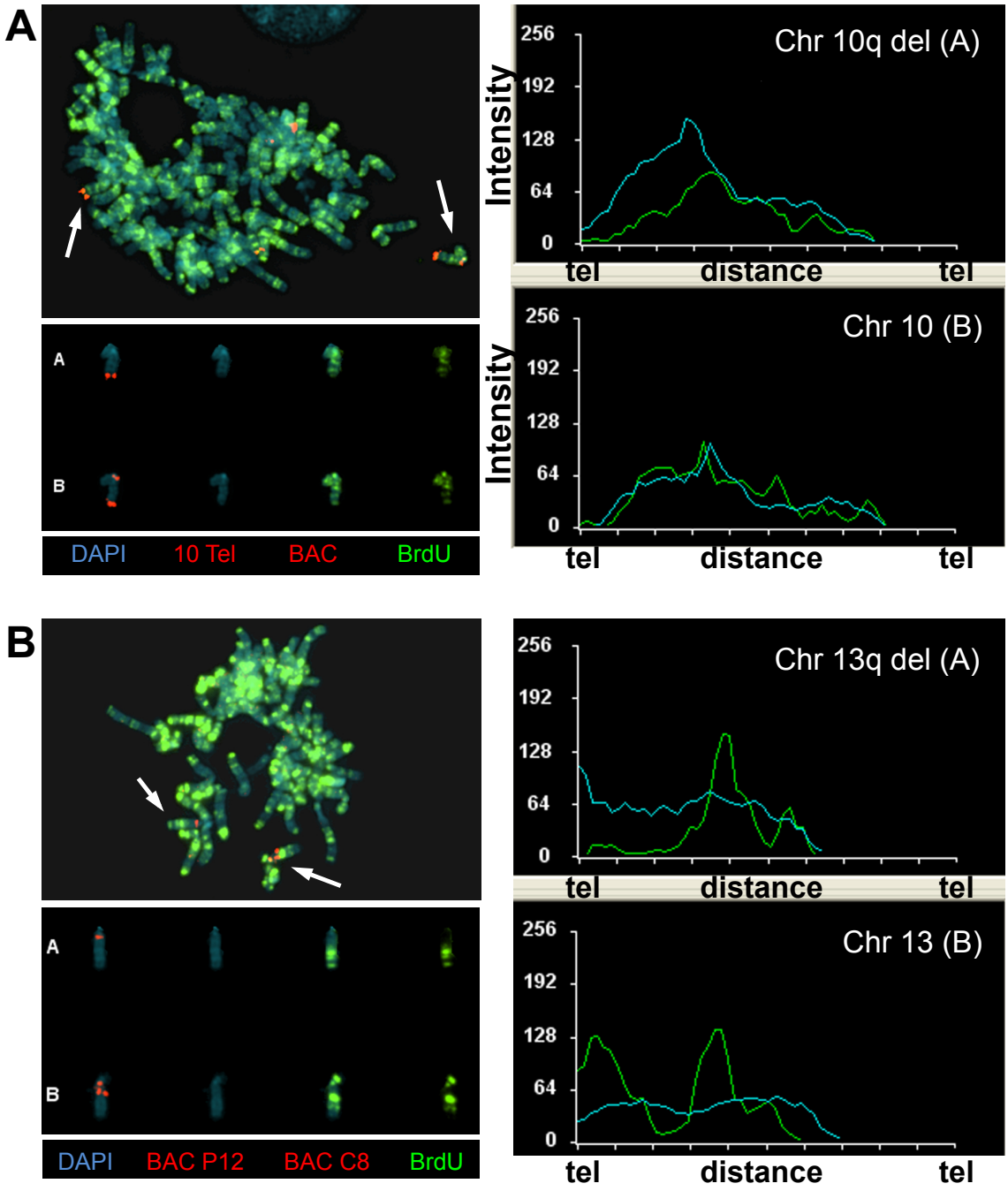


Table 3.2: Chromosome 13 deletion clones. The 21 clones with a deletion distal or proximal from the *lox*P site on chromosome 13 (Δ186Δ3´ clones) are listed on the left of the table. Multiple clones exist for those with an asterisk (*) next to their name. The exact viral integration sites were identified for some of the clones based off LAM-PCR data. The deletion size is the difference between the original 5´AP-*lox*P integration site and the viral integration site. Note how some clones harbor a distal deletion, some contain a proximal deletion and one contains a deletion in both orientations on the same chromosome. To the right is a summary of the screening methods that have been performed (Y) or not performed (N) for each clone as well as which have been analyzed for DRT/DMC. Bp positions are from the NCBI Build GRCh37/hg19. (Table on next page).

Table 3.2: Chromosome 13 deletion clones.

Chromosome 13 Lentiviral Integration Sites in P186		Deletion Size (in bp)		Screened by		Analyzed
Clone	Integration Site (bp)	Proximal	Distal	Southern	BAC-FISH	for DRT/DMC
Δ186Δ3'-5a	45,788,690	349,892	N/A	Y	Y	Y
Δ186Δ3'-1c*	45,860,823	2,778	N/A	Y	Y	Y
5'AP-*lox*P	46,138,582	N/A	N/A	N/A	N/A	N/A
Δ186Δ3'-4a*	46,155,178	N/A	16,596	Y	N	N
Δ186Δ3'-5c	46,583,938	N/A	445,356	Y	Y	Y
Δ186Δ3'-4b	46,850,932	N/A	712,350	Y	N	N
Δ186Δ3'-6b	47,002,920	N/A	864,338	Y	Y	Y
Δ186Δ3'-4d	55,976,429	N/A	9,837,847	Y	N	N
Δ186Δ3'-3a	45,812,118 + one addl.	326,464	> 12,361	Y	N	N
Δ186Δ3'-3d	unknown	?	?	Y	N	N
Δ186Δ3'-3b	unknown	?	?	Y	N	N
Δ186Δ3'-2d	unknown	?	?	Y	N	N
Δ186Δ3'-3c	unknown	?	?	Y	N	N
Δ186Δ3'-2b	unknown	?	?	Y	N	N
Δ186Δ3'-5b	unknown	?	?	Y	N	N
Δ186Δ3'-2a	unknown	?	?	Y	N	N
Δ186Δ3'-6a	unknown	?	?	Y	N	N
Δ186Δ3'-1e	unknown	?	?	Y	N	N
Δ186Δ3'-6c	unknown	?	?	Y	N	N
Δ186Δ3'-6d	unknown	?	?	Y	N	N
Δ186Δ3'-6e	unknown	?	?	Y	N	N
Δ186Δ3'-1a	unknown	?	?	Y	N	N

Chromosome 16 Lentiviral Integration Sites in P268		Deletion Size (in bp)		Screened by		Analyzed
Clone	Integration Site (bp)	Proximal	Distal	Southern	BAC-FISH	for DRT/DMC
Δ268Δ3'-2e	47,231,746	7,175,489	N/A	Y	N	N
Δ268Δ3'-2b	unknown	?	N/A	Y	N	N
Δ268Δ3'-2d	unknown	?	N/A	Y	N	N
Δ268Δ3'-2c	47,338,650	7,068,585	N/A	Y	N	N
Δ268Δ3'-5a	unknown	?	N/A	Y	N	N
Δ268Δ3'-5b	unknown	?	N/A	Y	N	N
Δ268Δ3'-5c	unknown	?	N/A	Y	N	N
Δ268Δ3'-6a	53,878,215	529,020	N/A	Y	N	N
Δ268Δ3'-6c	unknown	?	N/A	Y	N	N
5'AP-*lox*P	54,407,235	N/A	N/A	N/A	N/A	N/A

Table 3.3: Chromosome 16 deletion clones. The 9 clones with a deletion proximal from the *lox*P site on chromosome 16 (Δ268Δ3´ clones) are listed on the left of the table. The exact viral integration sites were identified for some of the clones based off LAM-PCR data. The deletion size is the difference between the original 5´AP-*lox*P integration site and the viral integration site. To the right is a summary of the screening methods that have been performed (Y) or not performed (N) for each clone as well as which have been analyzed for DRT/DMC. Bp positions are from the NCBI Build GRCh37/hg19.

Figure 3.6: Strategy used to make deletions in chromosome 15. In the top left corner is a depiction of chromosomes 15 (red) and 16 (green) in the P268 cell line. For visualization purposes, chromosome 15 has been enlarged and flipped on its side. **Step 1**) The original *lox*P-3´RT integration site is indicated in a magnified view of chromosome 15 in P268 (this cassette integrated in an antisense orientation). P268 cells were infected with a lentivirus containing a Blasticidin resistance gene (Blast), a *lox*P site (blue triangle) and exons 1 and 2 of mouse APRT. **Step 2**) Infected clones were pooled and transfected with Cre recombinase. **Step 3**) Following *lox*P recombination, APRT was reconstituted. If an intrachromosomal deletion occurred, Blast will be removed from the genome along with the deleted region, resulting in sensitivity to Blasticidin. Therefore, clones that were APRT positive and Blast sensitive were isolated and expanded for further screening. (Figure on next page).

Figure 3.6: Strategy used to make deletions in chromosome 15.

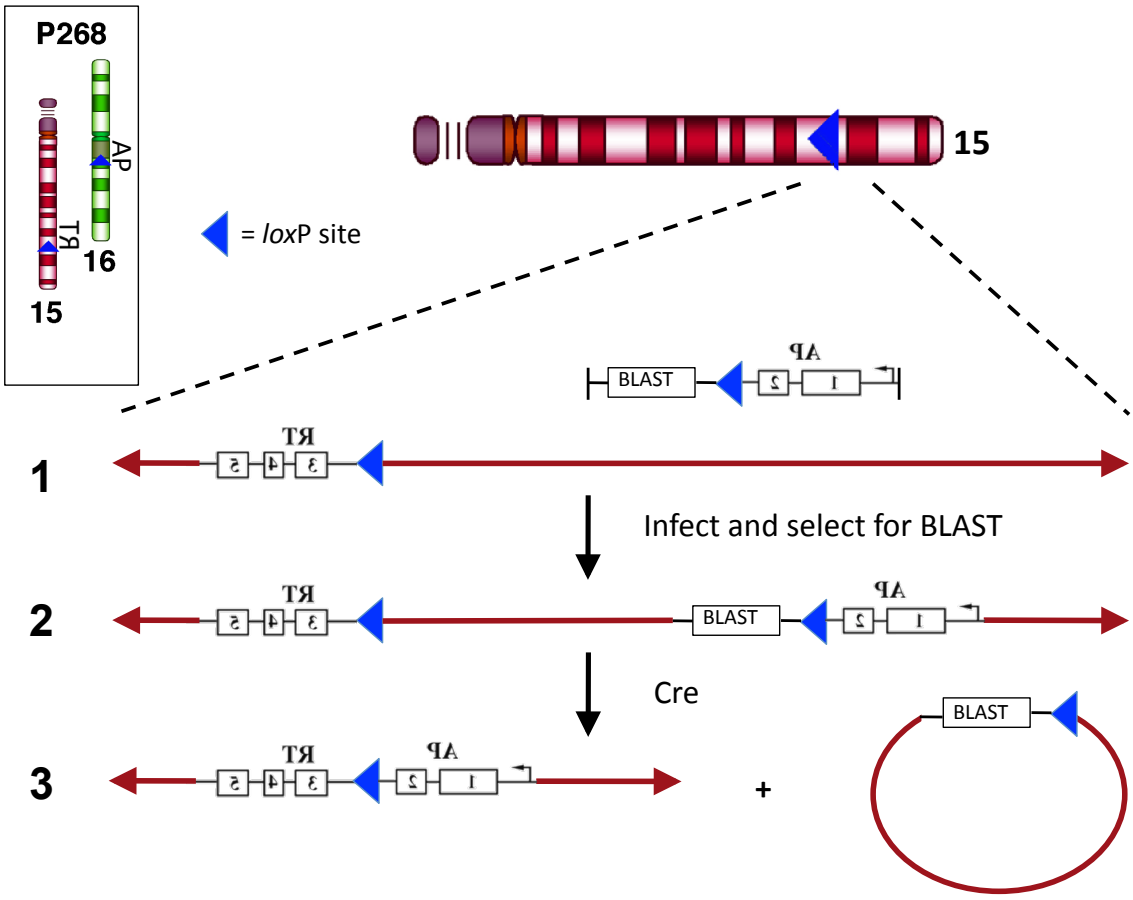
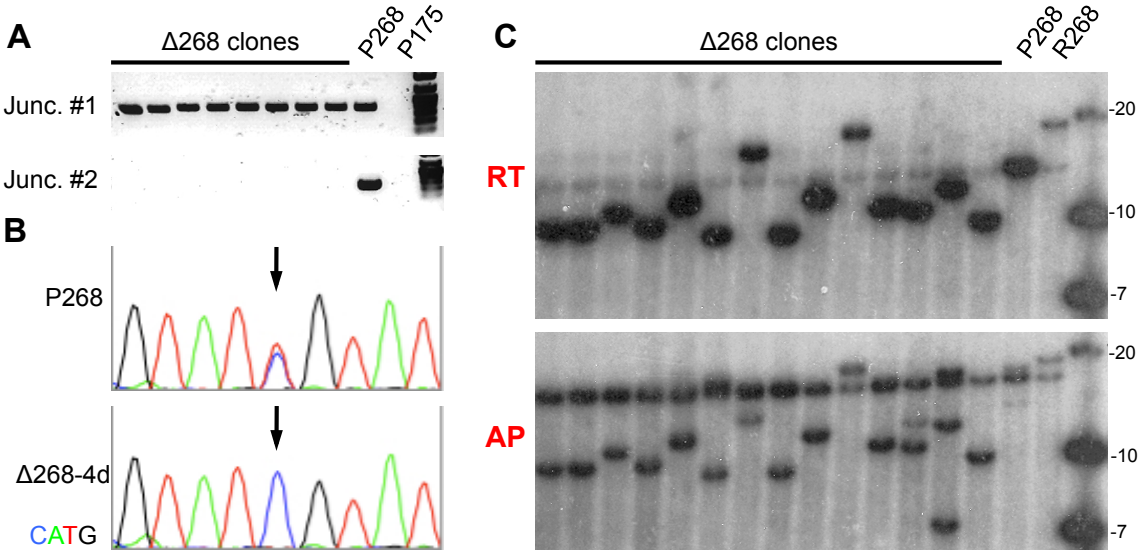


Figure 3.7: Deletion confirmation for ∆268 clones. A) Junction PCR to detect the loss of one *lox*P-3´RT-genome junction. ∆268 clones were subjected to PCR with primers that amplified the centromeric junction of the original *loxP*-3´RT integration site in P268 (Junc. #1) or the telomeric junction (Junc. #2). All of the deletion clones have lost the telomeric junction indicating that a deletion distal from the *lox*P site was made. **B**) LOH in deletion clones. Sequence trace for SNP rs2881582 (indicated by arrow) indicated that P268 contains both a T and C at this position while a deletion clone has lost the T allele. **C**) Southern blot of deletion clones, P268 and R268. Genomic DNA was digested with BclI and run on a gel with a DNA ladder (right lane, labeled in Kb). The top blot was probed with exons 3, 4, and 5 of mouse APRT (RT) and the bottom blot was probed with exons 1 and 2 of APRT (AP). Note how the size of the RT band has shifted in size compared to P268, indicating a rearrangement took place at this locus. Also, each clone has at least two AP bands (the original on chromosome 10 and a new one from the lentiviral infection. (Figure on next page).

Figure 3.7: Deletion confirmation for Δ268 clones.



Chromosome 15 Lentiviral Integration Sites in P268		Deletion Size (base pairs)		DRT/DMC		
Clone	Integration Site	Proximal	Distal	positive	negative	unstable
Δ268F-6a	btw cen-58,733,512	>18,125,231		0	1	0
Δ268F-5a	btw 76,732,559-76,791,779	btw 66,964-126,184		0	3	0
Δ268F-6b	btw 76,732,559-76,791,780	btw 66,964-126,184		0	3	0
loxP-3'RT	76,858,743-76,858,758			N/A	N/A	N/A
Δ268-15a	76,860,843		2,085	0	2	0
Δ268-18a	76,982,805		124,047	0	4	0
Δ268-4c	76,994,171		135,413	3	3	0
Δ268-4a	77,020,070		161,312	2	4	1
Δ268-18t	77,113,970		255,212	0	2	0
Δ268-6e	btw 77,451,350-tel		>592,592	0	1	0
Δ268-1c	btw 77,451,350-tel		>592,592	0	0	1
Δ268-3g	btw 77,451,350-tel		>592,592	0	0	1
Δ268-5c	82,480,936		5,622,178	0	0	2
Δ268-4d	89,692,218		12,833,460	1	0	0
Δ268-18b	100,154,623		23,295,865	0	2	0

Table 3.4: Deletion size and DRT/DMC status of chromosome 15 deletion clones. Positions are listed for the original *lox*P-3´RT integration site in P268 and lentiviral integrations in each deletion clone. Δ268F clones have a deletion proximal from the *lox*P site and Δ268 clones have a deletion distal from the *lox*P site. Specific integration sites were determined by LAM-PCR and predicted integration sites were estimated by LOH analysis. The deletion size is the difference between the original *lox*P-3´RT integration site and the lentiviral integration site. On the right is a summary of how many clones exhibited DRT/DMC (positive), did not display DRT/DMC (negative), or were too unstable to analyze (unstable). Each deletion clone listed on the left represents one of many clones with the same deletion; therefore multiple clones were analyzed for DRT/DMC for each deletion size. Bp positions are from the NCBI Build GRCh37/hg19.

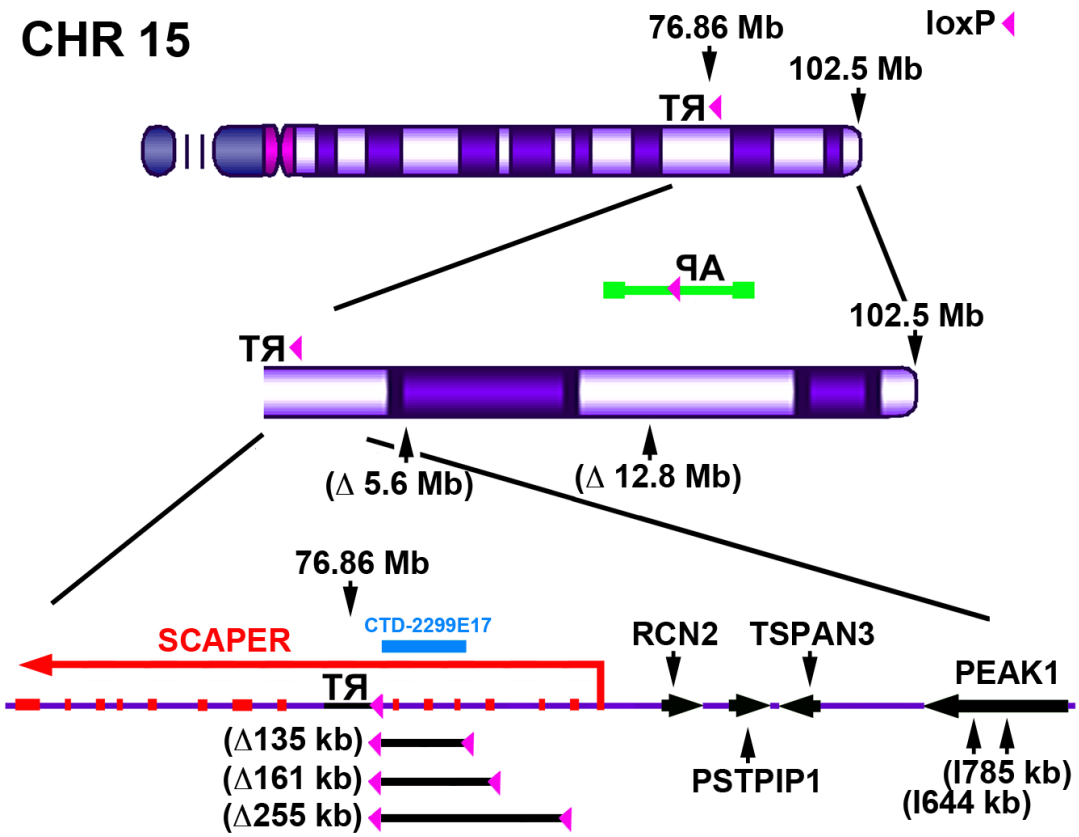
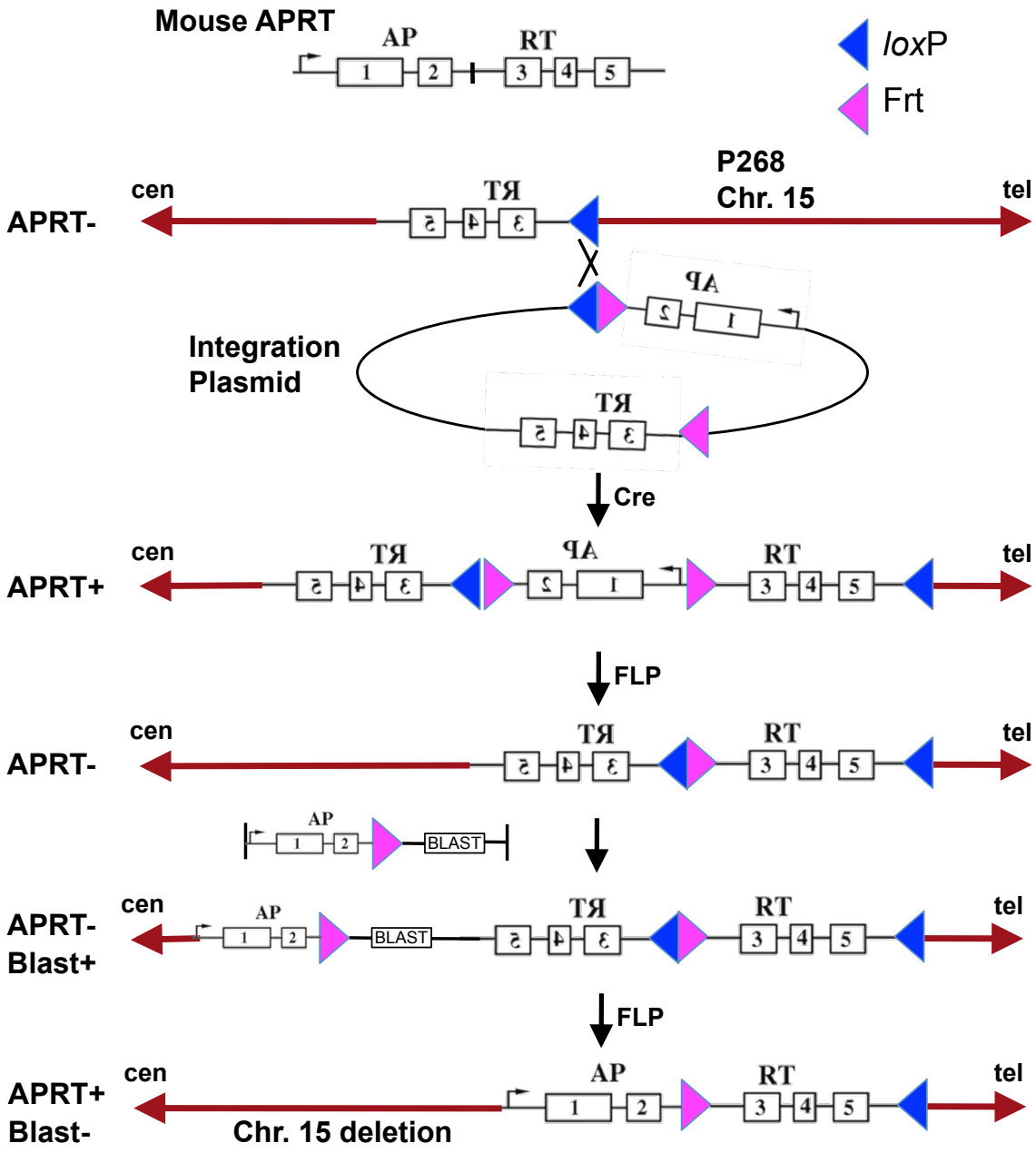


Figure 3.8: Chromosome 15 distal deletions. Diagram of chromosome 15 showing the orientation and integration site of the original *lox*P-3´RT cassette (RT+purple triangle), the 5´AP-*lox*P lentivirus (green) and the locations of the genomic integrations of the virus (arrows). The *lox*P-3´RT integration site is magnified to show the locations and orientations of the genes: SCAPER (red), RCN2, PSTPIP1, TSPAN3 and PEAK1. The BAC used for deletion analysis is shown in blue.

Figure 3.9: Strategy to make deletions proximal from the *lox*P integration site in chromosome 15 in P268. Schematic diagram of the mouse APRT gene, the location and orientation of the original *lox*P-3´RT site, the 'integration' plasmid (containing a *lox*P site, an AP-Frt cassette and an Frt-RT cassette), the lentiviral vector (L6-AP-Frt-Blast), and FLP-mediated deletion and reconstitution of APRT. P268 cells were transfected with the integration plasmid and Cre then selected for the presence of the APRT gene. Cells were transfected with FLP, selected for the absence of APRT, and infected with the 5´AP lentivirus. Blast-resistant cells were transfected with FLP and selected for the presence of APRT and Blast sensitivity. Individual clones were isolated, expanded and screened for deletion. This strategy resulted in a deletion of genomic material between the *lox*P-3´RT cassette and the centromere. Figure 3.9 adapted from [122]. (Figure on next page).

Figure 3.9: Strategy to make deletions proximal from the *lox*P integration site in chromosome 15 in P268.



Chromosome 15 Identical Clones	
Clone	Same As
Δ268-4c	4f, 4k, 4n, 4o, 4q, 4r, 4u
Δ268-4a	4e, 4g, 4m, 4p, 4s, 4t
Δ268-3g	3e, 3f, 3h
Δ268-5c	5d
Δ268-15a	15c, 15d, 15f
Δ268-18b	18e, 18f, 18g, 18h
Δ268-18t	18x
Δ268-18a	18c, 18d, 18m, 18n, 18p, 18r, 18s, 18w
Δ268F-5a	5b, 5c, 5d, 5e, 5f, 5g, 5h, 5i, 5k, 5m, 5n, 5o 5p, 5q, 5r, 5s, 5t, 5u, 5z
Δ268F-6b	6c, 6d, 6e, 6f

Table 3.5: List of chromosome 15 deletion clones with identical deletions.

The deletion clones in Table 3.4 that contain the same deletion as at least one other clone are listed on the left. The names of the identical clones are listed on the right.

Chromosome 15 Lentiviral Integration Sites in P268		Inversion Size (base pairs)		DRT/DMC		
Clone	Integration Site	Proximal	Distal	positive	negative	unstable
Inv268-3c	74,775,308	2,083,435		1	0	0
loxP-3'RT	76,858,743-76,858,758			N/A	N/A	N/A
Inv268-3a	77,502,462		643,704	0	0	1
Inv268-6c	77,643,709		784,951	1	0	0

Table 3.6: Inversion size and DRT/DMC status of chromosome 15 inversion clones. Positions are listed for the original *lox*P-3´RT integration site in P268 and lentiviral integrations in each inversion clone. Inv268-3c contains an inversion proximal from the *lox*P site and Inv268-3a and -6c contain an inversion distal from the *lox*P site. Integration sites were determined by LAM-PCR. The inversion size is the difference between the original *lox*P-3´RT integration site and the lentiviral integration site. On the right is a summary of how many clones exhibited DRT/DMC (positive), did not exhibit DRT/DMC (negative), or were too unstable to analyze (unstable). Bp positions are from the NCBI Build GRCh37/hg19.

Figure 3.10: Chromosome 15 replicates normally in P268 cells. P268 cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a whole-chromosome-15 paint (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). **A**) Mitotic chromosome spread with DNA stained with DAPI. **B**) The spread from panel **A** stained for chromosome 15, BrdU and DAPI. Each chromosome 15 is labeled A, B or C (note that chromosome 15 is triploid in P268 cells). **C**) Chromosomes A, B and C from panel **B** were cut out and aligned with each color displayed separately or in combination. Note the comparable BrdU staining between the three chromosomes. **D**) Quantification of the BrdU incorporation in seven different cells. The red bars represent each chromosome 15 in each cell. The y-axis values represent the total number of pixels (area x intensity) x 1000. A comparison of the three 15s in each cell reveals similar amounts of BrdU incorporation, indicating that chromosome 15 replicates synchronously in P268 cells. (Figure on next page).

Figure 3.10: Chromosome 15 replicates normally in P268 cells.

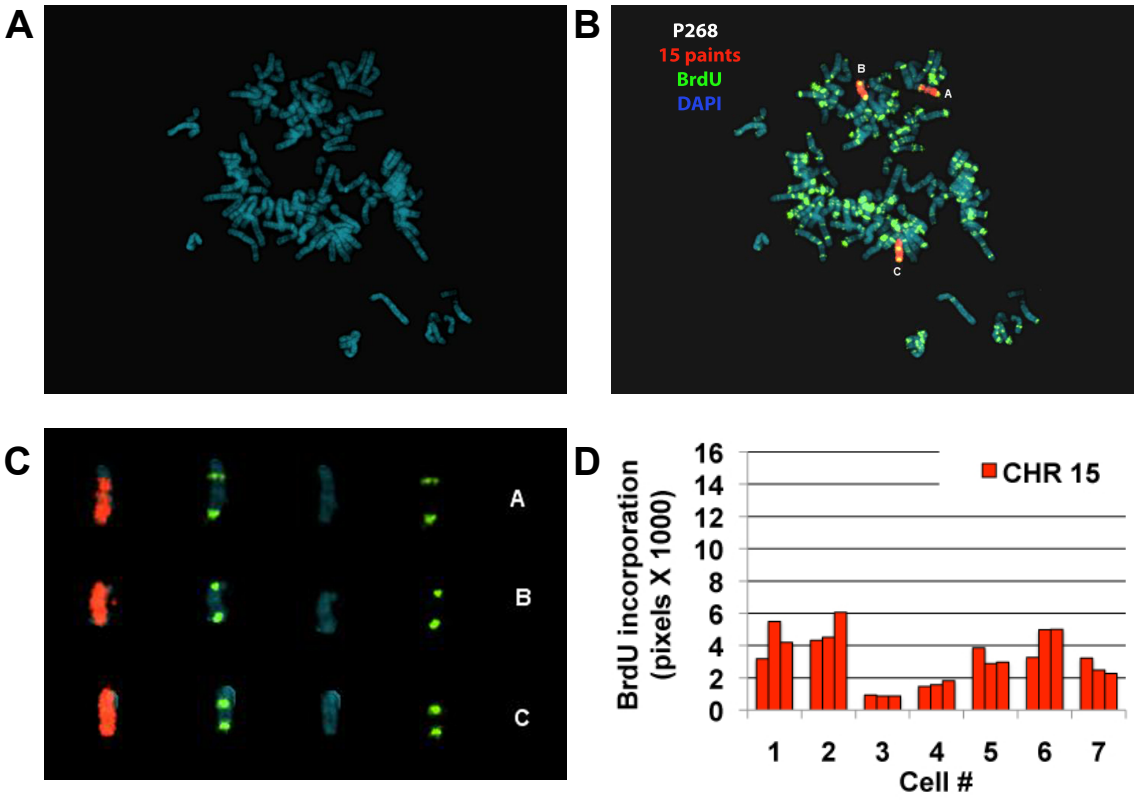


Figure 3.11: Delayed replication timing on an ~161 Kb deletion. ∆268-4g cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a chromosome-15 centromeric probe (red) and a BAC that lies within the deleted region of *SCAPER* (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). **A**) Mitotic spread of a cell containing an ~161 kb deletion distal from the *lox*P site on chromosome 15 stained with DAPI, and both probes. The three chromosome 15s are labeled i, ii and iii (arrows). **B**) Same spread in panel **A** with BrdU staining. **C**) The three chromosome 15s from panel **B** were cut out and aligned. The deleted chromosome 15 (top chromosome with no SCAPER BAC signal) is indicated by an asterisk. Pixel-intensity profiles of BrdU incorporation (green line) and DAPI staining (blue line) along the three chromosome 15s from the left panel. The lower right table shows the pixel-intensity (area x average intensity) for each chromosome showing the total amount of BrdU incorporation or DAPI staining. **D**) Quantification of the BrdU incorporation in seven different cells. The blue bars represent the intact chromosome 15s and the red bar represents the deleted chromosome 15. The y-axis values represent the total number of pixels (area x intensity) x 1000. Note how the deleted chromosome contains much more BrdU incorporation in late S phase than the other homologs. **E**) Chromosome 15 deletions led to multiple translocation events involving chromosome 15. ∆268-4g cells were subjected to DNA-FISH using a whole-chromosome-15 paint (red). Arrows point to all regions of hybridization to the 15-paint. (Figure on next page).

Figure 3.11: Delayed replication timing on an ~161 Kb deletion.

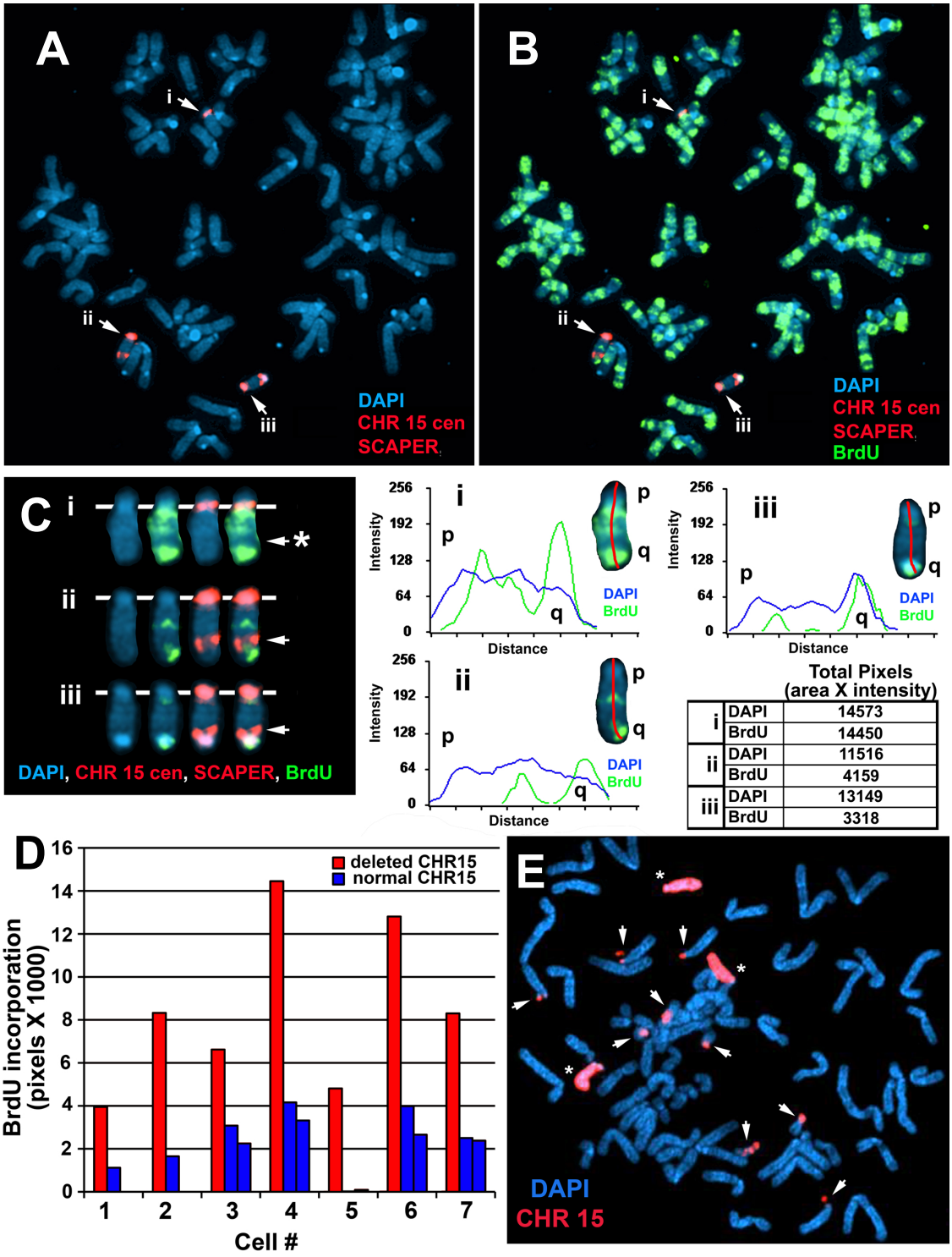


Figure 3.12: Chromosome 15 deletions cause DRT/DMC. A) ∆268-4c cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a chromosome-15 centromeric probe (red) and a BAC that lies within the deleted region (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Note how a chromosome-15 derivative displayed DRT and DMC. **B**) ∆268-4f cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a whole-chromosome-15 paint (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Note how a chromosome-15 derivative displayed extremely late replication. **C**) ∆268-5c cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a chromosome-15 centromeric probe (red) and a BAC that lies within the deleted region (red), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Note how a chromosome-15 derivative displayed DRT and DMC. (Figure on next page).

Figure 3.12: Chromosome 15 deletions cause DRT/DMC.

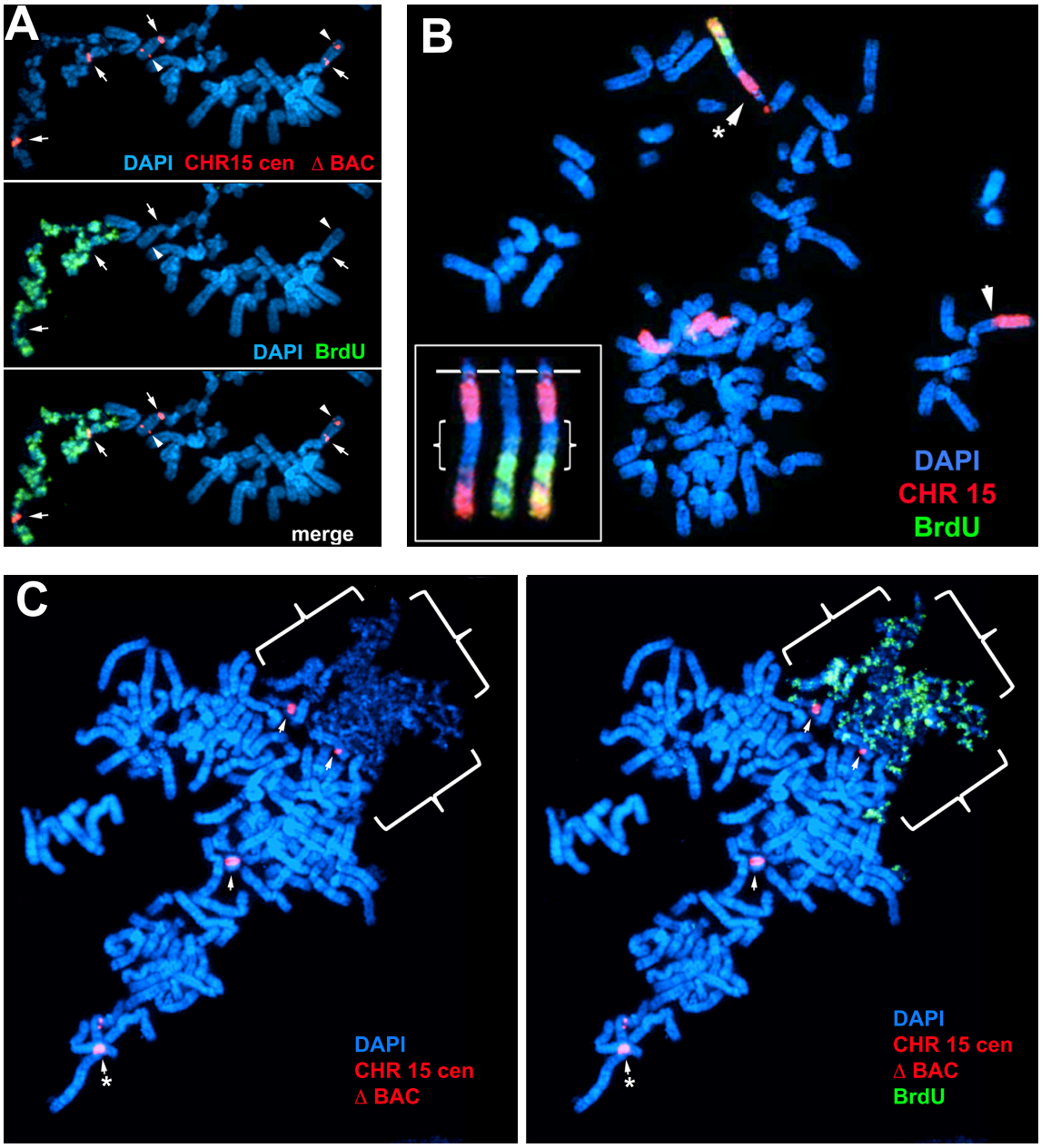


Figure 3.13: Some chromosome 15 deletions do not cause DRT/DMC. A) ∆268-4e cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a chromosome-15 centromeric probe (red) and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). **B**) The three chromosome 15s from panel **A** were cut out and aligned with each color displayed separately or in combination. Note the comparable BrdU staining between the three chromosomes. **C**) Quantification of the BrdU incorporation in seven different cells. The red bars represent each chromosome 15 in each cell. The y-axis values represent the total number of pixels (area x intensity) x 1000. A comparison of the three 15s in each cell revealed similar amounts of BrdU incorporation, indicating that chromosome 15 replicated synchronously in ∆268-4e cells. **D**) Box plot representing chromosome-15 replication asynchrony in P268, ∆268-4e and ∆268-4d cells. For each mitotic spread, a ratio was calculated for the total BrdU signal present on two of the chromosome 15s. A ratio of 1 indicates that the two chromosomes are at the exact same stage of replication and a ratio of <1 indicates replication asynchrony. Seven mitotic spreads were scored for each of the three cell lines and the mean ratios are labeled in each box. P-values were calculated using the Mann-Whitney U test. ∆268-4e did not display a significant replication delay on the deleted chromosome 15 ($p=0.44$) and ∆268-4d did ($p=0.003$) when compared to the parental cell line. (Figure on next page).

Figure 3.13: Some chromosome 15 deletions do not cause DRT/DMC.

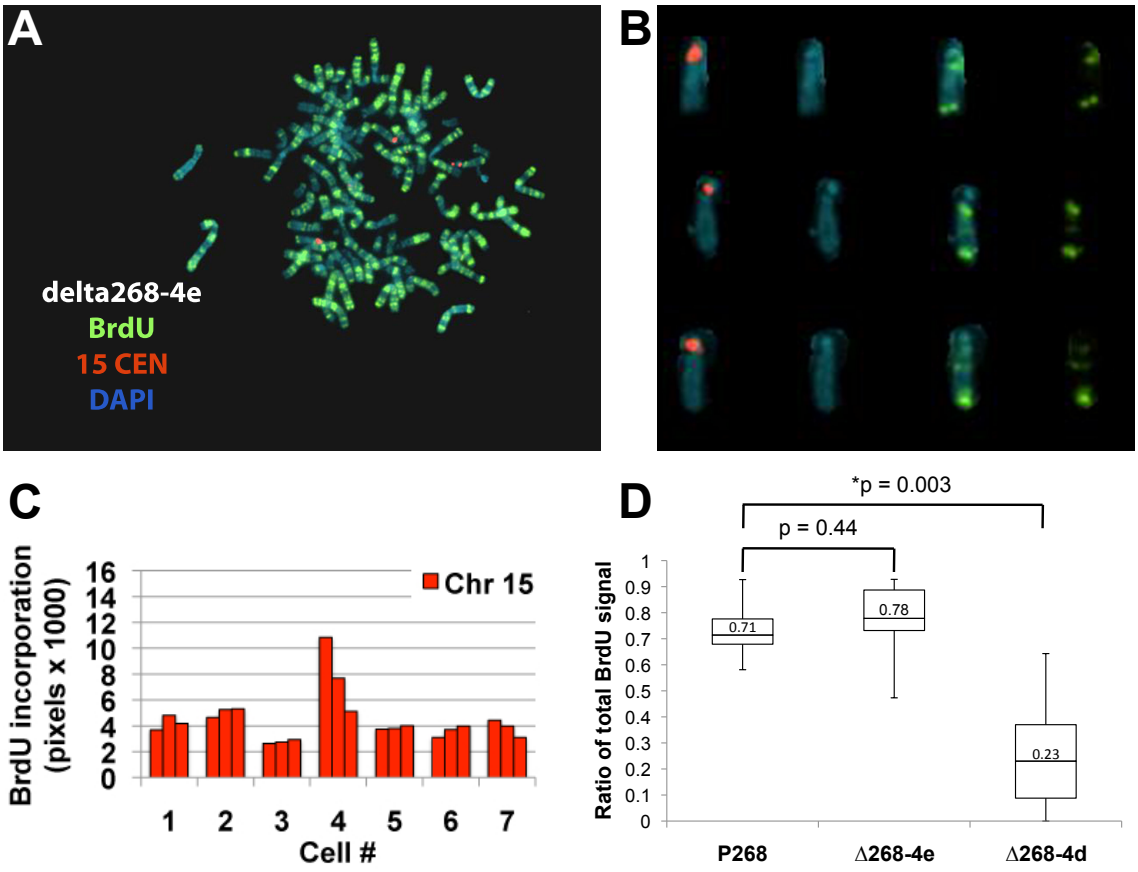
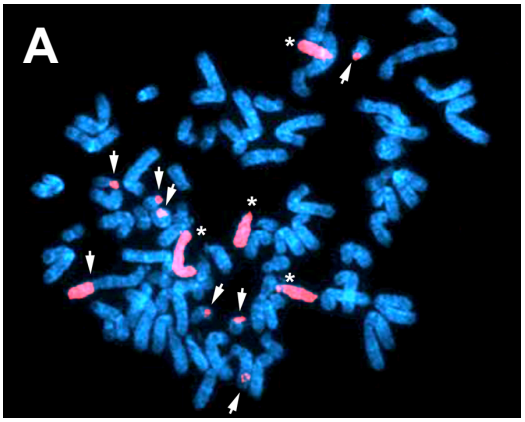
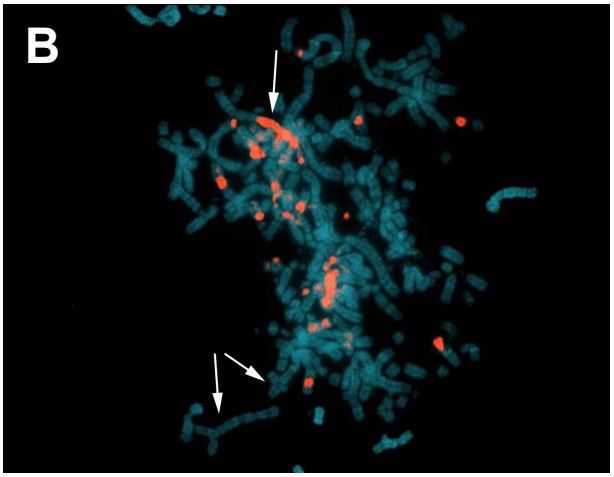
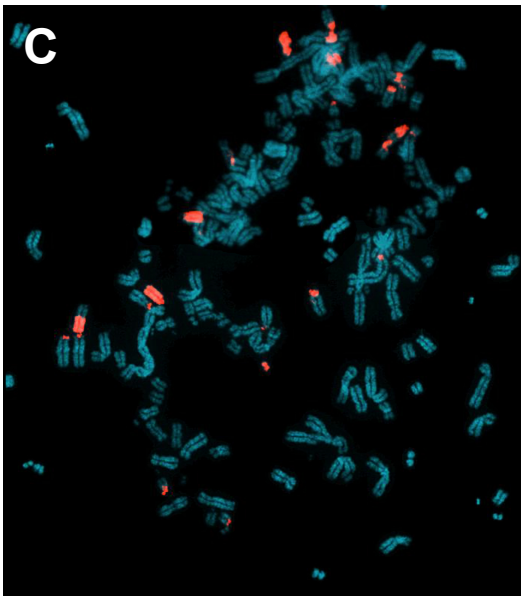


Figure 3.14: Chromosome 15 rearrangements induce abnormal karyotypes. **A**) ∆268-4f mitotic cells were harvested and processed for DNA-FISH using a whole-chromosome-15 paint (red). The DNA was stained with DAPI (blue). Despite having three intact chromosome 15s before an ~135 kb deletion, this cell now has four intact chromosome 15s (asterisks) and multiple chromosome 15 rearrangements (arrows). **B-C**) Inv268-6c mitotic cells were harvested and processed for DNA-FISH using a whole-chromosome-15 paint (red). The DNA was stained with DAPI (blue). This inversion on chromosome 15 has induced multiple chromosome 15 rearrangements and resulted in an increased number of chromosomes. **D**) Chromosome structure abnormalities observed in clones that displayed DRT. Inversion or deletion clones that displayed DRT or were too unstable to analyze for DRT were assayed for new translocations involving chromosome 15. One hundred cells for each clone were scored for whether they exhibited at least one new translocation involving chromosome 15. The percentage of cells that were translocation-positive for each clone is indicated on the right. Some clones had very few cells that had three intact 15s and, therefore, could not be analyzed for DRT (too unstable). Data on P268 new translocations was taken from [77]. (Figure on next page).

Figure 3.14: Chromosome 15 rearrangements induce abnormal karyotypes.







D

Δ268-clones:	DRT	% New Translocations
Δ268-1c	too unstable	100%
Δ268-3g	too unstable	100%
Inv268-3a	too unstable	97%
Δ268-4m	too unstable	97%
Δ268-5d	too unstable	96%
Δ268-5c	too unstable	76%
Δ268-4k	DRT	58%
Δ268-4f	DRT	50%
Inv268-6c	DRT	42%
Inv268-3c	DRT	38%
Δ268-4c	DRT	32%
Δ268-4g	DRT	12%
Δ268-4d	DRT	8%
Δ268-4p	DRT	8%
P268	No DRT	0%

Figure 3.15: Delayed replication and abnormal karyotype in HCC1143. A) Track from the Sanger Cancer Genome Project website showing LOH and copy number data for chromosome 15 in the HCC1143 cell line. The upper (dark blue) line indicates total copy number (on the y-axis) and the lower (light blue) line indicates minor allele copy number. When the light blue line drops to zero, there has been LOH at that region. The location of SCAPER is indicated with an arrow and a black line. **B**) Magnified image from panel **A** showing the location of SCAPER and the two BACs used in the analysis in panel **C**. **C**) DNA-FISH analysis of HCC1143 cells probed for BAC #1 (red) and BAC #2 (green). Although some chromosomes have co-localization of the two BACs (arrows), some have either only green or only red signals (#), indicating that a translocation event(s) took place between the two BACs. **D-G**) Replication-timing analysis of HCC1143. HCC1143 cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a whole-chromosome-15 paint (red) and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Each color is displayed separately or in combination. Arrows point to the chromosome 15 derivatives and asterisks (*) denote the two late replicating chromosomes. Note that there were two regions of extremely late replication in this cell: one was a chromosome 15 derivative and the other was from unknown origin. (Figure on next page).

Figure 3.15: Delayed replication and abnormal karyotype in HCC1143.

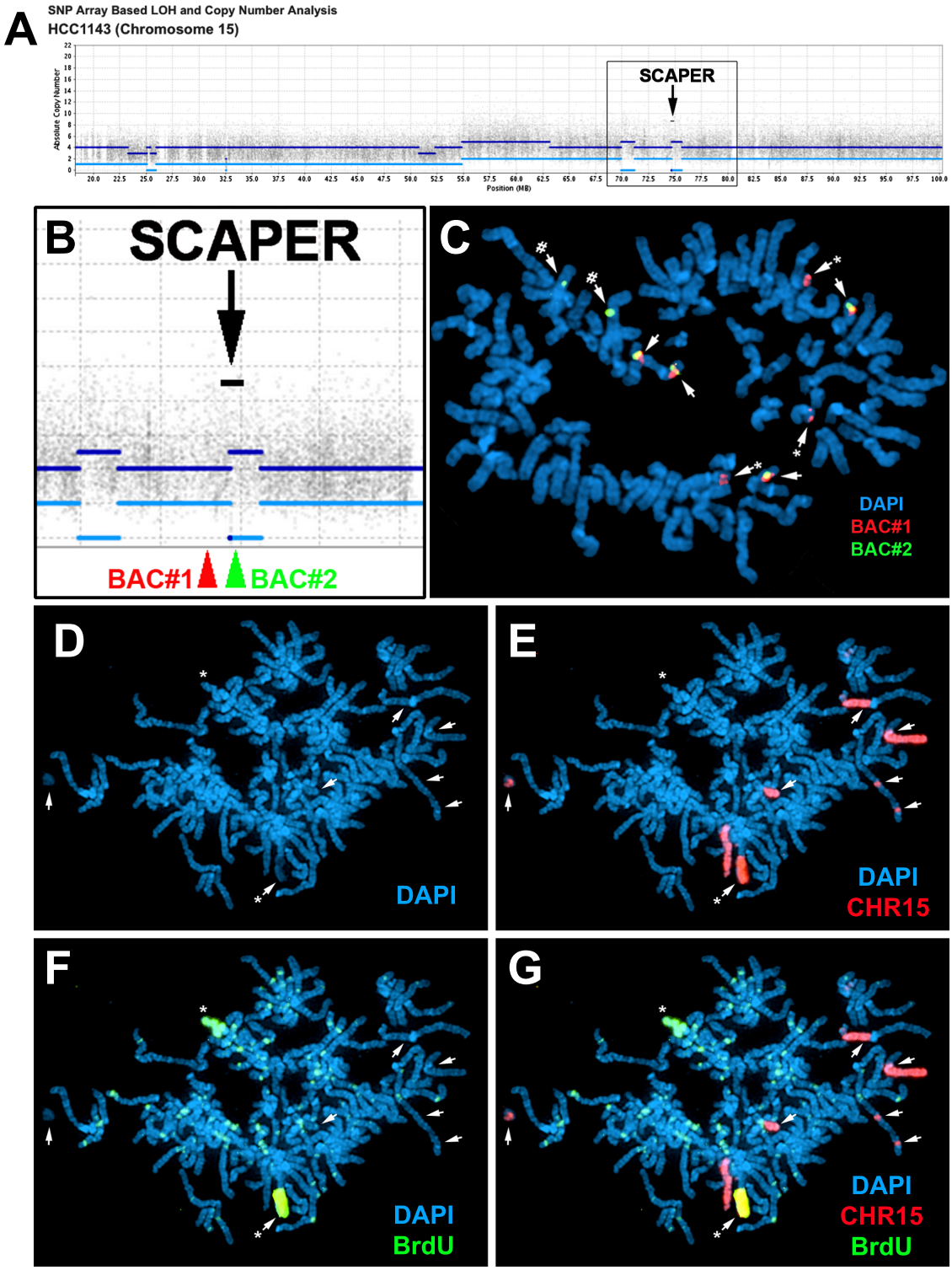


Figure 3.16: Delayed replication and abnormal karyotype in HCC1395. A) Track from the Sanger Cancer Genome Project website showing LOH and copy number data for chromosome 15 in the HCC1395 cell line. The upper (dark blue) line indicates total copy number (on the y-axis) and the lower (light blue) line indicates minor allele copy number. When the light blue line drops to zero, there has been LOH at that region. The location of SCAPER is indicated with an arrow and a black line. **B**) Magnified image from panel **A** showing the location of SCAPER and the two BACs used in the analysis in panel **C**. **C**) DNA-FISH analysis of HCC1395 cells probed for BAC #1 (green) and BAC #2 (red). Although some chromosomes have co-localization of the two BACs (arrows), one has only a green signal (*), indicating that a translocation or deletion event took place between the two BACs. **D-G**) Replication-timing analysis of HCC1395. HCC1395 cells were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a whole-chromosome-15 paint (red) and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Each color is displayed separately or in combination. An arrow points to a chromosome 15 derivative that displayed late replication. (Figure on next page).

Figure 3.16: Delayed replication and abnormal karyotype in HCC1395.

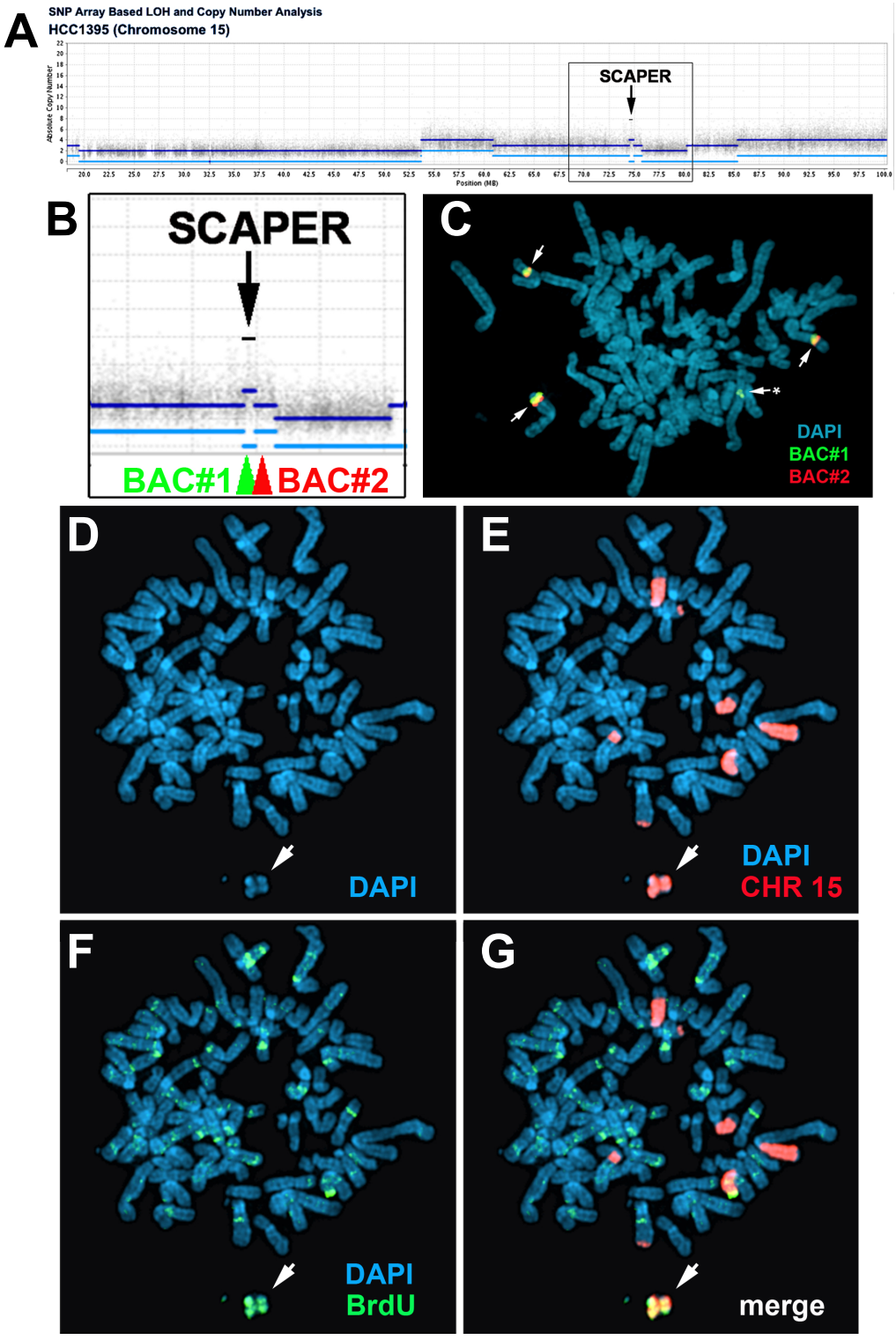
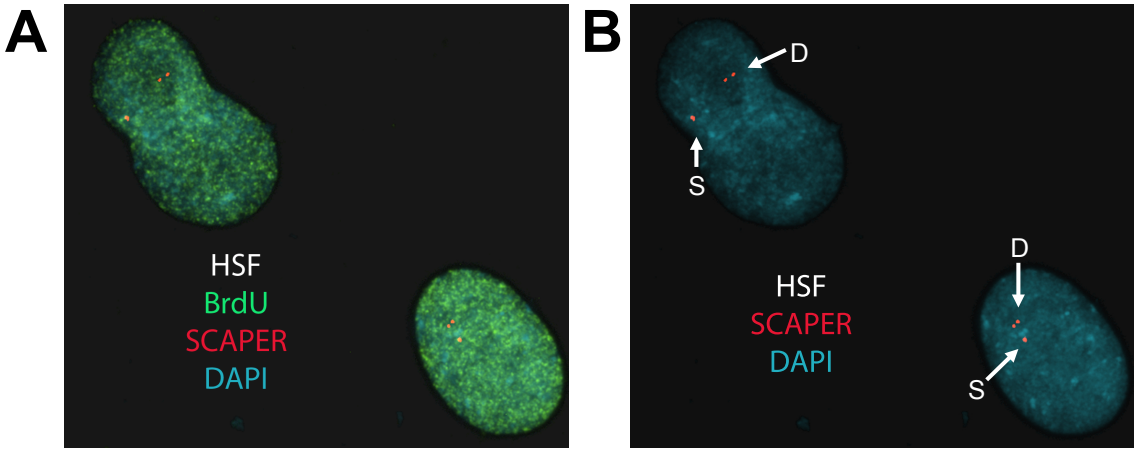


Figure 3.17: Replication asynchrony and coordination on chromosome 15. **A-B**) HSFs were incubated with BrdU, harvested for mitotic cells, processed for DNA-FISH using a BAC within SCAPER (CTD-2299E17, red) and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). Only BrdU-positive cells were used in the SD assay. In these two cells, SCAPER has been replicated on one allele (D) and not the other (S), which indicates replication asynchrony. **C**) Table outlining coordination of replication asynchrony. Each locus on the left was scored for the percentage of BrdU-positive cells that displayed an SD pattern. The data on synchronously-replicating loci, PTK6, LARP and C9orf43, was taken from [122]. On the right shows the percentage of cells that display replication coordination with BAC CTD-2299E17. (Figure on next page).

Figure 3.17: Replication asynchrony and coordination on chromosome 15.



C

Locus	Probe	Position	%SD	Coordination	
				cis	trans
NOR*	18s rDNA	15p	N/A	85%	15%
MYO1E	RP11-1089J12	15q22.2	40	72%	28%
PTPN9	CTD-2323K18	15q24.2	48	71%	29%
SCAPER	CTD-2117F7	15q24.3	35	89%	11%
SCAPER	CTD-2299E17	15q24.3	47	N/A	N/A
PEAK1	RP11-94P14	15q24.3	35	74%	26%
PTK6	RP11-95N13	20q13.3	19	N/A	N/A
LARP	CTD-2546N4	5q33.2	24	N/A	N/A
C9orf43	CTD-2167I10	9q32	19	N/A	N/A

The percentage of the single-double (%SD) pattern was determined using FISH. Coordinated asynchronous replication was scored against BAC CTD-2299E17.
*** indicates coordination measured by Re-TISH**

Figure 3.18: Replication coordination between *SCAPER* and the NOR on chromosome 15. A) Outline of the ReTISH assay. PBLs were pulsed with BrdU for the entirety of S-phase and G2 (14 hours) or just late S-phase and G2 (5 hours) and harvested for mitotics. The 5-hour timepoint identifies which allele is the late-replicating allele for any asynchronously-replicating gene. The combination of multiple probes can uncover if the late-replicating alleles of multiple genes are on the same chromosome. Panel A was adapted from [123]. **B-E**) Mitotic spreads were hybridized with three different FISH probes. First, the ReTiSH assay included a centromeric probe on chromosome 15 (the two larger green signals). Each assay also included BAC probes representing the 18s rDNA in the NOR (red) and SCAPER (RP11-356A10, smaller green signal). The NOR exists on 5 different chromosomes (13, 14, 15, 21 and 22). Both 15s were labeled along with the 13, 14, 21 and 22 alleles that had late-replicating NORs. The chromosome 15 that had the late-replicating NOR is the same one that contained the late-replicating SCAPER. Data are summarized in Figure 3.17C. (Figure on next page).

Figure 3.18: Replication coordination between *SCAPER* and the NOR on chromosome 15.

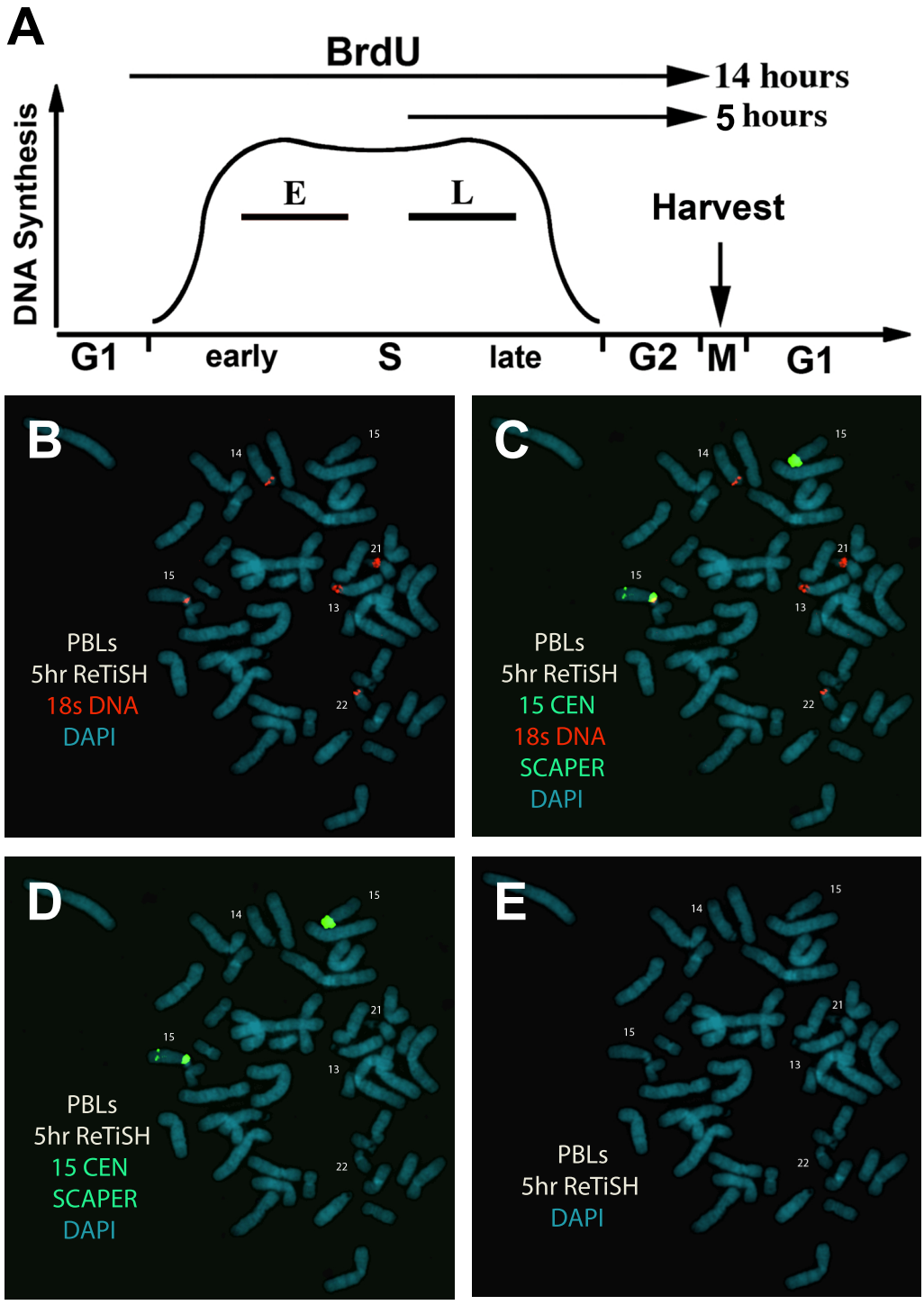
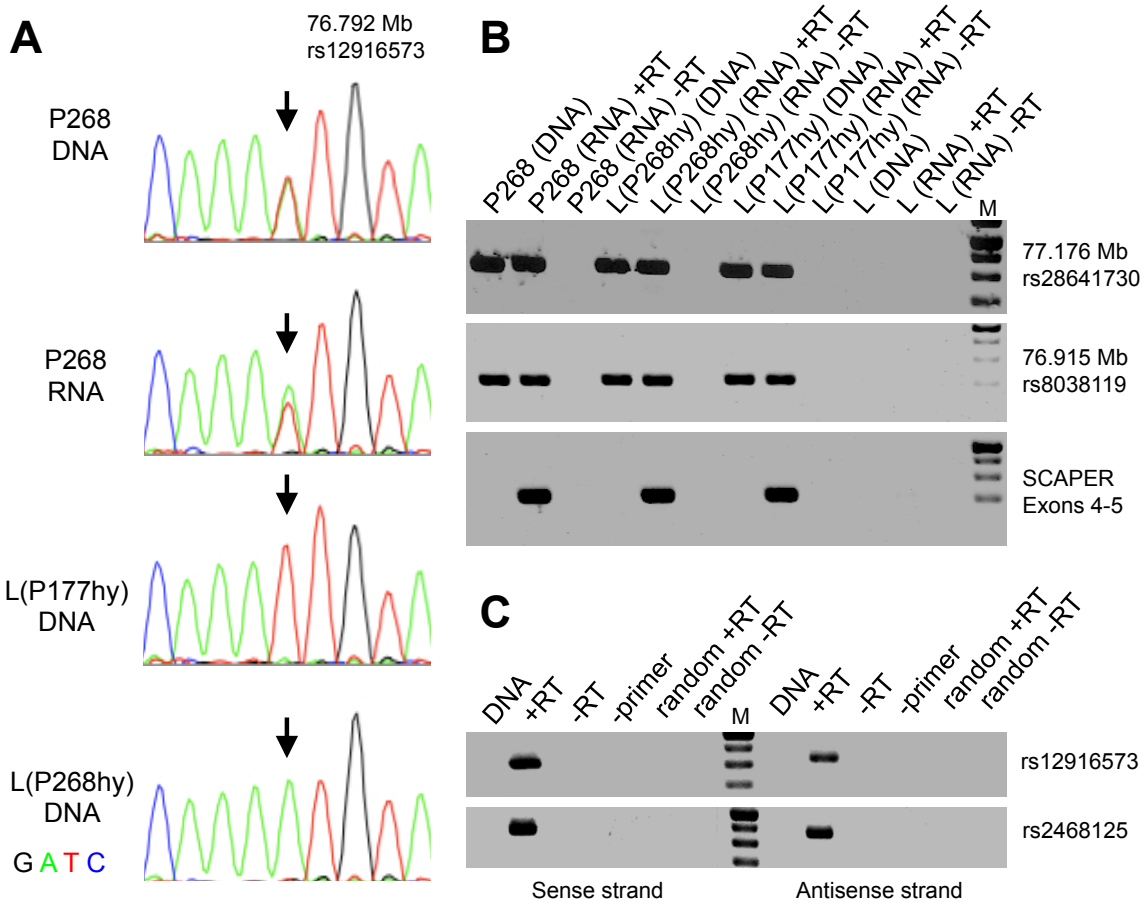


Figure 3.19: Biallelic expression of *SCAPER*. A) DNA sequencing traces from PCRs designed to detect SNP rs12916573. PCRs were carried out on genomic DNA isolated from P268 (top), cDNA from P268 RNA (2nd from the top) and two mono-chromosomal hybrids containing the two different chromosome 15s from HTD114 [L(P268hy)A4 contains the *lox*P-3′RT integrated chromosome 15, and L(P177hy)P5 contains the other chromosome 15] (bottom two). The arrows mark the location of the SNP. **B**) Biallelic expression of SCAPER in monochromosomal hybrids and in P268. RT-PCRs were carried out on RNA isolated from P268, L(P268hy)A4, L(P177hy)P5, and L cells (mouse cells used in hybrid generation). Genomic DNAs from each cell line served as positive controls. PCRs were carried out using primers designed to detect SNP rs28641730 and rs8038119. An additional primer set was used to detect spliced SCAPER by spanning the intron between exons 4-5 of RefSeq SCAPER. Sequencing of the PCR products confirmed that all primers amplified the correct chromosome position (data not shown). **C**) Strand-specific RT-PCR of SCAPER. First strand cDNA was synthesized from sense or antisense SCAPER RNA in HTD114 cells with (+RT) or without (-RT) reverse transcriptase and without cDNA primer (- primer). cDNA was also made with random hexamers with (random +RT) or without (random –RT) reverse transcriptase. Strand-specific primers had a T7 primer tag on the end to ensure specific amplification. cDNA was amplified with T7 and a genomic primer to detect expression of two SNPs (rs12916573 and rs2468125) within SCAPER. Rs2468125 is homozygous in HTD114 cells so allelic expression could not be identified from that SNP. Rs12916573 is heterozygous in HTD114 cells

and sequencing indicated that both the sense and antisense transcripts were biallelically expressed in HTD114 cells (data not shown). (Figure on next page).

Figure 3.19: Biallelic expression of *SCAPER*.



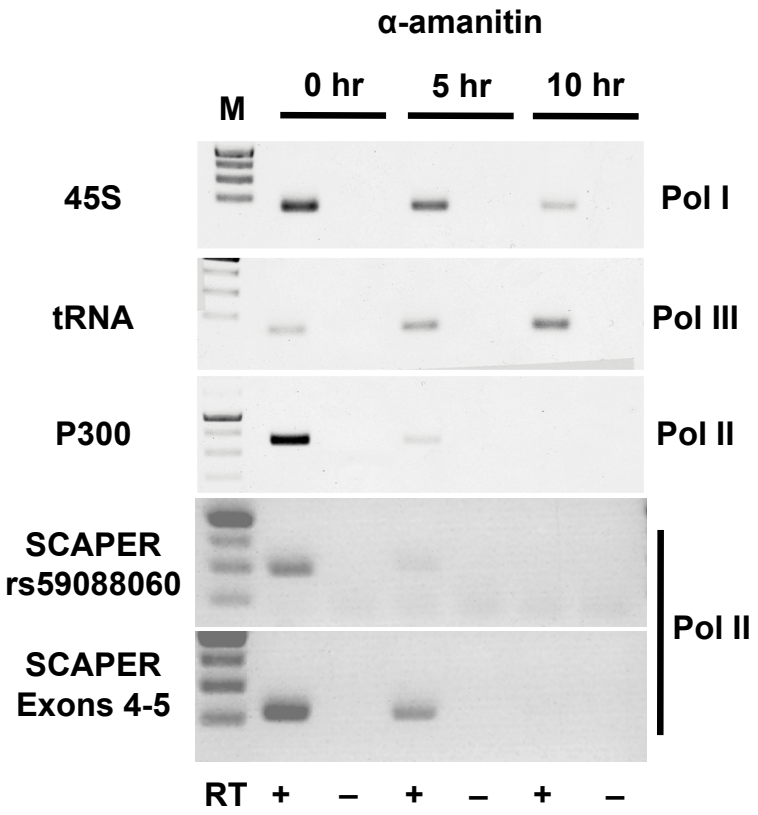


Figure 3.20: *SCAPER* is transcribed by RNA Pol II. P175 cells were exposed to 20 ug/mL of α-amanitin for 0, 5, and 10 hours. Total RNA was subjected to reverse transcriptase reactions (RT) in the presence (+) or absence (-) of reverse transcriptase followed by semi-quantitative PCR using primers to 45S RNA, a tRNA gene, P300, a SNP within SCAPER and primers spanning exons 4-5 of SCAPER. Note that SCAPER expression decreased at 5 hours of treatment and was not detectable after 10 hours of treatment. Figure 3.20 is adapted from [123]. The same cDNA preps used to detect expression of 45S, tRNA and P300 in [123] were used to assay SCAPER expression.

Table 3.7: Expression of genes surrounding the *lox*P-3´RT integration site in P268 cells. Gene expression on chromosome 15 with primers that amplify a heterozygous SNP, a homozygous SNP, or span an intron. For each gene, it is indicated whether or not the DNA feature is expressed (+) or not expressed (-). If the SNP was heterozygous, it was sequenced and the expression pattern was indicated as biallelically expressed (B). Assays used to determine biallelic expression are as follows: r (RT-PCR), h (RT-PCR in hybrid cell lines), s (RT-PCR and sequencing) and as (allele-specific RT-PCR). Bp positions are from the NCBI Build GRCh37/hg19. (Table on next page).

Table 3.7: Expression of genes surrounding the *lox*P-3´RT integration site in P268 cells.

DNA Feature	Location on Chr 15	Gene	P268		L(P268hy)A4	L(P177hy)P5	Assay(s)
			Expressed	Pattern	Expressed	Expressed	
rs8028182	75718669	SIN3A	+	B			r,s
exons 5-8	75782603-75809701	PTPN9	+		+	+	r,h
rs4322627	75810659	PTPN9	+		+	+	r,h
rs55720833	75899640	SNUPN	+		+	+	r,h
rs28412110	75907636	SNUPN	+		+	+	r,h
rs4073214	75948671	SNX33	+	B			r,s
rs10163112	76015397	intergenic	+	B			r,s
rs4503758	76017172	ODF3L1	+	B			r,s
rs2655129	76103729	intergenic	-				r
rs963249	76179923	UBE2Q2	+		+	+	r,h
rs7172445	76373545	C15orf27	+	B			r,s
exons 5-8	76566756-76578783	ETFA	+		+	+	r,h
rs1801591	76578762	ETFA	+		+	+	r,h
exons 1-4	76629242-76632855	ISL2	+		+	+	r,h
exons 3-4	76630664-76632855	ISL2	+		+	+	r,h
rs874224	76631797	ISL2	+		+	+	r,h
rs4886801	76701353	SCAPER	+	B			r,as,s
rs17363364	76732559	SCAPER	+	B			r,s
rs12916573	76791779	SCAPER	+	B	+	+	r,s,h
rs8038119	76914528	SCAPER	+		+	+	r,h
exons 4-5	77096875-77134266	SCAPER	+		+	+	r,h
rs28641730	77175727	SCAPER	+		+	+	r,h
rs28405838	77226053	RCN2	+		+	+	r,h
rs4886509	77354847	TSPAN3	+		+	+	r,h
rs17466989	77416549	PEAK1	+	B			r,ss
exons 3-4	77450856-77471334	PEAK1	+		+	+	r,h
rs12910419	77451350	PEAK1	+	B			r,as,s
rs11856513	77471105	PEAK1	+	B			r,s
rs2100054	77506427	PEAK1	+	B			r,s
rs16968730	77535785	PEAK1	+	B	+	+	r,s,h
rs7182221	77589535	PEAK1	+	B			r,s
rs17384809	77622006	PEAK1	+	B			r,s
rs17470228	77660345	PEAK1	+	B			r,s
rs10519167	77723027	HMG20A	+	B			r,s
rs17471697	77762493	HMG20A	+	B			r,s
rs12917175	77818193	intergenic	+	B			r,s
rs12916515	77818518	intergenic	+	B			r,s
rs12901471	77818992	intergenic	+	B			r,s

SNP	Location	Gene	P268	
			Expressed	Pattern
rs11634496	25139203	SNRPN	-	
rs4243769	25140684	SNRPN	-	
rs2201839	25145346	SNRPN	-	
rs2739835	25415168	intergenic	+	M
rs2014053	25964045	ATP10A	-	
rs3930739	28040342	OCA2	+	B
rs11071087	54843022	UNC13C	-	
rs166362	58733512	LIPC	-	
rs7174277	59442589	MYO1E	+	B
rs12594481	59567137	MYO1E	+	B
rs11631030	68595801	ITGA11	+	B
rs2114716	80454745	FAH	+	B
rs2665118	82396347	intergenic	+	B
rs9972374	93182422	FAM174b	+	M
rs12443091	93193012	FAM174b	-	

Table 3.8: Expression status of predicted or known monoallelically-expressed genes on chromosome 15. Heterozygous SNPs in P268 cells were amplified by RT-PCR and expressed genes were sequenced to determine expression pattern (biallelic=B and monoallelic=M). Monoallelic expression of SNPs rs2739835 and rs9972374 was also confirmed in monochromosomal hybrids (data not shown).

CHAPTER FOUR

Supplementary Analysis of the *ASAR6* Locus Reveals a Complex Genetic Mechanism

Introduction

The Thayer lab has developed a chromosome engineering system that allows for the systematic analysis of human chromosomes with DRT/DMC [75, 77, 80, 122]. To engineer such chromosomes, the lab employed a strategy using the Cre/*lox*P recombinase system. This strategy, as outlined in Figure 1.2, generated random interchromosomal translocations (ICTs) via Cre/*lox*P-mediated, site-specific recombination [77]. Using this system, the Thayer lab previously identified four balanced translocations, each displaying DRT/DMC on one of the two derivative chromosomes (Figure 1.3) [77].

These data led the lab to propose two models for how ICTs can cause DRT/DMC: 1) the ICT could delete or disrupt a *cis*-acting genetic element that acts to ensure proper DNA replication timing, mitotic chromosome condensation, and chromosome stability (loss-of-function model, Figure 4.1A) or 2) the ICT could generate a dominant interfering element that disrupts DNA replication timing, mitotic chromosome condensation, and chromosome stability (gain-of-function model, Figure 4.1B) [74, 77].

The identification of specific chromosomal loci that are involved in the acquisition of DRT/DMC allowed for the first genetic characterization of this phenotype. The P175 cell line was used for further characterization (Figure 1.4). P175 contains a *lox*P site in one allele of chromosome 6 and 10 [77]. Following *lox*P recombination, a balanced translocation, t(6;10)(q14-15;q11.2), was generated that displays DRT/DMC [77]. To identify which *lox*P integration site

was responsible for the acquisition of DRT/DMC (the chromosome 10 locus, the chromosome 6 locus, or both), the Thayer lab generated alternative translocation partners and found that DRT/DMC segregated with the chromosome 6 *lox*P integration site [122]. This indicated that the chromosome 6 locus was required for generating the DRT/DMC phenotype in P175 and the chromosome 10 locus was playing a passive role.

Subsequently, intrachromosomal deletions (ICDs) were made on chromosome 6 using two different methods. One generated deletions anchored at the *lox*P site and extending towards the chromosome 6 centromere (proximal deletions) and the other generated deletions anchored at the *lox*P site and extending towards the q-arm telomere (distal deletions). Deletions of many different sizes were made in each direction and it was found that distal deletions never displayed DRT/DMC, whereas proximal deletions almost always displayed DRT/DMC (Figure 1.5) [122, 123]. All of the proximal deletions that caused DRT/DMC disrupted a monoallelically-expressed ncRNA gene called *ASAR6* [122]. It is currently unknown whether the disruption of the ASAR6 RNA is necessary for this phenotype.

Therefore, the Thayer lab has generated 4 different ICTs (from the alternative partner analysis) and >15 ICDs involving chromosome 6 that cause the DRT/DMC phenotype in *cis* [77, 122]. The Thayer lab has taken the deletion data to support the loss-of-function model (Figure 4.1A) because the loss of genomic information triggers the DRT/DMC phenotype. However, the gain-of-function model has not yet been ruled out (Figure 4.1B) because it is possible

that all of the deletions and translocations were producing interfering elements that dominantly disrupted replication timing by juxtaposing incompatible domains. Furthermore, during the deletion and translocation analysis, it became clear that the chromosomal derivatives that displayed DRT/DMC always contained the distal portion of chromosome 6 (Figure 4.2). Therefore, the DRT/DMC phenotype always segregated with the distal portion of chromosome 6. This observation supports the gain-of-function model because the phenotype would always occur on the chromosome-derivative containing the interfering element (which would be present on the distal portion of chromosome 6) (Figure 4.1B).

As a result, it is currently unknown whether DRT/DMC is caused by the loss-of-function of a timing element or the gain-of-function of an interfering element or a combination of both. Without this basic knowledge, it will be impossible to gain a greater understanding of the mechanism that governs this process. In this chapter, we provide data to support the interfering element (gain-of-function) model. By making both proximal and distal deletions on the same chromosome, we show that distal deletions negated the effects of a subsequent proximal deletion on chromosome 6. Highlighting the complexity of this process, we show that a distal deletion, followed by the original 6;10 translocation in R175, resulted in the opposite chromosome derivative displaying DRT (the 6-centromere derivative instead of the 10-centromere derivative). These data uncover a previously unappreciated aspect of DRT/DMC and open up new avenues for the study of this phenotype.

Results

A Distal Deletion in Chromosome 6 Negates the Effects of a Proximal Deletion

As mentioned in the introduction, the Thayer lab had generated 4 interchromosomal translocations and >15 intrachromosomal deletions involving chromosome 6 that caused the DRT/DMC phenotype in *cis* [77, 122]. During this analysis, it became clear that the chromosomal derivative displaying DRT/DMC always contained the distal portion of chromosome 6 (Figure 4.2). This raises the possibility that the distal region of chromosome 6 can dominantly disrupt replication timing when put in a non-native chromosomal context. In this scenario, the region on the other side of the *lox*P integration site (proximal to the centromere) would be functioning perhaps as an "insulating element," separating the distal portion of chromosome 6 (interfering element) from other regions of the chromosome (Figure 4.1B). On the other hand, it is also possible that these deletions and translocations disrupt a timing element that functions to promote the normal replication-timing program (Figure 4.1A).

To test these two possibilities, I engineered proximal deletions in three different cell lines that already contained distal deletions on chromosome 6 (Δ175F-7a, Δ175F-8a, and Δ175F-12a) (Figure 4.3). The strategy used to generate these "double deletions" is outlined in Figure 4.4. In effect, this strategy generated derivative chromosomes that underwent an ICD distal from the *lox*P site in chromosome 6 followed by an ICD proximal from the *lox*P site (referred to

as double-deletion derivatives). If DRT/DMC is caused by a dominant interfering activity of the distal portion of chromosome 6, perhaps by removing it from an "insulating element" (gain-of-function model), then I expect the double-deletion derivatives to not display the phenotype. This is because the first deletion distal from the original *lox*P site should delete the region that is causing the gain-of-function phenotype; thereby, negating any effect of the second deletion proximal from the *lox*P site. If the simple loss-of-function model is correct, then I expect the double-deletion derivatives to display the DRT/DMC phenotype because the proximal deletion would disrupt the timing element, regardless of the presence of a distal deletion (Figure 4.5).

I have generated double-deletion derivatives in the Δ175F-7a, Δ175F-8a and Δ175F-12a cell lines. All deletions were confirmed using Southern blot hybridizations, junction PCR with primers spanning the *lox*P-genome junction and LOH analysis (Figure 4.6 and data not shown). The size of the distal deletions in each cell line had been determined previously to be 18 Kb, 150 Kb, and 1.3 Mb in the Δ175F-7a, Δ175F-12a, and Δ175F-8a cell lines, respectively [122] (Figure 4.3). Loss of heterozygosity (LOH) analysis was used to estimate the size of the proximal deletions in each clone (Table 4.1).

The replication timing of the chromosome 6s were then analyzed in double-deletion derivative clones from each cell line (Δ7aΔ3a, Δ8aΔ4c and Δ12aΔ4a). Cultures were incubated with BrdU and mitotic cells were harvested, processed for BrdU incorporation and subjected to DNA-FISH using a whole-chromosome-6 paint. Interestingly, DRT/DMC was not detected in any of the

double-deletion derivatives (Figure 4.7). These results suggest that the proximal deletions (removing *ASAR6*) on chromosome 6 are sufficient to induce DRT/DMC only when the distal portion of the chromosome is intact. This is the first example of disruption of *ASAR6* not causing DRT/DMC on chromosome 6. I take these data to support the gain-of-function model.

A Distal Deletion in Chromosome 6 Alters the Effects of the Original t(6;10)

The surprising result that a distal deletion on chromosome 6 negated the effects of a proximal deletion led us to ask whether a distal deletion would have a similar effect on translocation-induced replication delay. From previous studies it was known that a balanced t(6;10) generated in the R175 cell line displayed DRT/DMC on the chromosome 10-centromere derivative (Figure 1.4). Therefore, I induced the t(6;10) in cells that contained distal deletions on chromosome 6, thereby generating chromosomes that have undergone a distal deletion followed by a translocation (deletion-translocation derivatives).

Again, I used the Δ175F-7a, Δ175F-8a, and Δ175F-12a cell lines to induce the t(6;10) as outlined in Figure 4.8. Despite multiple attempts, I was unable to isolate translocation clones from the Δ175F-8a and Δ175F-12a cell lines. Regardless, two clones were isolated from the Δ175F-7a cell line that contained deletion-translocations (R7a-1a and R7a-1b). The t(6;10) was confirmed using Southern blot hybridizations, junction PCR with primers spanning the *lox*P-genome junction and DNA-FISH using whole-chromosome 6 and 10 paints (data not shown).

The replication timing of the derivative chromosomes was analyzed in the two deletion-translocation derivative clones. Interestingly, DNA replication delay was detected on one translocation derivative, as was the case in R175; however, the translocation derivative that displayed DRT was the 6-centromere derivative, opposite of what is seen in R175 cells (Figure 4.9). This result is an unexpected observation and highlights the complicated nature of the ASAR6 locus and the DRT/DMC phenotype.

Discussion

Mammalian cells replicate their genomes every cell cycle during a defined replication-timing program. It is clear that the determinants of replication timing are not encoded within the sequence of the origins of replication, but rather the timing of origin firing is dictated by chromosomal location [10, 189]. Recent studies indicate that at least half of the genome is subject to changes in the temporal sequence of DNA replication during development [15]. The current thinking is that replication timing is directly linked to complex higher-order features of chromosome architecture [13, 190]. However, both the mechanisms and the significance of this temporal replication program remain poorly defined.

Experiments were designed to test the two models that the Thayer lab previously put forth to explain how ICDs and ICTs can cause DRT/DMC. By generating double-deletion and deletion-translocation derivative chromosomes, I have concluded that neither model alone can explain the chromosome-wide replication delay. Although the double-deletion data supports the gain-of-function model, the deletion-translocation data likely reflects the complexity of the mechanisms involved and new models will need to be developed that can explain these observations.

The double-deletion and deletion-translocation data identified the region distal from the *lox*P site on chromosome 6 as playing a vital role in the acquisition of DRT/DMC. Surprisingly, a distal deletion of only 18 kb in the Δ175F-7a cell line was able to negate the effects of a proximal deletion. Due to the small size of this

deletion, the expectation was that this distal deletion would have no effect on subsequent proximal deletions, and the larger distal deletions in Δ175F-12a, and Δ175F-8a cells (150 Kb, and 1.3 Mb) may or may not effect subsequent proximal deletions. Regardless, these data have identified an ~18 kb region on chromosome 6 that has a "dominant interfering" function on replication timing. It is currently unclear how this element may be functioning. Therefore, further experiments to study the activity of this region are warranted. For example, the *ASAR6* transgene used in the experiments in Chapter Two of this thesis contain this 18 kb region. It would be interesting to determine if a BAC that does not contain this 18 kb or a transgene that only contains this 18 kb region could induce DRT/DMC when ectopically integrated into a chromosome.

As mentioned in the Chapter Two discussion, I propose the possibility that the chromosome 6 locus has dual functions. One function would be ASAR6-transcript-dependent and promote late replication via ASAR6 RNA coating and gene silencing. The other function of this locus would be transcript-independent and promote early DNA replication via an unknown mechanism. I believe the data presented in this chapter support this model. The double-deletion derivatives are the first example of chromosomes that have disrupted ASAR6 expression and do not display DRT/DMC. Furthermore, the deletion-translocation derivatives that display replication delay (6-centromere derivatives) have an intact, expressed *ASAR6* gene. This suggests that the *ASAR6* RNA is not involved in the acquisition of the DRT/DMC phenotype and that something else within this locus is the main driver. Since this is the first time that DRT/DMC has

been decoupled from *ASAR6* disruption on chromosome 6, it provides a unique opportunity to investigate the transcript-dependent function of this locus, which may be transcriptional silencing. If the dual function hypothesis is correct, then the double-deletion derivatives should have reactivation of the previously silenced alleles of monoallelically-expressed genes on chromosome 6 even though they do not display DRT/DMC. The idea being that if *ASAR6* is disrupted, as it is in the double-deletion derivatives, then the transcript-dependent function of this locus would be disrupted as well and the silenced alleles of monoallelically-expressed genes would be reactivated.

Importantly, these data do not directly conflict with the 'inactivation/stability center' (I/S center) model mentioned in Chapters One, Two and Three. However, the genetics of the I/S center are likely to be complicated and, therefore, require a more in-depth analysis.

Materials and Methods

Development of Double-deletion Derivative Cell Lines

A plasmid with a pCR2.1 backbone containing an Frt site, Hyg^R gene, *lox*P site, and exons 3,4 and 5 of the mouse APRT gene was generated using common molecular techniques (Frt-hygro-*lox*P-RT in pCR2.1). This plasmid was co-transfected with a Flippase (Flp) expression plasmid into Δ175F-7a, Δ175F-8a, and Δ175F-12a cells [122] with Lipofectamine™ 2000 (Invitrogen) using the manufacturer's protocol. Cells recovered for 48 hours and were subsequently put in media containing 100μg/mL Hygromycin B and 80μg/mL 2,6-diaminopurine (DAP) until single clones could be isolated and expanded. These clones (distal deletion+vector clones) were screened for proper vector integration by PCR and Southern blotting. Distal deletion+vector clones were infected with pL6-frt-5´AP-frt-lox lentivirus at a multiplicity of infection (MOI) of <0.05, allowed to recover for 48 hours and cultured in media containing 7μg/mL Blasticidin S HCL (Invitrogen). Infected clones were pooled and transfected with a Cre recombinase expression plasmid (pBS185 Gibco) and a green fluorescence protein (GFP) expression plasmid with Lipofectamine™ 2000 (Invitrogen) using the manufacturer's protocol. Cells recovered for 48 hours and were subsequently selected for reconstitution of APRT by culturing in media containing 10μg/mL Azaserine and 10μg/mL Adenine. APRT+ clones were isolated and expanded. Clones were screened for proximal deletions by PCR, Southern blotting and LOH analysis.

Development of Deletion-translocation Derivative Cell Lines

Distal deletion+vector clones (described in previous section) were transfected with a Cre recombinase expression plasmid (pBS185 Gibco) and a green fluorescence protein (GFP) expression plasmid with Lipofectamine™ 2000 (Invitrogen) using the manufacturer's protocol. Cells recovered for 48 hours and were subsequently selected for reconstitution of APRT by culturing in media containing 10μg/mL Azaserine and 10μg/mL Adenine. APRT+ clones were isolated and expanded. Clones were screened for the 6;10 translocation by PCR, Southern blotting and DNA-FISH.

PCR and Southern Blotting to Screen for Targeted Recombinants

DNA was isolated from cells and subjected to PCR in a 25-50 μl volume using 50-100 nanograms (ng) of DNA, 1x Standard Taq Buffer (New England Biolabs, Inc.), 200 μM each deoxynucleotide triphosphate, 0.2 μM of each primer, and 3 U of Taq DNA Polymerase (New England Biolabs, Inc.) under the following reaction conditions: 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 45 sec, and 72°C for 1 min, with a final extension time of 10 min at 72°C. PCR products were separated on 1% agarose gels, stained with ethidium bromide, and photographed under ultraviolet light illumination. Heterozygous SNPs used for LOH analysis (Table 4.1) were rs6904580 (~20 Mb away), rs4296866 (~10 Mb away), rs806272 (~5 Mb away), rs9353921 (~3 Mb away), rs2750030 (~1 Mb away), and rs9482760 (~138 Kb away).

BclI endonuclease digestion was performed according to the manufacturer's protocols (Fermentas) and digested DNA was run on a 0.7% agarose gel overnight at low voltage and subjected to Southern blotting using standard protocols. Blots were exposed to film for 10-30 hours at -80ºC.

DNA-FISH

Trypsinized cells were centrifuged at 1,000 rpm for 10 minutes in a swinging bucket rotor. The cell pellet was re-suspended in 75 milliMolar (mM) potassium chloride for 15-30 minutes at 37°C, re-centrifuged at 1,000 rpm for 10 minutes and fixed in 3:1 methanol:acetic acid. Fixed cells were added drop-wise to microscope slides to generate mitotic chromosome spreads using standard methods [192]. Slides with mitotic spreads were baked at 85°C for 20 minutes and then treated with 0.1 mg/ml RNAase for 1 hour at 37°C. After RNAase treatment, the slides were washed in 2xSSC (1xSSC is 150 mM NaCl and 15 mM sodium citrate) with 3 changes for 3 minutes each and dehydrated in 70%, 90%, and 100% ethanol for 3 minutes each. The slides were denatured in 70% formamide in 2xSSC at 70°C for 3 min followed by dehydration in ice cold 70%, 90% and 100% ethanol.

Centromeric probe (CEP) cocktails (Vysis) and whole-chromosome-paints (WCPs) (MetaSystems) were denatured at 75^{O}C for 10 minutes and prehybridized at 37^{O}C for 30 minutes. Probes were applied to slides and incubated overnight at 37^{O}C. Post-hybridization washes consisted of three 3-minute rinses in 50% formamide/2XSSC, three 3-minute rinses in 2XSSC, and

finally three 3-minute rinses in PN buffer (0.1M Na2HPO4 + 0.0M NaH2PO4, ph 8.0, +2.5% Nonidet NP-40), all at 45^{O}C. Slides were then counterstained with either propidium iodide (2.5ug/ml) or DAPI (15ug/ml) and viewed under UV fluorescence (Olympus).

Replication Timing Assay

The BrdU replication timing assay was performed on exponentially dividing cultures as follows: asynchronously growing cells were exposed to 20 ug/ml of BrdU (Sigma) for 5, 5.5 or 6 hours [193]. Mitotic cells were harvested in the absence of colcemid, treated with 75 mM KCl for 15-30 minutes at 37°C, fixed in 3:1 methanol:acetic acid and dropped on wet, ice-cold slides. The chromosomes were denatured in 70% formamide in 2xSSC at 70°C for 3 minutes and processed for DNA-FISH, as described above. The incorporated BrdU was then detected using a FITC-labeled anti-BrdU antibody (Becton Dickinson). Slides were stained with propidium iodide (0.3 mg/ml), cover slipped, and viewed under UV fluorescence.

All images were captured with an Olympus BX Fluorescent Microscope using a 100X objective, automatic filter-wheel and Cytovision workstation. Individual chromosomes were identified by hybridization with WCPs or CEPs. Utilizing the Cytovision workstation, each chromosome was isolated from the metaphase spread and a line drawn along the middle of the entire length of the chromosome. The Cytovision software was used to calculate the pixel area and intensity along each chromosome for each fluorochrome occupied by the DAPI

and BrdU (FITC) signals. The total amount of fluorescent signal was calculated by multiplying the average pixel intensity by the area occupied by those pixels.

Figures and Tables

Figure 4.1: Two models for how interchromosomal translocations and intrachromosomal deletions can cause DRT/DMC. A) Loss-of-function model. This model assumes that every chromosome contains a "timing element" that functions to promote the normal DNA replication-timing program in *cis*. In the P175 cell line, the *lox*P site (blue triangle) on chromosome 6 (red chromosome) would be within or adjacent to the chromosome 6 timing element (yellow rectangle) while the chromosome 10 (green chromosome) *lox*P site would be distant from the chromosome 10 timing element. Following a 6;10 translocation (R175) or a chromosome 6 deletion (not shown), the chromosome 6 timing element would be disrupted resulting in one derivative chromosome displaying DRT/DMC. Therefore, the loss-of-function of a timing element would result in DRT/DMC. **B**) Gain-of-function model. This model assumes that some, possibly all, chromosomes contain an "interfering element" (black rectangle). Normally this interfering element resides adjacent to an insulating element (purple rectangle) and has no negative impact on chromosome function. In the P175 cell line, the *lox*P site on chromosome 6 would lie between an interfering element and its insulator. Following a 6;10 translocation (R175) or a chromosome 6 deletion (not shown), the interfering element becomes detached from its insulator and dominantly interferes with the replication timing of one derivative chromosome (red x). Therefore, the gain-of-function of the aberrant interfering-element activity disrupts DNA replication and results in DRT/DMC. (Figure on next page).

Figure 4.1: Two models for how interchromosomal translocations and intrachromosomal deletions can cause DRT/DMC.

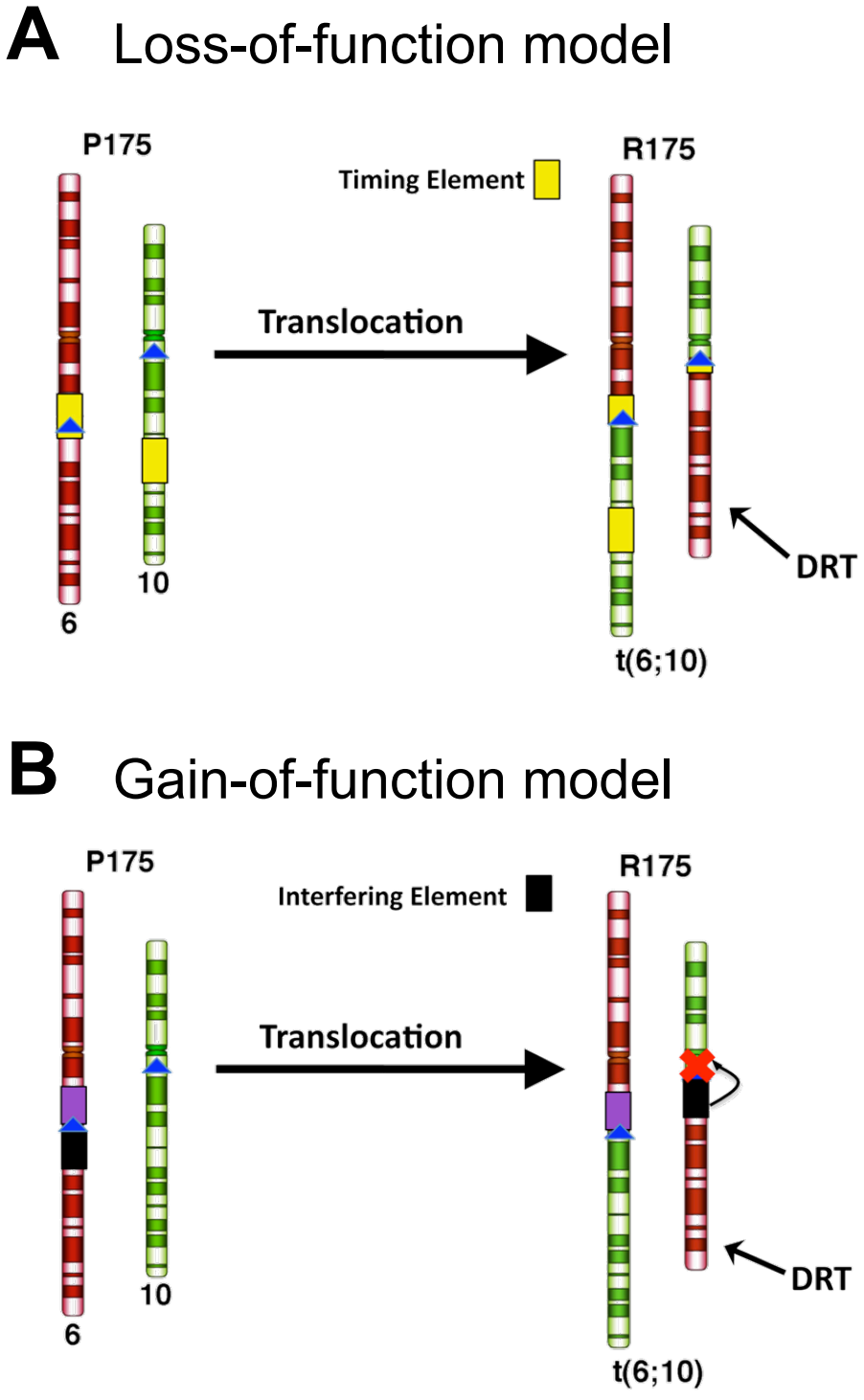
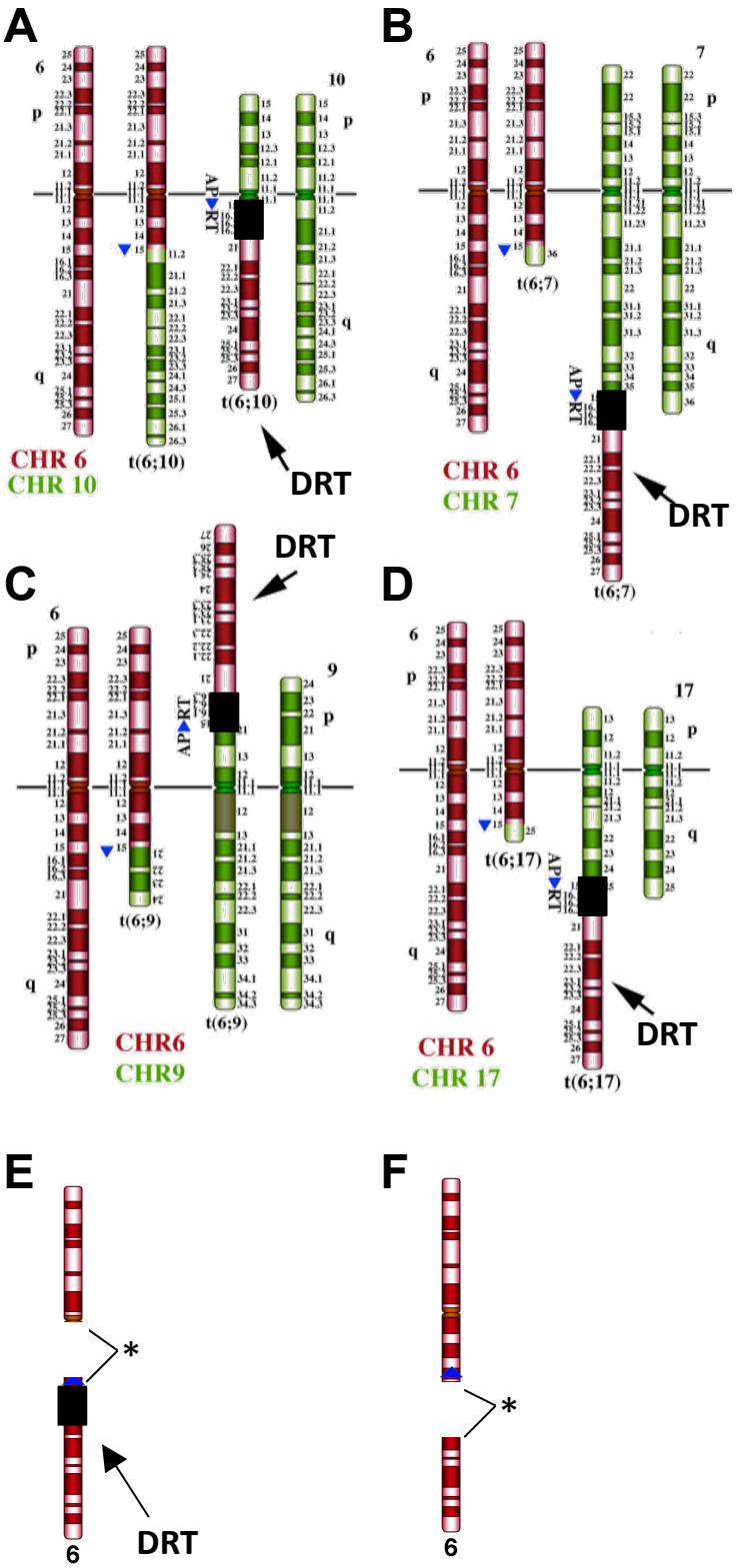


Figure 4.2: Schematic representation of chromosomes that display DRT/DMC in multiple different cell lines. A) Cells that have undergone a 6;10 translocation display DRT on the 10-centromere derivative (arrow) [122]. **B**) Cells that have undergone a 6;7 translocation display DRT on the 7-centromere derivative (arrow) [122]. **C**) Cells that have undergone a 6;9 translocation display DRT on the 9-centromere derivative (arrow) [122]. **D**) Cells that have undergone a 6;17 translocation display DRT on the 17-centromere derivative (arrow) [122]. **E**) Cells that have a chromosome 6 deletion proximal from the *lox*P site display DRT on the derivative chromosome (arrow) [122]. **F**) Cells that have a chromosome 6 deletion distal from the *lox*P site do not display DRT on the derivative chromosome. **A-F**) A black rectangle was added to the distal portion of chromosome 6 on each derivative chromosome for ease of visualization. The distal portion of chromosome 6 is where an interfering element would reside if the gain-of-function model were correct (Figure 4.1B). Note that every derivative chromosome that contains a black rectangle displays DRT and that all of the derivative chromosomes that do not contain a black rectangle do not display DRT. This indicates that the DRT phenotype always segregates with the distal portion of chromosome 6. (Figure on next page).

Figure 4.2: Schematic representation of chromosomes that display DRT/DMC in multiple different cell lines.



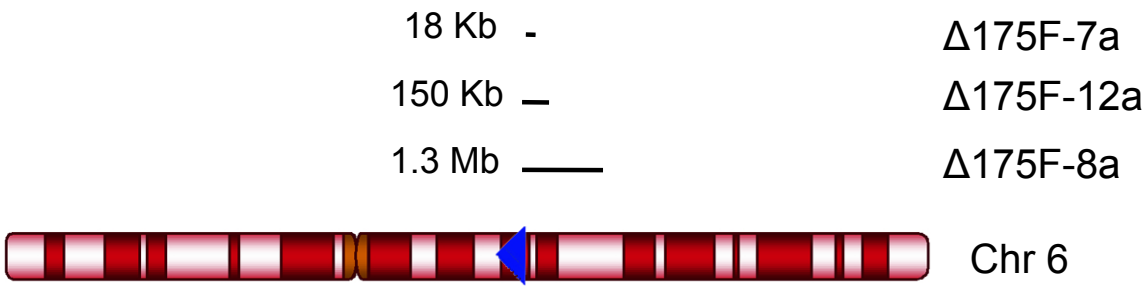


Figure 4.3: Visual representation of chromosome 6 in Δ175F-7a, Δ175F-8a, and Δ175F-12a cells. The Δ175F-7a, Δ175F-8a, and Δ175F-12a cells [122] each have a distal deletion anchored at the *lox*P site (blue triangle) on chromosome 6 (red chromosome) and extending towards the q-arm telomere. The relative deletion size is represented with a black line above the chromosome for each cell line. The Δ175F-7a cell line harbors an ~18 kb distal deletion on chromosome 6, the Δ175F-12a line harbors an ~150 kb distal deletion, and the Δ175F-8a line harbors an ~1.3 Mb distal deletion. Note that the black lines are not drawn to scale.

Figure 4.4: Strategy used to generate double-deletion derivatives. A) The chromosome 6 locus in Δ175F-7a, Δ175F-8a, and Δ175F-12a cells contains exons 1 and 2 of mouse APRT (AP), an Frt site (green triangle), exons 3,4 and 5 of mouse APRT (RT), and a *lox*P site (purple triangle) in an antisense orientation between the *ASAR6* (light blue rectangle) and *FUT9* (dark blue rectangle) genes. These sequences are remnants that were leftover from generating the distal deletions on chromosome 6 (red) [122]. Δ175F-7a, Δ175F-8a, and Δ175F-12a cells are APRT+, Blasticidin (Blas) sensitive and Hyg sensitive. An integration plasmid containing an Frt site, a Hyg^R gene (orange rectangle), a *lox*P site and RT was introduced into Δ175F-7a, Δ175F-8a, and Δ175F-12a cells along with flippase (Flp). **B**) Hyg^R clones were isolated that underwent recombination between the two Frt sites followed by plasmid integration into chromosome 6. Plasmid integration also disrupted the APRT gene that was present in Δ175F-7a, Δ175F-8a, and Δ175F-12a cells. Therefore, this step produced cells that were APRT-, Blas sensitive, and Hyg^R. **C**) Cells from step **B** were infected with a lentivirus containing AP, a *lox*P site, and $Blas^R$ and selected in media containing Blasticidin. APRT-, $Blas^R$ and Hyg^R cells were then transfected with Cre to generate **D**) double-deletion derivatives that are APRT+, Blas sensitive and Hyg sensitive. This leaves an intact APRT transgene with a *lox*P site in the middle and a leftover AP-Frt fragment. Double-deletion derivatives were screened to make sure a proximal deletion occurred on chromosome 6. Note that the proximal deletion generated by this method deletes or disrupts the *ASAR6* gene. Image adapted from [216]. (Figure on next page).

Figure 4.4: Strategy used to generate double-deletion derivatives.

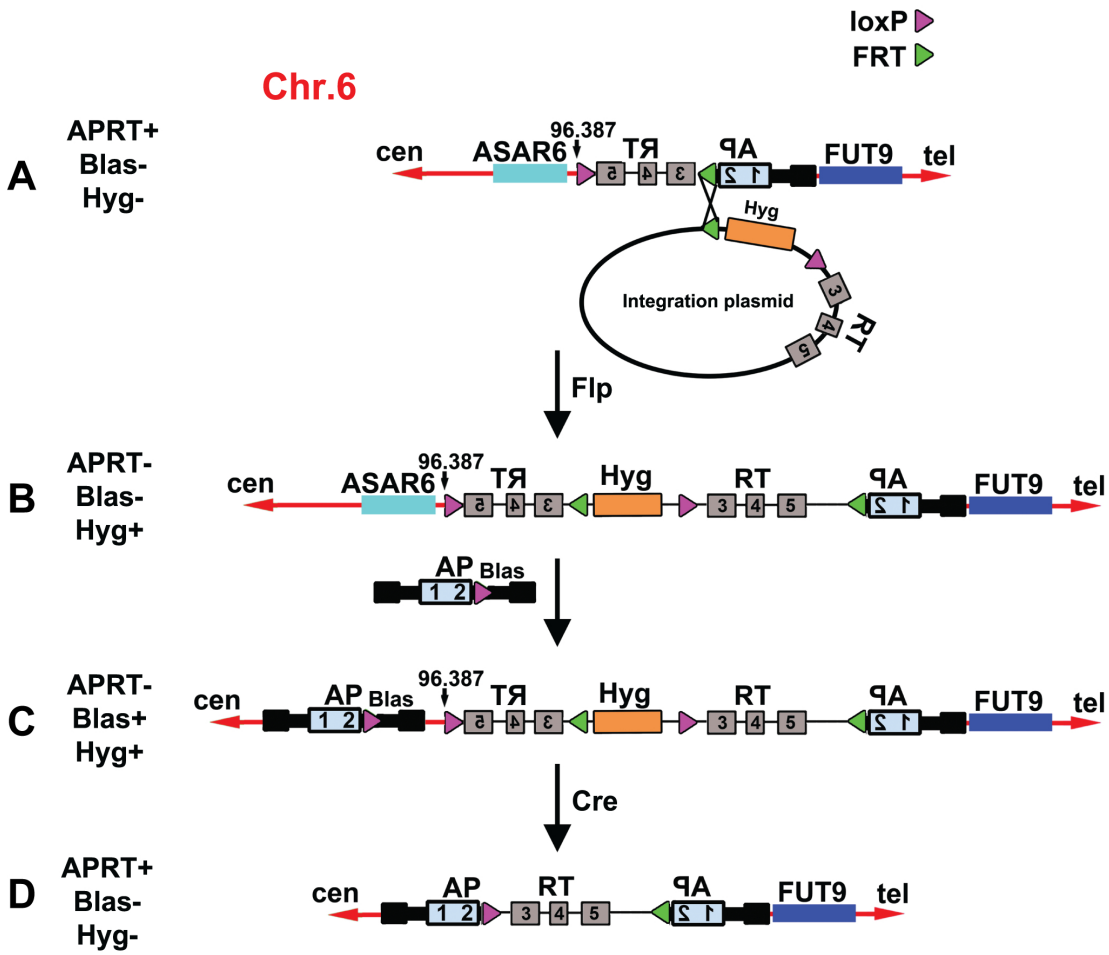


Figure 4.5: Schematic representation of chromosome 6 following a distal and proximal deletion. A hypothetical timing element (yellow rectangle) and interfering element (black rectangle) are aligned on chromosome 6 (red chromosome) in relation to the *lox*P site (blue triangle) that is present in the P175 cell line (Figure 1.4). **Step 1**) Deletions distal from the *lox*P site on chromosome 6 were made in P175 cells to generate the Δ175F-7a, Δ175F-8a, and Δ175F-12a cell lines. These distal deletions would remove the interfering element if one were to exist. **Step 2**) We know from [122] that the distal deletions in the Δ175F-7a, Δ175F-8a, and Δ175F-12a cell lines do not cause DRT. **Step 3**) In this study, we inserted another *lox*P site into chromosome 6 and made deletions proximal from the *lox*P site in the Δ175F-7a, Δ175F-8a, and Δ175F-12a cell lines. **Step 4**) These double-deletion derivatives will be analyzed for DRT. (Figure on next page).

Figure 4.5: Schematic representation of chromosome 6 following a distal and proximal deletion.

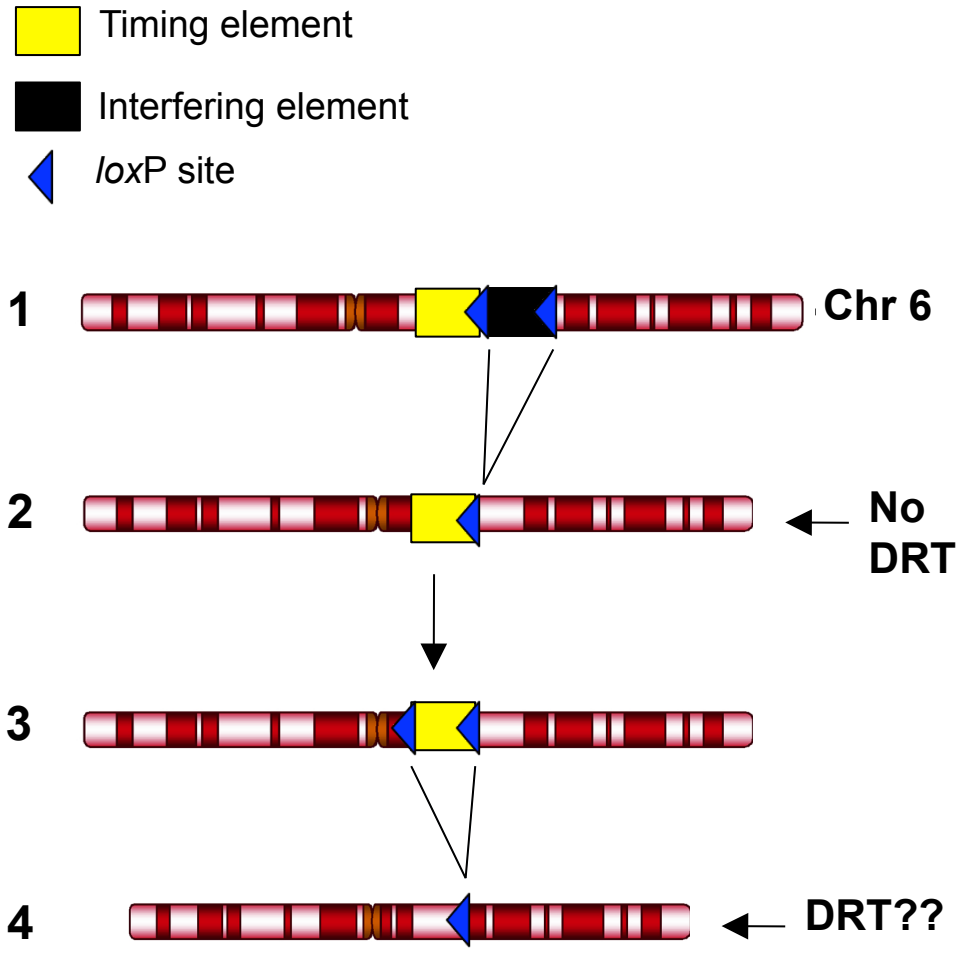
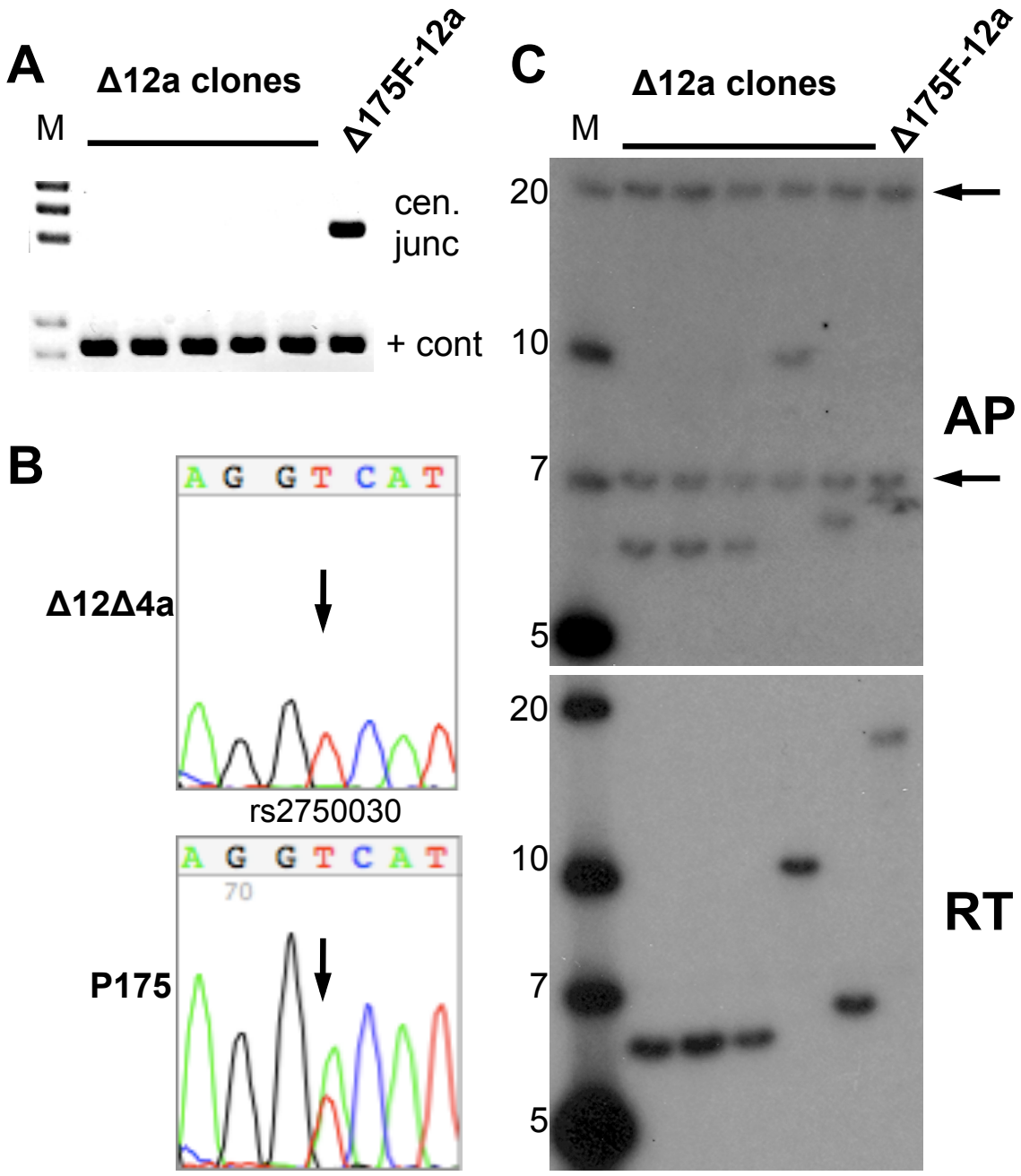


Figure 4.6: Molecular confirmation of double-deletion clones. A) Diagnostic PCR on five double-deletion derivatives derived from the Δ175F-12a cell line. The cen. junc is the junction that will be deleted following a proximal deletion and the + cont is the junction that was formed following the distal deletion. Note how all five clones have lost the cen. junc, indicating they harbor a proximal deletion on chromosome 6. **B**) SNP rs2750030 that is heterozygous in the P175 parent line was assayed in the double-deletion derivative clone Δ12Δ4a. This SNP lies ~1Mb proximal to the centromere from the chromosome 6 *lox*P site. The P175 cell line and the Δ175F-12a cell line harbor both an A and T residue at this position (bottom arrow) while Δ12Δ4a cell line has lost the A allele (top arrow) indicating that Δ12Δ4a cells have a proximal deletion that is larger than 1 Mb. **C**) The same five double-deletion clones in panel **A** were digested with BclI and Southern blotting was performed with a probe to exons 1 and 2 of mouse APRT (AP, top blot) or exons 3,4 and 5 of mouse APRT (RT, bottom blot). The size of the marker in kb is indicated on the left. Note that the Δ175F-12a cell line has two AP bands, the top one is the AP fragment at the chromosome 6 locus (Figure 4.4 A) and the bottom one is the AP fragment on chromosome 10 in these cells. The double-deletion derivatives gained an AP fragment during the lentiviral infection (see Figure 4.4). This new AP fragment co-migrates with the new RT fragment that is generated following the proximal deletion. (Figure on next page).

Figure 4.6: Molecular confirmation of double-deletion clones.



Clone Name	Lentiviral Integration Site (proximal) in bp	Size of proximal Deletion in bp	Size of distal Deletion in bp
Δ7aΔ1a	btw 91,204,215-96,279,600	<5,075,385	18,525
Δ7aΔ2a	btw 91,204,215-96,279,600	<5,075,385	18,525
Δ7aΔ2b	btw 91,204,215-96,279,600	<5,075,385	18,525
Δ7aΔ2c	btw cen-76,114,316	>20,165,284	18,525
Δ7aΔ2d *	btw cen-76,114,316	>20,165,284	18,525
Δ7aΔ3a *	btw 95,278,533-96,141,384	btw 138,216-1,001,067	18,525
Δ7aΔ4a *	btw 95,278,533-96,141,384	btw 138,216-1,001,067	18,525
Δ7aΔ4b	btw 91,204,215-96,279,600	<5,075,385	18,525
Δ7aΔ4c *	btw 95,278,533-96,141,384	btw 138,216-1,001,067	18,525
Δ7aΔ4d *	btw 95,278,533-96,141,384	btw 138,216-1,001,067	18,525
Δ8aΔ1a *	btw 86,279,737-91,204,215	btw 5,075,385-9,999,863	1,371,246
Δ8aΔ1d	btw 76,114,316-86,279,737	btw 9,999,863-20,165,284	1,371,246
Δ8aΔ2a	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ2b	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ2c	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ3a *	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ3b	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ3c *	btw 76,114,316-86,279,737	btw 9,999,863-20,165,284	1,371,246
Δ8aΔ4a *	btw 76,114,316-86,279,737	btw 9,999,863-20,165,284	1,371,246
Δ8aΔ4b *	btw 91,204,215-96,279,600	<5,075,385	1,371,246
Δ8aΔ4c *	btw 93,259,842-95,278,533	btw 1,001,067-3,019,758	1,371,246
Δ12aΔ4a *	btw 93,259,842-95,278,533	btw 1,001,067-3,019,758	152,786
Δ12aΔ4d	btw 91,204,215-96,279,600	<5,075,385	152,786
Δ12aΔ4e	btw 91,204,215-96,279,600	<5,075,385	152,786

Table 4.1: Deletion sizes of double-deletion derivative clones. Clones whose name begins with Δ7a were derived from the Δ175F-7a cell line, Δ8a clones were derived from the Δ175F-8a cell line and Δ12a clones were derived from Δ175F-12a cells. The size of the proximal deletion in each clone was estimated by LOH analysis. Multiple clones exist for those with an asterisk (*) next to their name. Bp positions are from the NCBI Build GRCh37/hg19.

Figure 4.7: Double-deletion derivatives do not display DRT/DMC. A-C) The Δ7aΔ3a (**A**), Δ8aΔ4c (**B**) and Δ12aΔ4a (**C**) cell lines were incubated with BrdU for 6 hours, harvested for mitotic cells, and processed for DNA-FISH using a whole-chromosome-6 paint as a probe (red) and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The chromosome 6s (A and B) from each spread (small panel on top) were cut out and aligned with each color displayed separately or in combination. The graphs on the right measure quantification of the BrdU incorporation in six different cells. The blue bars represent the intact chromosome 6 and the red bars represent the truncated chromosome 6. The values represent the total number of pixels (area x intensity) x1000. Note how both chromosome 6s contain comparable amounts of BrdU incorporation in each cell, indicating there is no replication delay. (Figure on next page).

Figure 4.7: Double-deletion derivatives do not display DRT/DMC.

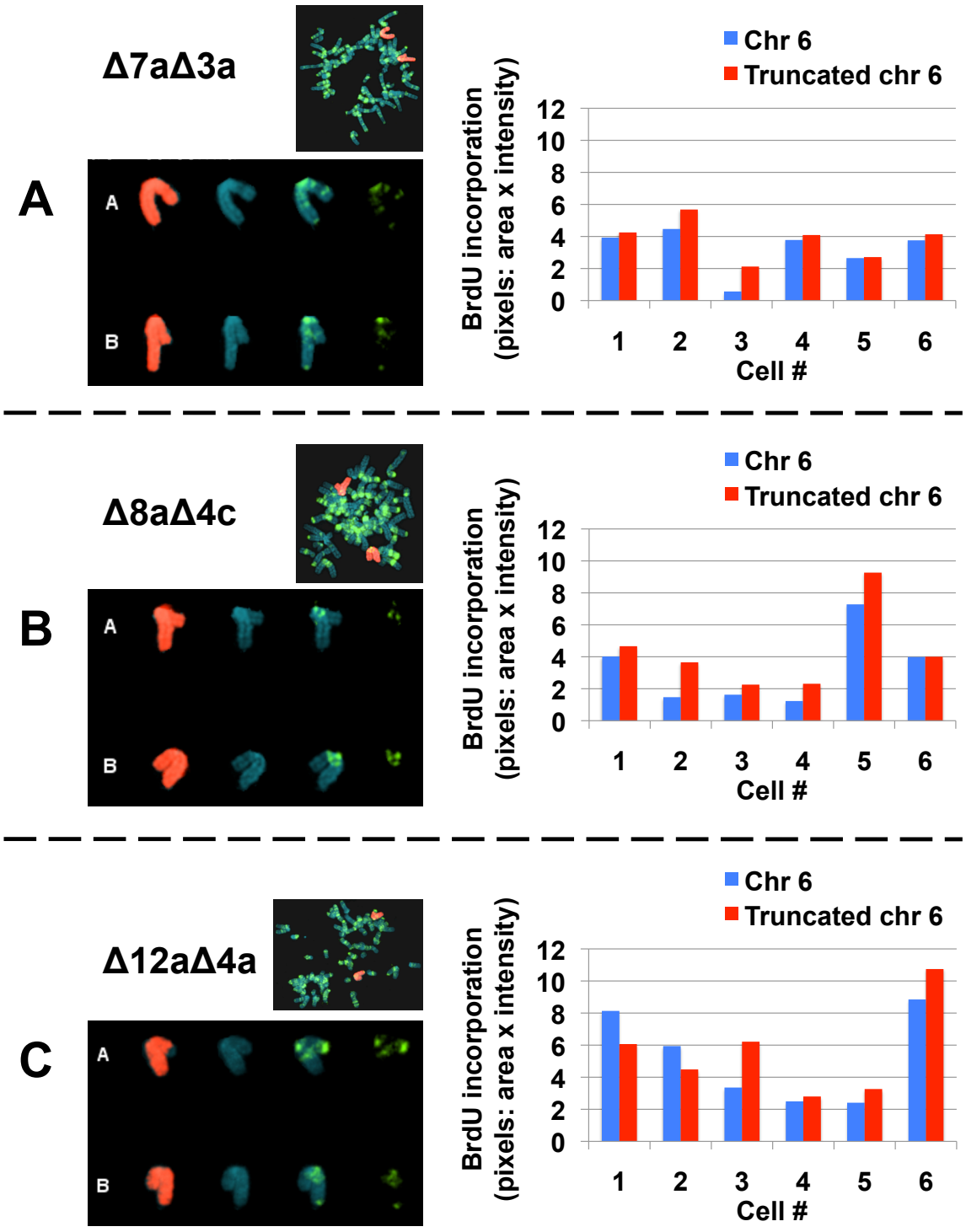


Figure 4.8: Strategy used to generate deletion-translocation derivatives. A) The chromosome 6 locus in Δ175F-7a, Δ175F-8a, and Δ175F-12a cells contains exons 1 and 2 of mouse APRT (AP), an Frt site (green triangle), exons 3,4 and 5 of mouse APRT (RT), and a *lox*P site (purple triangle) in an antisense orientation between the *ASAR6* (light blue rectangle) and *FUT9* (dark blue rectangle) genes. These sequences are remnants that were leftover from generating the distal deletions on chromosome 6 (red) [122]. Δ175F-7a, Δ175F-8a, and Δ175F-12a cells are APRT+ and Hyg sensitive. An integration plasmid containing an Frt site, a Hyg^R gene (orange rectangle), a *lox*P site and RT was introduced into Δ175F-7a, Δ175F-8a, and Δ175F-12a cells along with flippase (Flp). **B**) Hyg^R clones were isolated that underwent recombination between the two Frt sites followed by plasmid integration into chromosome 6. Plasmid integration also disrupted the APRT gene that was present in Δ175F-7a, Δ175F-8a, and Δ175F-12a cells. Therefore, this step produced cells that were APRT- and Hyg^R. **C**) Cells from step **B** were transfected with Cre recombinase and selected for *lox*P recombination. Since there were three *lox*P sites in these cells, two in chromosome 6 and one in chromosome 10 (green), recombination floxed out Hyg^R, FRT and RT and mediated a translocation between chromosome 6 and 10. This produced **D**) Deletion-translocation derivatives that are APRT+ and Hyg sensitive. This leaves an intact APRT transgene with a *lox*P site in the middle and a leftover AP-Frt fragment. Deletion-translocation derivatives were screened to ensure the 6;10 translocation occurred. Image adapted from [216]. (Figure on next page).

Figure 4.8: Strategy used to generate deletion-translocation derivatives.

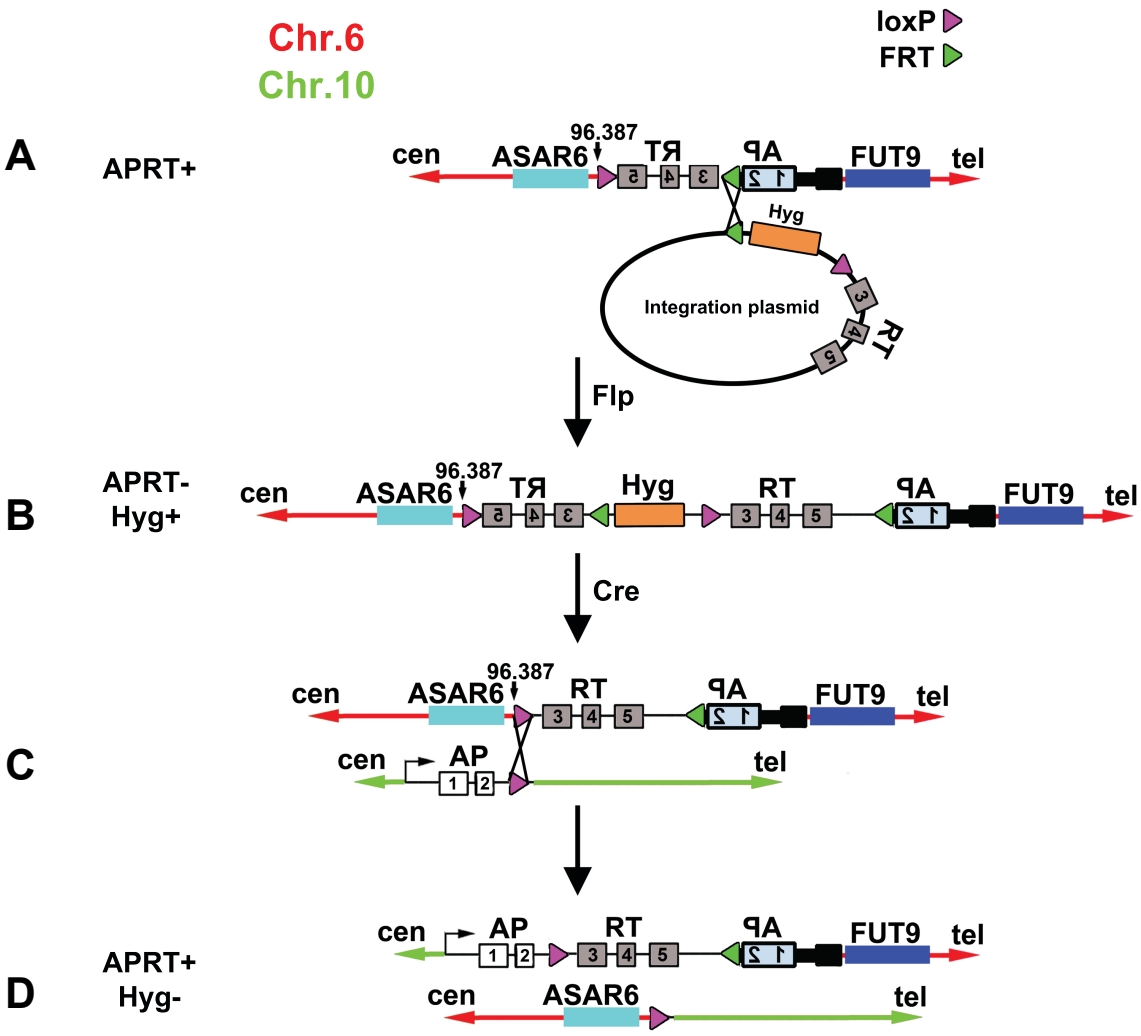
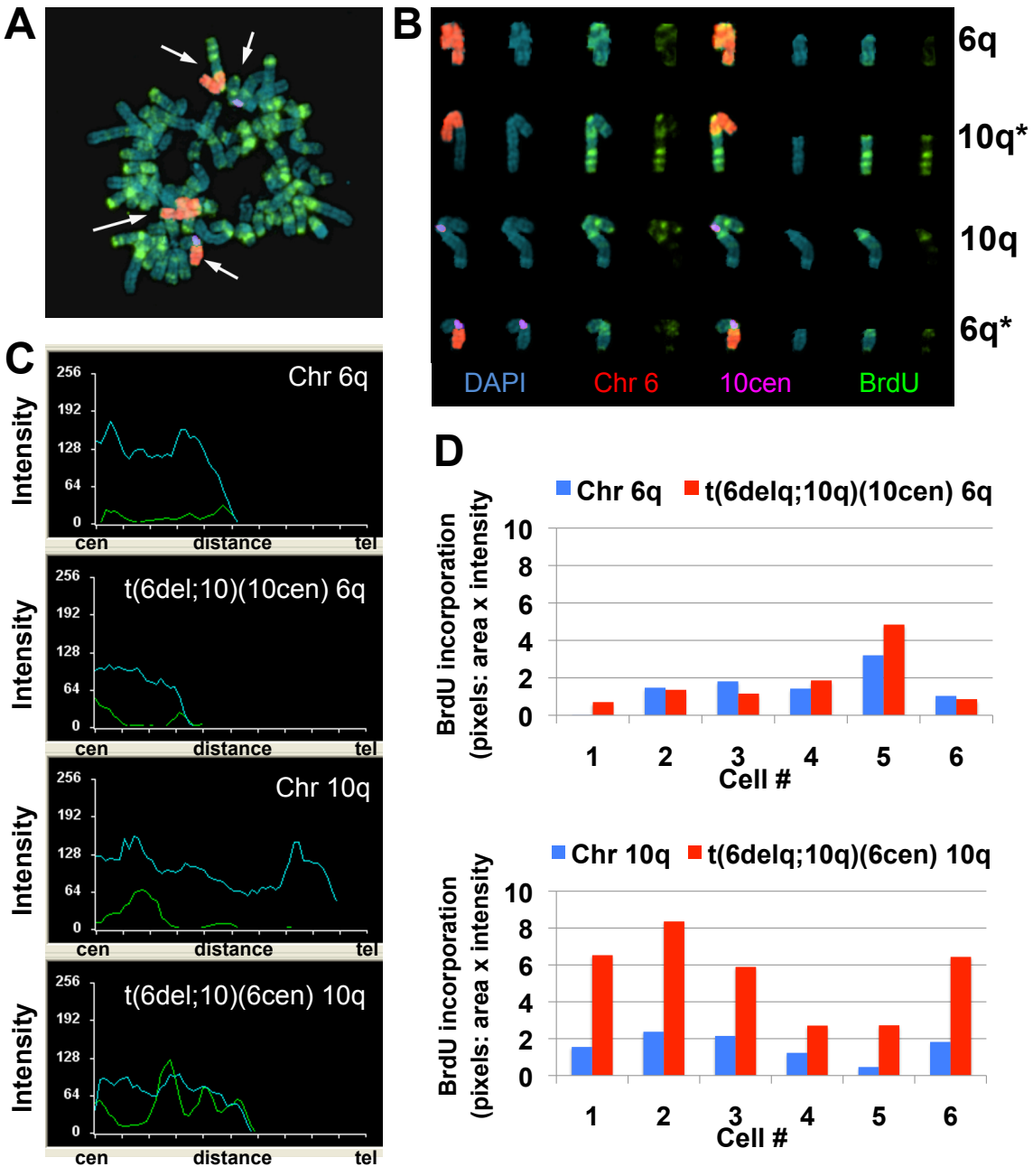


Figure 4.9: Deletion-translocation derivatives display DRT on the 6-centromere derivative. R7a-1b cells were incubated with BrdU for 5 hours, harvested for mitotic cells, processed for DNA-FISH using a whole-chromosome-6 paint (red) and a chromosome-10-centromere probe (purple), and for BrdU incorporation using an antibody against BrdU (green). The DNA was stained with DAPI (blue). The two translocation-derivatives and the intact chromosome 6 and 10 in panel **A** are indicated by arrows. **B**) The chromosomes from panel **A** were cut out and aligned with each color displayed separately or in combination. Chromosome 6 (top) is labeled 6q on the right, the chromosome-6-centromere derivative (2nd from the top) is labeled 10q* on the right, chromosome 10 (3rd from the top) is labeled 10q on the right and the chromosome-10-centromere derivative (bottom) is labeled 6q* on the right. In order to compare the BrdU incorporation between the translocation derivatives and the intact chromosome 6 and 10, we only analyzed BrdU incorporation on the q arms of each chromosome (last three images in each row). The BrdU incorporation in 6q from the intact chromosome 6 was compared with the BrdU incorporation in 6q from the chromosome-10-centromere derivative (6q and 6q*). The BrdU incorporation in 10q from the intact chromosome 10 was compared with the BrdU incorporation in 10q from the chromosome-6-centromere derivative (10q and 10q*). **C**) Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the four q arms from panel **B** are shown. Pixel intensity is on the Y-axis while the X-axis denotes the distance along the chromosome fragment. **D**) Quantification of the BrdU incorporation in six different cells. The blue bars

represent the q arm of the intact chromosome and the red bars represent the q arm of the derivative chromosome. The values represent the total number of pixels (area x intensity) x 1000. A comparison of Chr 6q and t(6delq;10q)(10cen) 6q reveals similar amounts of BrdU incorporation while a comparison of Chr 10q and t(6delq;10q)(6cen) 10q reveals more BrdU incorporation in the derivative chromosome, indicating that the 6-centromere-derivative displays DRT. This is opposite of what is seen in R175 (Figure 1.4) where the 10-centromere-derivative displays DRT. (Figure on next page).

Figure 4.9: Deletion-translocation derivatives display DRT on the 6-centromere derivative.



CHAPTER FIVE

Conclusions

Chapter Two

In this chapter, I show that, like *Xist*, ectopic integration of *ASAR6* into an autosome caused a chromosome-wide delay in DNA replication. Similar to what was observed with *Xist* transgenes, this effect was only observed when *ASAR6* was integrated in multi-copy arrays. ASAR6 also formed an RNA “cloud” and silenced transcription on the integrated autosome, indicating that ASAR6 RNA coated the integrated chromosome in *cis* and resulted in silencing of gene expression. When this transgene lacked the 29 kb region encompassing the promoter and 5´ portion of *ASAR6*, it could no longer induce these changes. This is the first demonstration of an autosomal transcript exhibiting functions that were thought to be associated only with Xist. This indicates that human chromosome 6 harbors an *Xist*-like gene that functions to regulate DNA replication timing and monoallelic expression in *cis*.

One interesting aspect of *ASAR6* and *Xist* is that the loss-of-function and gain-of-function both give the same phenotype. Deletions at the native loci cause DRT and ectopic integrations on different chromosomes also cause DRT [79, 81, 122, 152]. I believe these data suggest that these loci are multi-functional. For example, on the X chromosome, it is known that a small deletion at the *Xist* locus causes chromosome-wide DRT, indicating that the *Xist* locus functions to promote early replication of the X chromosome [79, 81]. However, it is also known that expression of *Xist* causes chromosome-wide silencing and late replication [138]. Interestingly, a deletion at the *Xist* locus on the Xa (the X

chromosome that does not express Xist) causes DRT, suggesting that the Xist transcript is not involved in promoting early transcription [79]. Therefore, it is possible that one function of the *Xist* locus is to produce the Xist transcript, which silences gene expression and delays DNA replication, while another function of this locus is to promote early DNA replication through a transcript-independent mechanism. Thus, the default state of the X chromosome would be to replicate early (what we would call normal replication timing) and this could be dominantly interfered with via *Xist* transcription. This idea would be the same for the *ASAR6* locus. The ASAR6 locus may function in a transcript-dependent manner to silence expression of one allele of random monoallelically-expressed genes on chromosome 6 during the process of random autosomal inactivation (random AI) and promote late-replication of these loci. The other function of this locus would be transcript-independent and promote early replication-timing via an unknown mechanism.

Previous studies have indicated that ASAR6 does not coat chromosome 6 in differentiated lymphocytes or fibroblasts [122, 123]. In this study, it was found that ectopic integration of an *ASAR6* transgene resulted in ASAR6 RNA cloud formation in the nucleus and accumulation of the transcript around the integrated chromosome. These data indicate that ASAR6 RNA has the ability to coat chromosomes, much like the Xist transcript, only it does not do so in at least two differentiated cell types. But why would ASAR6 RNA retain this activity if it does not use it? I believe that ASAR6 RNA does coat chromosome 6 at some point early in development. In the case of X-inactivation, Xist begins to coat the

chromosome around the time that pluripotent cells within the inner cell mass of the blastocyst differentiate, which initiates gene inactivation of ~90% of the genes on the future inactive X chromosome. This process could be the same for chromosome 6, and every autosome for that matter. In this scenario, ASAR6 RNA would coat one allele of chromosome 6 around the time of pluripotent stem cell differentiation, when random AI is being established. This would induce silencing of random monoallelically-expressed genes on one allele of chromosome 6 (~5-10% of genes on chromosome 6). Then, unlike Xist, ASAR6 RNA would no longer associate with the chromosome once autosomal inactivation had been established. Recently, a novel long ncRNA was found to coat the active X chromosome only during ESC differentiation and dissociate shortly after [217]. This study lacked any functional analysis of this transcript but, nevertheless, set the precedent that RNA coating can be a transient, developmentally-regulated event. One piece of evidence that goes against this theory is the fact that random monoallelic gene expression is not coordinated at the whole-chromosome level [163]. Therefore, if one function of *ASAR6* is to silence gene expression in *cis*, it would only be to silence one allele of some of the monoallelically-expressed genes on chromosome 6. The *ASAR6* counterpart on the other chromosome 6 allele would be responsible to silence expression of the remaining monoallelically-expressed genes on chromosome 6.

The demonstration that *ASAR6* has the ability to silence gene transcription in *cis* correlates with previous observations that deletion of *ASAR6* results in the activation of previously silent alleles of linked monoallelically-expressed genes

[122]. Therefore, I suggest that *ASAR6* normally functions to silence gene expression of random monoallelically-expressed genes on chromosome 6. This process would begin early in development, and would result in ASAR6 RNA coating chromosome 6 and initiating silencing of random monoallelically-expressed genes. Following differentiation, once silencing of one allele is established, ASAR6 RNA would no longer associate in *cis* with chromosome 6. So, unlike XCI, where Xist associates with the Xi throughout all stages of development, ASAR6 RNA associates with chromosome 6 at the early stages of development and not in adult tissues. Of course, this is not the only function of the *ASAR6* locus. This locus, through some unknown mechanism, also controls chromosome-wide DNA replication timing by promoting early DNA replication. This function is supported by data indicating that chromosomal deletions and translocations at this locus result in DRT/DMC [77, 122].

Chapter Three

In this chapter, I generated intrachromosomal deletions at four loci known to be involved in translocation-induced DRT/DMC in an attempt to identify other "inactivation/stability" (I/S) centers. Deletions and inversions at a locus on chromosome 15q24 resulted in DRT/DMC, identifying it as a novel I/S center. Gene expression analysis at this locus revealed biallelic expression of a protein-coding gene called *SCAPER*. Yet, despite being biallelically-transcribed from both sense and antisense strands, this locus was asynchronously replicated and

enriched for LINE-1 repeats, similar to the other I/S centers that have been identified to date.

This study has identified a new I/S center on human chromosome 15 and supports my hypothesis that all chromosomes contain at least one of these centers. While it is evident that many deletions in *SCAPER* caused DRT/DMC, it is noteworthy that some did not. There are several explanations for this lack of penetrance: One is that the ploidy imbalance of chromosome 15 is affecting the phenotype in some way. One trend we have noticed is that deletion clones that remain mostly triploid for chromosome 15 tend to display DRT/DMC more often than deletion clones that are mostly tetraploid for chromosome 15. This is an intriguing possibility because deletion of both *Xist* alleles was shown to have a much more dramatic impact on DNA replication timing than deletion of just one allele, indicating there are *trans*-effects that affect the severity of the DRT/DMC phenotype on the X chromosome [79, 81]. Under this scenario, it would be hypothesized that the additional copy of chromosome 15 can rescue the phenotype through a *trans*-acting mechanism. However, it should be noted that this trend with chromosome 15 ploidy and DRT/DMC was not observed in every clone. Furthermore, the existence of genetically identical deletion clones that have different phenotypes does not support this model.

Another possibility is that the generation of DRT/DMC might be via a *cis*-acting epigenetic mechanism. The best evidence for this possibility is that I generated clones that have the exact same deletion in chromosome 15, yet some displayed DRT/DMC and some did not. Assuming these clones are

genetically identical, the only possibility would be that an epigenetic mechanism is somehow differentially regulated in these clones. If this region harbors an epigenetic mark that regulates chromosome-wide replication timing, it is possible that a new I/S center could arise *de novo* elsewhere on the chromosome in some cells following the excision of an existing one. A similar phenomenon has been observed for another *cis*-acting chromosomal element, the centromere. Neocentromere formation has been documented in multiple organisms following excision of a centromere or the generation of an acentric translocation derivative [208]. Since the centromere is essentially an epigenetic structure [209], it appears capable of forming on any DNA sequence [210]. Another cis-acting chromosomal element, the DNA replication origin, has essentially no sequence specificity in higher eukaryotes [7]. Although the replication origin is technically a genetic element, the pre-RC is an epigenetic feature that forms on all origins and is essential for their function. New replication origins can also form upon the deletion of existing ones [7]. So it is formally possible that I/S centers represent an epigenetic feature of every chromosome and that new ones can arise following the loss of an existing one. One piece of evidence against the epigenetic model is that in Chapter Two it was demonstrated that ectopic integration of multiple copies of BAC DNA could dominantly disrupt DNA replication of the integrated autosome. Since BAC DNA contains no proteins, this would argue that the DNA sequence itself is the important functional mediator of chromosome-wide replication timing.

A third possibility is that the instability associated with the chromosome 15

deletions is somehow impacting the DRT/DMC phenotype. The Thayer lab has observed that once the DRT/DMC chromosome has experienced secondary translocations, it often ceases to display DRT/DMC [74, 77]. Therefore, it is possible that some of the clones that do not display DRT/DMC have undergone chromosomal rearrangements that have, in effect, rescued the phenotype. For the DNA replication timing analysis, only chromosome 15s that had not undergone an interchromosomal translocation were analyzed. However, from the DNA-FISH analysis, it was impossible to determine if intrachromosomal events have occurred. Therefore, it is possible that some clones have had their delayed chromosomes rescued by intrachromosomal events. Because the chromosome 15 deletions are much more unstable than the chromosome 6 deletions, this possibility could explain why the deletion analysis on chromosome 6 was not as complicated. This could also explain the differences in phenotype between clones that have the same deletion. Since these clones grew independently of one another, they were able to accumulate different rearrangements that could impact replication delay differently.

The Thayer lab has characterized many common features of the *Xist*, *ASAR6* and *SCAPER* loci: 1) disruption of these loci results in delayed replication timing of entire chromosomes in *cis*, 2) disruption of all three loci results in dramatic instability of the affected chromosome, 3) all three loci have a high LINE-1 content, and 4) each locus displays asynchronous replication that is coordinated with other linked asynchronously replicated loci. Surprisingly, *SCAPER* was biallelically expressed in all cell types examined. However, given

the variability observed with many monoallelic genes [163], it is possible that *SCAPER* is monoallelically expressed in some tissues and biallelically expressed in HTD114 cells and lymphoblastoid cells.

Random monoallelic expression is regulated differently in different cell types [163, 165]. On the other hand, asynchronous DNA replication is a much more consistent chromosomal feature. Therefore, it is possible that asynchronous DNA replication is the more important feature of these loci and monoallelic gene expression is a by-product of differential replication timing. However, it is known that the monoallelically-expressed *Xist* gene, and possibly the *ASAR6* gene, is essential for allelic inactivation of the silent alleles of linked monoallelically-expressed genes. As mentioned in Chapter Two, it is possible that these regions have dual functions: one being transcript-independent and controlling DNA replication timing and chromosome stability and the other one being transcript-dependent and controlling allelic inactivation. It is possible that the SCAPER locus only has the transcript-independent function and does not control allelic inactivation in *cis*.

Chapter Four

The Thayer lab has proposed two models for how ICTs can cause DRT/DMC: 1) the ICT could delete or disrupt a *cis*-acting genetic element that acts to ensure proper DNA replication timing, mitotic chromosome condensation, and chromosome stability (loss-of-function model, Figure 4.1A) or 2) the ICT

could generate a dominant interfering element that disrupts DNA replication timing, mitotic chromosome condensation, and chromosome stability (gain-of-function model, Figure 4.1B) [74, 77]. In this chapter, evidence is provided to support the interfering element (gain-of-function) model. By making both proximal and distal deletions on the same chromosome it was shown that a distal deletion negated the effects of a subsequent proximal deletion. Highlighting the complexity of this process, it was shown that generating a distal deletion followed by the original t(6;10) translocation resulted in DRT on the opposite chromosome derivative (the 6-centromere derivative instead of the 10-centromere derivative). By generating double-deletion and deletion-translocation derivative chromosomes, I have concluded that neither model alone can explain the DRT/DMC phenotype. Although the double-deletion derivative data support the gain-of-function model, the deletion-translocation derivative data does not support either and suggests a more complicated mechanism. Furthermore, because the data presented in Chapter Four does not represent an exhaustive analysis of double-deletion or deletion-translocation clones, these conclusions should be taken as preliminary.

This analysis identified the region distal from the *lox*P site on chromosome 6 as playing a vital role in the acquisition of DRT/DMC. Surprisingly, a distal deletion of only 18 kb in the Δ175F-7a cell line was able to negate the effects of a proximal deletion. Due to the small size of this deletion, the expectation was that this distal deletion would have no effect on subsequent proximal deletions, and the larger distal deletions in Δ175F-12a, and Δ175F-8a cells (150 Kb, and 1.3

Mb) may or may not effect subsequent proximal deletions. Regardless, these data identify an ~18 kb region on chromosome 6 that has a “dominant interfering” function on replication timing. It is currently unclear how this element may be functioning. Therefore, further experiments to study the activity of this region are warranted.

As mentioned in the Chapter Two discussion, I propose the possibility that the chromosome 6 locus has dual functions. One function would be ASAR6-transcript-dependent and promote late replication via ASAR6 coating and gene silencing. The other function of this locus would be transcript-independent and promote early DNA replication via an unknown mechanism. I believe the data presented in this chapter support this model. The double-deletion derivatives are the first example of chromosomes that have disrupted ASAR6 expression and do not display DRT/DMC. Furthermore, the deletion-translocation derivatives that display replication delay (6-centromere derivatives) have an intact, expressed *ASAR6* gene. This suggests that the *ASAR6* RNA is not involved in the acquisition of the DRT/DMC phenotype and that something else within this locus is the main driver.

Importantly, these data do not conflict with the ‘inactivation/stability center’ model that our lab has proposed. However, the genetics of the ‘inactivation/stability center’ are likely to be complicated and, therefore, require a more in-depth analysis.

Overall Conclusions

The data presented in this thesis support the proposal that all mammalian chromosomes contain I/S centers that function to maintain proper replication timing, mitotic chromosome condensation, monoallelic gene expression and stability of individual chromosomes. Under this scenario every mammalian chromosome contains four distinct types of *cis*-acting elements, origins of replication, centromeres, telomeres, and I/S centers, all functioning to ensure proper replication, segregation and stability of individual chromosomes.

With a new locus indentified on chromosome 15, I believe it is only a matter of time before at least one of these regions has been identified on every human chromosome. Furthermore, it is known that mouse and hamster chromosomes can also display DRT/DMC [80, 82]. Therefore, it is possible that all mammalian chromosomes, and perhaps all eukaryotic chromosomes, contain similarly functioning regions. It appears that the reason each chromosome contains an I/S center is to ensure that chromosome replication is synchronized during S phase, because when one of these regions is disrupted, that chromosome no longer replicates in synch with its homolog. This would indicate that each chromosome regulates its replication independently of one another, not just at the origin level, but also at the whole chromosome level. This is in contrast to the classical model of chromosomal replication being dictated solely by origin licensing and activation by S-phase cyclin/CDK complexes [3]. These data indicate that this process is more complex and involves multiple facets of

chromosomal function. Whether the underlying mechanism of these I/S centers is epigenetic or genetic in nature still remains an unanswered question. Like most processes, however, it is likely to involve both genetic and epigenetic regulation.

The data provided in this thesis reveal a complex mechanism that governs chromosome-wide DNA replication timing. The two models that the Thayer lab has proposed previously can no longer explain all of the data. Therefore, I propose a new model that can explain nearly every one of our observations (Figure 5.1). This model assumes that the *lox*P locus on chromosome 6 contains two functional domains: the 29 kb region that spans the *ASAR6* promoter (region A) and the 18 kb region distal from the *lox*P site (region B). Region A has dual functions: one, mediated by ASAR6 RNA, is to promote the late replication and gene inactivation of individual loci that are subject to random AI (function 1) and the other is to promote chromosome-wide early replication (function 2). Region B functions to inhibit early DNA replication along the entire chromosome. These functions have a hierarchy, whereby function 1 of region A is dominant to function 2, and function 2 of region A is dominant to the function of region B.

Normally, the inhibitory function of region B is suppressed by function 2 of region A and random AI is initiated and maintained by function 1 of region A. Upon deletion of region A or separation of the two regions by a chromosomal translocation, region B inhibits early replication of the derivative chromosome and causes DRT. Upon deletion of region B, no replication timing phenotype is displayed. Ectopic integration of a single-copy BAC that spans both of these regions does not cause any replication timing defects; however, a multi-copy

integration can cause chromosome-wide gene silencing, ASAR6 RNA coating and DRT. This is due to the dominant activity of function 1 of region A. Normally this function is tightly regulated to only allow late replication and gene inactivation of specific loci in *cis*, however, when this region is amplified by multi-copy integration, a threshold is reached thereby deregulating function 1 of region A and promoting late-replication and gene silencing of the entire chromosome.

The strength of this model is that it can explain why the gain-of-function and the loss-of-function of this region result in similar, but distinct, phenotypes. As mentioned before, the gain or loss of the *ASAR6* locus results in chromosome-wide delayed replication. However, the gain of the *ASAR6* locus results in chromosome-wide gene silencing while the loss of *ASAR6* results in the activation of some genes [122]. Accordingly, this model proposes that the gain or loss of the ASAR6 locus results in chromosome-wide replication delay via two completely different mechanisms: 1) The loss of ASAR6 results in the activation of region B, which inhibits early DNA replication. The loss of ASAR6 also results in the activation of silent alleles of random monoallelic genes. 2) The gain of ASAR6 results in the deregulation of region A resulting in late replication and gene inactivation of the entire chromosome (Figure 5.1). Although these two regions are well mapped in the ASAR6 locus, it is too early to tell exactly where they exist within the chromosome 15 locus.

The severity of the genomic instability observed on chromosome 15 deletions and inversions underscores just how essential these regions are to normal chromosomal function. Where *P53* has been called “the guardian of the

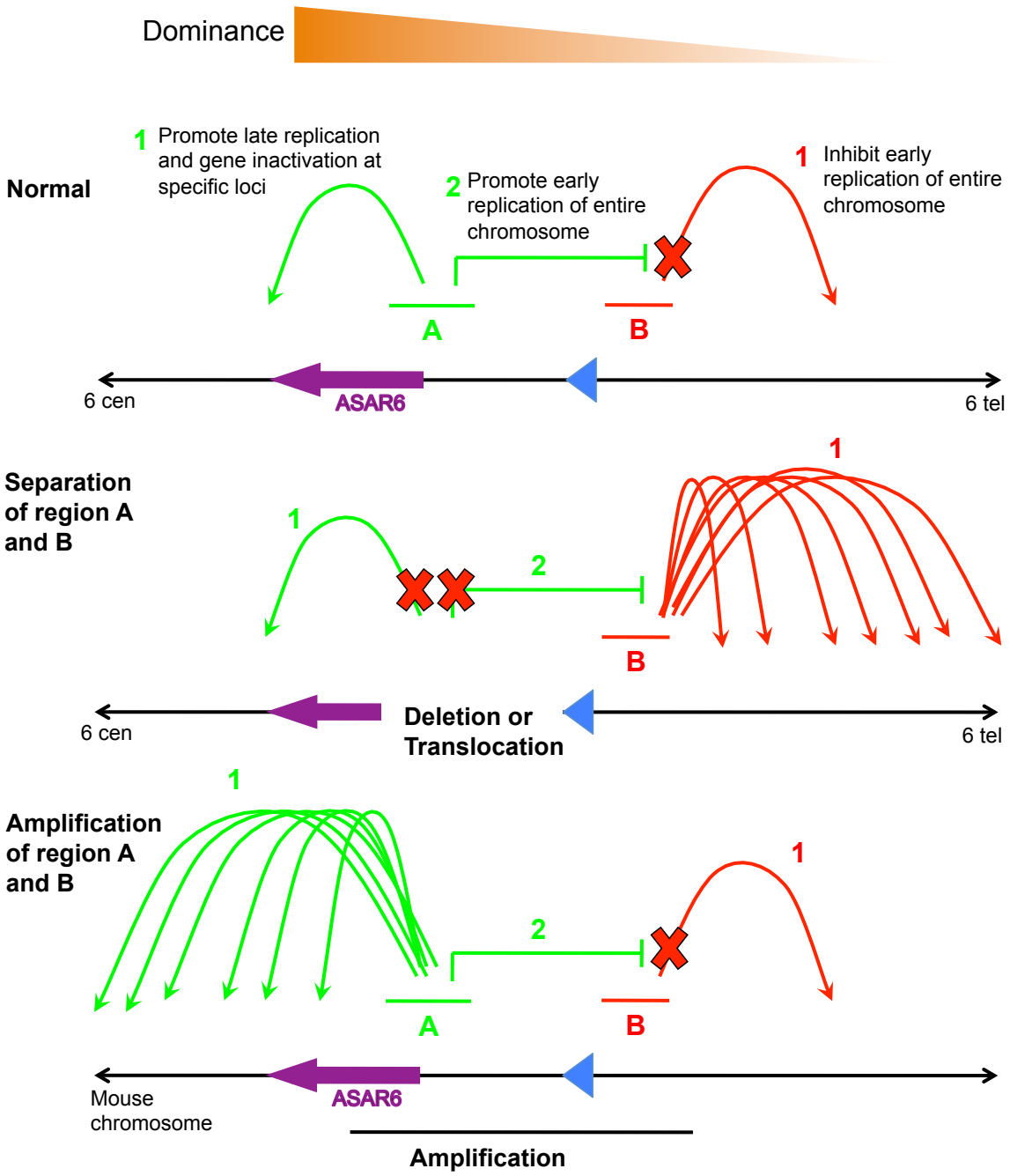
genome," I/S centers appear to be "the guardians of individual chromosomes." The chromosome 15 analysis supports the observation that DRT/DMC leads to chromosome number imbalances and chromosome structure instability [75, 80]. Unlike other mechanisms that cause genomic instability, DRT/DMC tends to be a transient phenomenon. The inherent instability of DRT/DMC chromosomes makes them prone to extreme fragmentation over a limited number of cell divisions, which eventually results in highly rearranged chromosomes that no longer display DRT/DMC [74]. This feature of DRT/DMC makes it an underappreciated, yet potentially important force driving mutagenesis.

Historically, cancer has been thought to develop according to a stepwise equilibrium, where mutations accumulate gradually over time. However, recent karyotypic analysis and whole-genome sequencing of tumor samples has indicated that the stepwise equilibrium of tumorigenesis may be punctuated by short events of increased mutagenesis [218]. This is called the punctuated equilibrium model and it has been used to explain the existence of catastrophic, single events that have impacted the genome in some cancers [115, 116]. I believe that DRT/DMC is one such process that is responsible for punctuated mutagenesis in cancer. It is known that DRT/DMC chromosomes are present in primary tumor cells [74] and that DRT/DMC can produce an abundance of mutational events in a very short period of time. I, therefore, propose that altered replication timing and chromosome condensation plays an important role in the acquisition of genetic aberrations in the development of human cancer.

Figures

Figure 5.1: Model for *cis* control of chromosome-wide replication timing and random monoallelic gene expression. In normal cells, function 2 of region A suppresses the inhibitory function of region B, thereby resulting in the early replication of one allele of chromosome 6. Function 1 of region A dominantly disrupts the early replication program only at loci that are subject to random AI, resulting in the gene inactivation and late replication at those specific loci. Following the separation of region A from region B, the function of region B inhibits early replication of one allele of chromosome 6, resulting in the late replication of the entire chromosome. The loss of region A also results in the reactivation of previously silent alleles of genes subject to random AI. Following the amplification of both regions, function 1 of region A becomes deregulated and promotes late replication and gene inactivation of the entire chromosome instead of just at specific loci. (Figure on next page).

Figure 5.1: Model for *cis* control of chromosome-wide replication timing and random monoallelic gene expression.



CHAPTER SIX

Future Directions

Genetic Analysis of the *ASAR6* Locus

Ectopic Integration

One contribution of this work is the development of a gain-of-function assay to detect DRT/DMC. Ectopic integrations of *Xist* transgenes have been vital in identifying the many different domains that give this gene its unique function [152]. I believe that this ectopic integration assay will be just as integral in identifying the different functional domains within I/S centers. The *ASAR6* locus has been mapped via random intrachromosomal deletions, but there is only so much information that can be learned from generating random-sized deletions. The Thayer lab has been successful in using a recombinant adeno-associated virus (rAAV) to make targeted deletions [123], however, this method is very inefficient and deletions still have to be anchored at the original *lox*P site, limiting the range of genetic manipulations that can be achieved. A gain-of-function assay, like ectopic integration, will allow us to make specific genetic manipulations to the DNA that we are integrating—something that is currently not possible with our loss-of-function assay.

Another contribution of this work is the identification of two separate functional domains at the *ASAR6* locus on chromosome 6. The first was identified initially by the deletion analysis and confirmed using the ectopic integration assay (the ~29kb region around the *ASAR6* promoter). The second is the ~18 kb region identified in chapter 4 that, upon deletion, negated the effects of *ASAR6* deletion on DNA replication timing (Figure 6.1). The RP11-767E7 BAC

that was used in the ectopic integration assay happened to contain both of these regions. The ectopic integration assay would be a great way to test the function of different transgenes containing various combinations of the genetic material at this locus. I have successfully used recombineering to insert a selectable marker in two other BACs and three other fosmids (Figure 6.1). These BACs and fosmids have been transfected into C2C12 cells and clones have been isolated that have ectopic integrations. These clones have not yet been analyzed for DRT.

Rescue of DRT/DMC

Although ectopic integration has its benefits, it also has significant downsides. The phenotype of transgene integrations can be prone to position and copy-number effects that are not fully understood at this point. Therefore an alternate genetic analysis is warranted. One experiment that has not been performed yet is a genetic rescue experiment. This is the gold standard for any genetic analysis but, due to the nature of these studies, it is a difficult experiment to perform. Our lab has developed two strategies to rescue the DRT/DMC phenotype. The first is a straightforward rescue. It involves transfecting the RP11-767E7+hyg BAC into cells that contain an ~76 kb proximal deletion in chromosome 6 (Δ175-23a cells [122]) along with Cre recombinase (Figure 6.2). This would answer an important question, which is: once DRT/DMC is established, can it be reversed?

The problem with the above strategy is that once the DRT/DMC phenotype is established, the chromosome becomes very unstable. So, by the time we get around to integrating the BAC for rescue, the chromosome may have already undergone multiple rearrangements. To get around this problem, we have come up with a second strategy that involves duplicating this region before making the deletion (Figure 6.3). This would get around the instability effects of DRT/DMC and also answer the question: does a duplication of this region result in any replication timing defects? I have attempted the first strategy multiple times and have not been able to isolate clones that have the targeted BAC integration. My guess is that getting a BAC and Cre into the same cell and having recombination occur will be an extremely rare event, so it may just take many more times of trying until targeted BAC integration is possible. For the second strategy, I have successfully made the integration plasmid (rescue backbone e3.6) that needs to be integrated into the BAC via recombineering. Once the rescue backbone is integrated into the BAC to make the BAC integration plasmid then the strategy can be carried out according to Figure 6.3.

Inactivation/Stability Center Regulation During Development

One conclusion of this work is that ASAR6 RNA has chromosomal coating function and the ability to silence gene expression. However, ASAR6 RNA localization in differentiated cells does not reveal any evidence that ASAR6 RNA coats chromosome 6. It is my belief that ASAR6 RNA does, in fact, utilize its chromosomal coating and silencing function early in development. One way to confirm this hypothesis is to analyze ASAR6 RNA expression in human ESCs pre- and post-differentiation. Our expectation is that ASAR6 RNA would either be biallelically expressed or not expressed in ESCs. Expression would then become monoallelic and ASAR6 RNA would begin to cover chromosome 6 upon differentiation. Eventually the cloud would dissipate and expression would be similar to what we have observed in our differentiated cells. Since I believe this process to take place at the same time as X inactivation, using human ESC lines may be problematic for our analysis because the XCI status is questionable in many of these lines. Some culturing conditions have been established that keep cells in an X-inactivation-free state [219, 220], however, the random AI state of these cells is unknown (i.e. are random monoallelically-expressed genes that exist on autosomes biallelically expressed or not).

The difficulty surrounding these studies in human would not be present in mouse. True mouse ESCs can be easily isolated from developing embryos and analyzed. Unfortunately the mouse *Asar4* locus (the *ASAR6* locus exists on mouse chromosome 4) has not been characterized yet and it is unclear whether

deletions at the mouse locus would phenocopy those at the human locus. Looking to the future, the mouse *Asar4* locus needs to be characterized because it would make the genetic analysis so much easier and would allow us to perform experiments that are extremely hard to do in human cells. One assay to see if the mouse locus functions in a similar manner to the human locus is to ectopically integrate a mouse *Asar4* transgene into an autosome and see if it disrupts DNA replication timing. I have used recombineering to put a hyg cassette into a mouse BAC (RP23-117E19) that corresponds to the human *ASAR6* locus and this BAC just needs to be put into cells and the integrated chromosomes analyzed for DRT. I have also helped our collaborators make targeting constructs to generate transgenic mice that have a deletion within the mouse *Asar4* locus.

Instability Signature of DRT/DMC

The instability observed following DRT/DMC has a very unique cytogenetic signature, with most of the chromosomal rearrangements affecting the delayed chromosome [77]. This single-chromosome instability is reminiscent of two newly described instability signatures, “chromothripsis” and “kataegis,” which are present in some cancers [115, 116]. Chromothripsis and kataegis appear to be cataclysmic events in which a chromosome, chromosome arm or local region on a chromosome is fragmented or heavily mutated in a relatively short period of time. This clustering of mutational events occurs in *cis* and results in entire chromosomes or local regions that have undergone extreme fragmentation and mutagenesis [115, 116]. In the case of chromothripsis, the sequences at the rearrangement junctions show either a lack of homology or microhomology between the joined segments, suggesting that the DNA was repaired by non-homologous end joining (NHEJ) [115]. In addition, the complex chromosome rearrangements associated with genomic disorders in humans were recently found to resemble chromothripsis [117, 118].

Sequencing the breakpoints at these complex rearrangements identified characteristic features, including small templated insertions of nearby sequences and microhomologies, suggestive of replicative processes. These observations led the Lupski group to propose the term ‘chromoanasynthesis’ as an alternative descriptor to chromothripsis for the shattering and reassembly of chromosomes via replicative mechanisms [117]. The Lupski group proposed a microhomology

mediated break induced replication (MMBIR) and a related fork stalling and template switching (FoSTeS) model for the origin of these complex rearrangements [119]. The distinction between MMBIR/FoSTeS and NHEJ is that the microhomology junctions in MMBIR/FoSTeS are followed by stretches of DNA sequence derived from elsewhere, usually nearby. The MMBIR/FoSTeS models involve stalled DNA replication forks that are resolved by replication restart using short stretches of homology [119]. Furthermore, stalled DNA replication forks can also be resolved into DSBs providing a substrate that can be repaired by NHEJ [120].

Thus, if multiple stalled replication forks form on a single chromosome and are resolved via MMBIR, FoSTeS, or strand breakage followed by NHEJ, it would leave the chromothripsis and kataegis instability signatures. One possibility is that DRT/DMC is responsible for the formation of multiple stalled replication forks on a single chromosome. As mentioned in Chapter One, there are two models for how DRT/DMC can induce CSIN. One is that DRT/DMC results in checkpoint adaptation and the onset of mitotic chromosome condensation prior to the completion of DNA synthesis leading to stalled replication forks [85]. This would generate multiple stalled replication forks on a single chromosome and lead to multiple rearrangements generated at the stalled replication forks via NHEJ, MMBIR, and/or FoSTeS type mechanisms. The other way DRT/DMC can induce CSIN is by delaying mitotic spindle attachment, which can lead to chromosome mis-segregation and the formation of micronuclei [85]. Interestingly, it has been found that inducing micronuclei by nocodazole treatment can lead to extreme

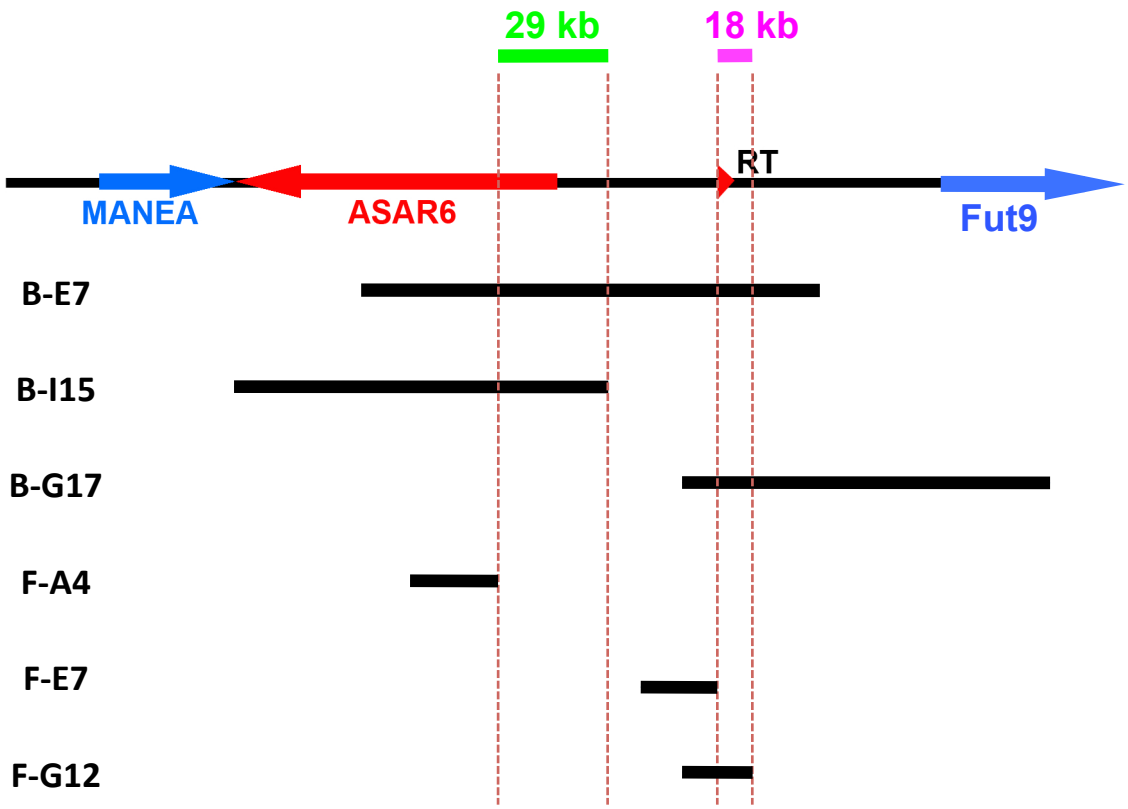
fragmentation of single chromosomes [84]. Therefore, DRT/DMC is not only present in some cancer cells but it can also explain certain unique instability signatures that are present in some cancer cells.

The data presented here confirms our belief that DRT/DMC is one possible mechanism responsible for chromothripsis and/or kataegis. The instability associated with the chromosome 15 deletions is even more severe than that of the chromosome 6 deletions. Cytogenetically, all of our data suggest that DRT/DMC causes chromosome shattering. However, since chromothripsis and kataegis were discovered and characterized by their molecular signatures, we need to do the same analysis on our cells. To detect the intrachromosomal events that are hard to see by DNA-FISH, the Thayer lab hopes to perform whole-genome sequencing on some of the chromosome 15 deletion clones to identify whether the breakpoints cluster on a single chromosome and what the breakpoint junctions look like. Once we have a better idea of the molecular impact of DRT/DMC, we should be able to confirm whether these cells have a chromothripsis-like signature or whether DRT/DMC results in another instability signature.

Figures

Figure 6.1: Positions of recently generated transgenes. The *ASAR6* locus is depicted with the location and orientation of the *lox*P-3´RT integration and three genes, *MANEA, ASAR6 and FUT9*. The 29 kb region identified in Chapter Two (green) and the 18 kb region identified in Chapter Four (purple) are indicated above. The RP11-767E7 BAC (B-E7) that was used in the experiments in Chapter Two spans both regions. BACs and fosmids that contain a Hyg resistance gene are listed below: RP11-374I15 (B-I15), RP11-48G17 (B-G17) G248P83419A4 (F-A4), G248P86031E7 (F-E7) and G248P86150G12 (F-G12). B-E7 and B-I15 contain the entire 29 kb region and B-E7, B-G17 and F-G12 contain the entire 18 kb region. All have been integrated into C2C12 cells and are awaiting analysis. (Figure on next page).

Figure 6.1: Positions of recently generated transgenes.



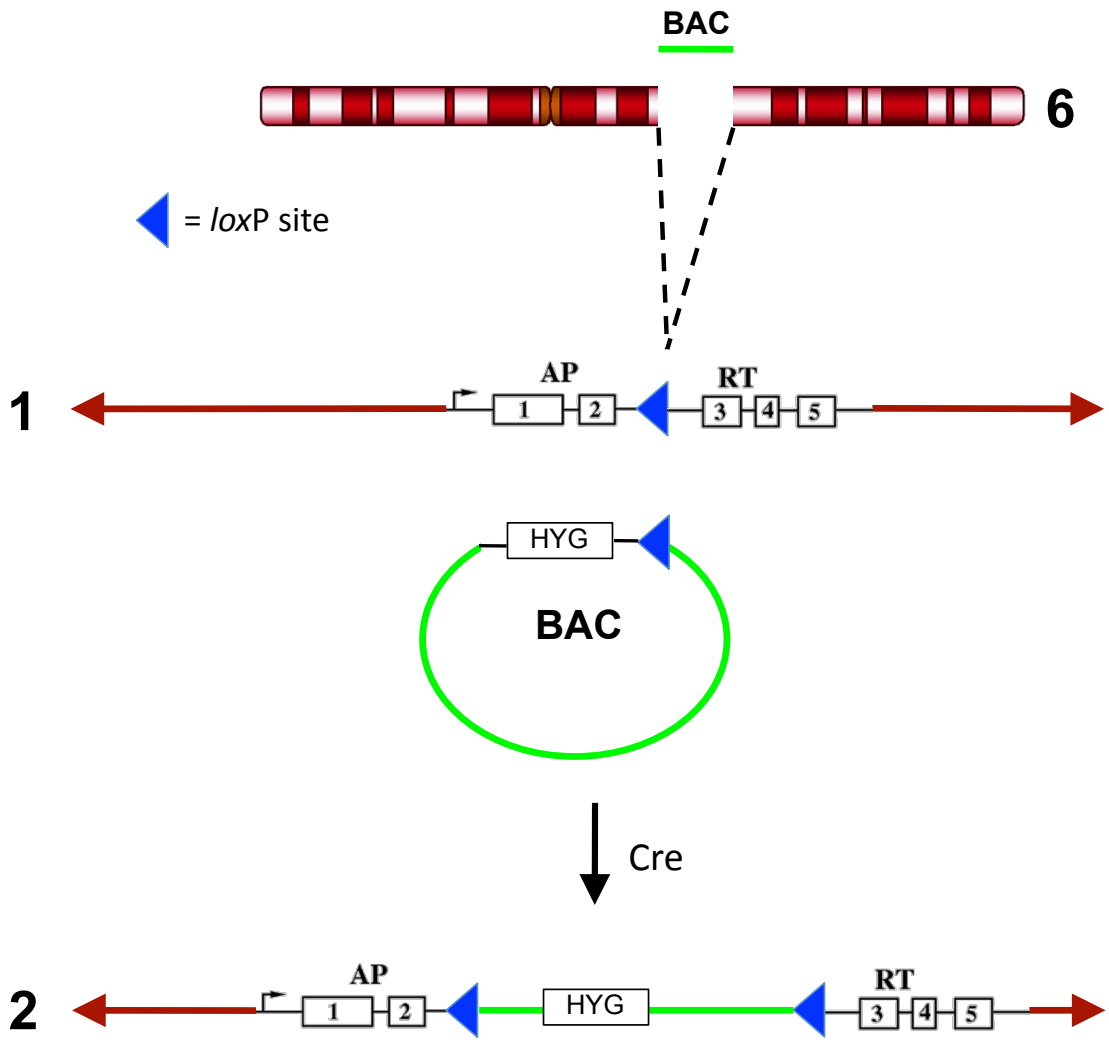
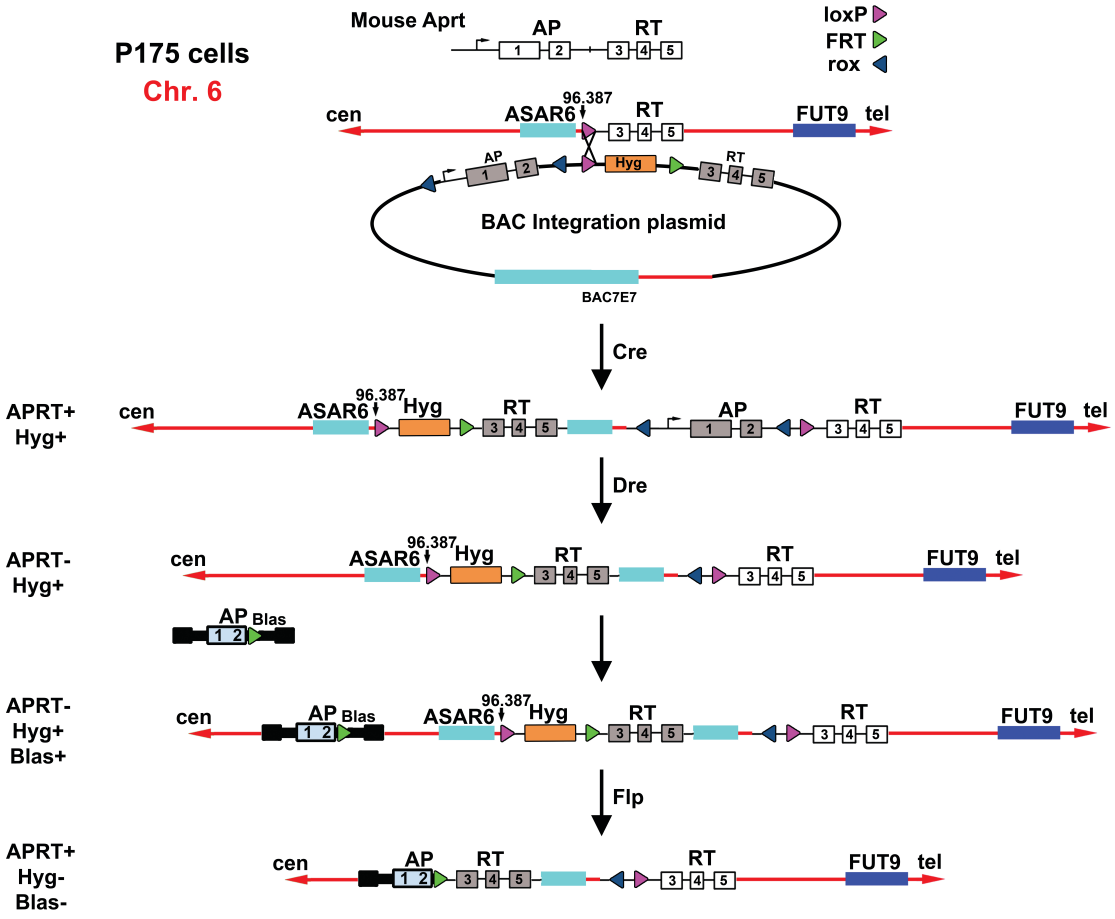


Figure 6.2: Rescue strategy number one. Strategy to integrate BAC into deleted locus. **1**) ∆175-23a cells that have a 76 kb deletion in chromosome 6 and display DRT contain an intact mouse APRT transgene with a *lox*P site (blue triangle) in the middle of it at the deleted locus [122]. A BAC that spans the deleted locus (RP11-767E7+hyg, green) is transfected into cells along with Cre recombinase. Cre mediates recombination between the *lox*P site on the BAC and the *lox*P site in chromosome 6. Cells are cultured in media containing Hygromycin B and DAP to select Hyg resistant APRT- cells.

Figure 6.3: Rescue strategy number two. First, an integration plasmid (rescue backbone e3.6) is inserted into the BAC RP11-767E7 (BAC7E7) to generate the BAC integration plasmid. The BAC integration plasmid contains a Hygromycin resistance gene, a 3′RT cassette, and the 5′AP cassette and will integrate at the original *lox*P-3′RT integration site on chromosome 6 in P175 cells via Cre-mediated recombination. The resulting cells will be both APRT+ and Hyg+. The integration plasmid also contains *rox* (*r*egion of x-over) sites that flank the 5′AP cassette. These *rox* sites are 32 bp recognition sequences for the site-specific recombinase Dre (D6 recombinase) [221, 222]. Following transfection with Dre, the 5′AP cassette gets excised, disrupting the APRT gene. APRT- cells are infected with a lentivirus containing the 5′AP cassette and an FRT site. Blast-positive cells are transfected with FLP recombinase. Proximal deletions on chromosome 6 result in reconstitution of the APRT selectable marker as well as Blast and Hyg sensitivity. These deletions excise various lengths of DNA on chromosome 6, proximal to the original *lox*P-3′RT integration site, and will include part of or all of the smallest deletion (~76 kb) that displays DRT/DMC. However, chromosome 6 will still contain slightly more than that smallest ~76 kb region of DNA from the integrated BAC7E7. In this way, the proximal deletions that result in DRT/DMC will be remade, but the minimal deleted region known to cause DRT/DMC will be provided on the BAC to see whether it is capable of rescuing the phenotype. Figure adapted from [216]. (Figure on next page).

Figure 6.3: Rescue strategy number two.



APPENDIX

DNA Replication Timing, Genome Stability and Cancer

Nathan Donley and Mathew J. Thayer

Department of Biochemistry and Molecular Biology

Oregon Health & Science University

Knight Cancer Institute

3181 S. W. Sam Jackson Park Road

Portland, Oregon 97239, USA

TEL: (503) 494-2447

FAX: (503) 494-7368

Email:thayerm@ohsu.edu

Subtitle: Late and/or delayed DNA replication timing is associated with increased genomic instability.

Semin Cancer Biol. (2013) **23** (2): p. 80-9

Abstract

Normal cellular division requires that the genome be faithfully replicated to ensure that unaltered genomic information is passed from one generation to the next. DNA replication initiates from thousands of origins scattered throughout the genome every cell cycle; however, not all origins initiate replication at the same time. A vast amount of work over the years indicates that different origins along each eukaryotic chromosome are activated in early, middle or late S phase. This temporal control of DNA replication is referred to as the replication-timing program. The replication-timing program represents a very stable epigenetic feature of chromosomes. Recent evidence has indicated that the replication-timing program can influence the spatial distribution of mutagenic events such that certain regions of the genome experience increased spontaneous mutagenesis compared to surrounding regions. This influence has helped shape the genomes of humans and other multicellular organisms and can affect the distribution of mutations in somatic cells. It is also becoming clear that the replication-timing program is deregulated in many disease states, including cancer. Aberrant DNA replication timing is associated with changes in gene expression, changes in epigenetic modifications and an increased frequency of structural rearrangements. Furthermore, certain replication timing changes can directly lead to overt genomic instability and may explain unique mutational signatures that are present in cells that have undergone the recently described processes of “chromothripsis” and “kataegis”. In this review, we will discuss how

the normal replication timing program, as well as how alterations to this program, can contribute to the evolution of the genomic landscape in normal and cancerous cells.

Introduction

In order to divide, a eukaryotic cell must undergo precise DNA replication to ensure that an exact copy of its genetic content is passed on to its daughter cells. This process occurs during S phase and proceeds via the coordinated initiation of DNA replication at hundreds of replication origins scattered throughout the length of each chromosome [223]. Interestingly, the cell begins preparation for DNA synthesis in telophase of the prior cell cycle [5]. This is when the pre-replicative complex (pre-RC) begins to form on each potential origin of replication. However, not all pre-RCs will go on to become active replication origins. In mid-G1, at the origin decision point (ODP), some pre-RCs are chosen to become initiators of DNA replication while others remain inactive throughout S-phase [6, 7]. The addition of other replication factors to a subset of the pre-RCs transforms them into pre-initiation complexes (pre-ICs) [224]. Shortly after the pre-IC is formed, DNA polymerase and primase are recruited to each origin and DNA synthesis begins in a bidirectional manner. DNA replication proceeds from each origin until the replication forks from two neighboring origins meet and the nascent DNA strands are ligated [3].

While DNA replication can initiate from any active origin within a given S-phase, the timing at which initiation takes place can vary widely between origins. Adjacent origins tend to initiate DNA replication at the same time resulting in large, synchronously replicating chromosomal domains called “replicon clusters” [9, 10]. Some replicon clusters begin replication at the onset of S-phase while

others begin later during the middle or near the end of S-phase. This coordination of the temporal control of DNA replication is referred to as the replication-timing program. The replication-timing program is established shortly after mitosis at a point in the G1 phase, preceding the ODP, called the timing decision point (TDP) [11, 12]. The TDP is established coincidently with a global reorganization of chromatin into specified regions within the nucleus [12].

The replication-timing program is mitotically stable, heritable and subject to differential regulation during differentiation and development, making it a robust epigenetic feature of all eukaryotic chromosomes [31]. The biological significance of this replication-timing program is currently unknown; however, the existence of aberrant replication timing in many different genetic diseases suggests that it is a vital cellular process [25-28, 225]. Not surprisingly, the 3-dimensional chromosome architecture in the nucleus is highly coordinated with DNA replication timing. In most, if not all, eukaryotic organisms, early-replicating DNA resides in the interior of the nucleus while the later-replicating regions remain at the nuclear periphery or near the nucleolus [12-14]. Molecular analysis has also revealed that late-replicating regions tend to cluster with other late-replicating regions in the nucleus and vice-versa [15]. Additional complex associations have been observed with genome sequence, structure and replication timing. For example, early-replicating regions tend to positively correlate with gene expression, G+C rich sequences, light-staining Giemsa bands, and active chromatin marks, while late-replicating regions tend to be gene-poor, A+T rich, and have repressive chromatin marks [13, 16, 17]. It should

be pointed out that while these correlations are significant they are not absolute, as some expressed genes with transcriptional active chromatin marks reside in late-replicating regions [10].

DNA synthesis occurs in replication factories within the 3-dimensional space of the nucleus. In these factories, regions of similar replication timing cluster together in the nucleus, with early-replicating regions residing in the nuclear interior and late-replicating regions remaining at the nuclear periphery or near the nucleolus [11, 13, 15, 190]. Additionally, replication-timing changes that occur during development are accompanied by changes in nuclear architecture, indicating that these two features are very closely linked [226]. Therefore, regions that replicate at comparable times in S phase tend to have a closer spatial association than regions that replicate at different times. This association has been highlighted using the HiC method, which probes the three-dimensional architecture of whole genomes by coupling proximity-based ligation with massively parallel sequencing [15, 227].

One prominent disease that is characterized by replication-timing aberrations is cancer. Cancer develops when normal cells acquire genetic and epigenetic alterations that lead to uncontrolled growth and the ability to evade cell death. These genetic and epigenetic alterations are generally thought to drive carcinogenesis by deregulating key pathways that control cell growth and proliferation [86]. Genetic alterations can arise during cancer progression through normal cellular processes, induced or spontaneous mutagenesis, or as a result of genomic instability. Mutagenesis refers to the process by which genetic changes

occur, either spontaneously or as a consequence of exposure to mutagens, resulting in a change in the DNA sequence. Genomic instability, on the other hand, refers to an increase in the rate of mutagenesis per unit time. While normal cells have a very low intrinsic mutation rate, any mechanism that increases the mutation rate can be said to cause genomic instability. Current models suggest that an underlying genomic instability is responsible for the rapid accumulation of the genetic and epigenetic changes that affect gene function in cancer [86]. Therefore, it is very difficult to understand cancer development without understanding the mechanisms that cause genomic instability.

In this review, we highlight research suggesting that the normal DNA replication-timing program has a profound impact on the distribution of mutations that arise during the evolution of species as well as during the evolution of cancer. Aberrant DNA replication timing is associated with altered gene expression, mutagenesis and genomic instability. Furthermore, we propose that certain DNA replication-timing aberrations can explain the newly described processes of “chromothripsis” and “kataegis”, which have been found to generate unique genomic signatures in the genomes of some tumor cells [115, 116].

DNA Replication Timing and the Evolution of the Genomic Landscape

The conventional view of evolution assumes that DNA mutations occur randomly throughout the genome and the eventual presence or absence of those DNA changes in the population is determined through the process of natural selection. While natural selection remains the most potent force shaping the evolution of the genomic landscape, the notion that DNA mutations occur randomly in the genome has become outdated. We now know that mutation rate varies widely throughout the genome and is influenced by many local genetic and epigenetic features such as recombination rate, CpG content, transcriptional status, repetitive-sequence content and chromatin conformation [228-230]. Although it was observed more than 20 years ago [231], a wealth of recent experimental data has confirmed that DNA replication timing is also a potent force that influences mutation rates.

An elegant series of experiments in yeast established that late-replicating regions of the genome have higher rates of spontaneous mutagenesis than early-replicating regions. By inserting an exogenous sequence into different regions of a chromosome and calculating the rate of mutations occurring in that sequence, Lang et al. demonstrated that there was a strong positive correlation between the time of replication and the rate of mutation [232]. Furthermore, by deleting an early-replicating origin, and consequently delaying replication, near one of these exogenous sequences they observed a slight increase in the rate of

mutagenic events. This indicates that delaying the initiation of DNA replication at a particular sequence is sufficient to increase its mutation rate [232].

Other experiments have demonstrated that endogenous loci from many different organisms show a similar correlation of mutagenesis and replication timing. Regions of single-nucleotide diversity in mice and humans are enriched in late-replicating regions [27, 233-235]. When comparing the human genome to multiple non-human primate genomes it was also observed that areas of single-nucleotide divergence between species disproportionately lie in late-replicating regions [233, 234]. A parallel correlation between divergence and late replication was also seen when comparing the genomes of mice and rats [234]. Similarly, regions that have a high density of duplications tend to be late replicating in flies [236]. And duplication hotspots that are shared between different species of flies also reside preferentially in regions of late replication [237]. Genomic domains prone to duplication events are also hotspots for neutral point mutations [238, 239], further supporting the idea that mutational events occur in spatial proximity to one another (i.e. in late-replicating genomic regions). Interestingly, a correlation between early replication timing and deletion variation in flies was also observed, and it will be interesting to determine if this holds true for other eukaryotes as well [236]. These data indicate that timing of DNA replication may be a driving force in copy number and single-nucleotide polymorphism (SNP) diversity that is observed within a species and between species. There have been several suggestions for why replication timing and neutral mutation rate correlate so closely. Most involve differential repair mechanisms being used at

different times during S phase such that error-prone DNA repair pathways are utilized more frequently during late S phase [232, 234].

Late-replicating regions aren't the only replication-associated sites of genomic change. There is also evidence that replication transition regions are also hotspots for spontaneous mutagenesis. Replication transition regions are areas that lie between early-replicating DNA and late-replicating DNA and, therefore, often replicate in mid to late S phase. Transition regions are void of origins and are passively replicated by a uni-directional fork that initiates at an adjacent early origin. This single fork replicates the entire transition region until it reaches the replication fork of an adjacent late origin [226]. One consequence of such a large replicon is an increase in the probability of replication fork stalling and DNA damage [226, 240]. Indeed, a survey of SNPs on human chromosomes 11 and 21 indicated that there is a higher density of SNPs in replication transition regions than there are in early-replicating regions [27]. In addition, frequent gene amplifications on these same chromosomes also lie within these replication transition regions [241]. The frequent gene amplification that is seen in late-replicating and transition regions may have functional implications during evolution, as gene duplications are considered to be an important factor during speciation [242]. Additionally, syntenic breakpoints in the mouse and human genomes appear to occur predominantly in transition regions [241], indicating that these regions may be sources of breakage during the generation of new chromosomes during evolution. Of course more comprehensive studies will be

needed to confirm these findings, but these observations nevertheless implicate DNA replication timing as a potent regulator of mutational dynamics.

Now that DNA replication timing is increasingly implicated in establishing a gradient of mutagenesis such that late replicating and transition regions have a higher rate of mutation than early replicating regions, it is important to understand why a replication-timing program exists at all. It is well established that most regions of active gene transcription are early replicating, whereas silenced genes, intergenic regions and repetitive sequences are late replicating [189]. Most silenced genes (late replicating) tend to be tissue-specific and only become expressed (early replicating) in the tissue where they function [16, 243]. Incidentally, many tissue-specific genes, e.g. receptors involved in sensing environmental changes and during the immune response, are much more divergent between species than genes involved in essential cell functions like metabolism and transcription [238, 244]. It is tempting to speculate that replication timing might be a way to optimize the intrinsic mutation rate such that housekeeping genes incur fewer mutations while tissue-specific genes, that are generally under greater pressure to adapt, receive an increased mutational load. Interestingly, it was found that different classes of genes tend to reside in regions with differential substitution rates between mouse and human [238]. Genes with “receptor type” functions (cell adhesion, immune function, olfactory receptors) generally reside in regions of high mutation density while genes involved in RNA binding, kinase activity and metabolism reside in regions of low mutation density

[238]. As we get deeper into the age of genomics it will be interesting to see what other trends emerge with respect to the evolution of the genomic landscape.

DNA Replication Timing and the Evolution of the Cancer Genome

The observations described above indicate that replication timing influences the mutation rate of different genomic regions in the germline, and over long periods of time differences in replication timing can contribute to the genetic variation within and between species. However, there is also increasing evidence that replication timing influences the mutation rate in somatic cells and may be a contributing factor to the distribution of genomic changes that arise during cancer development.

The Role of Normal DNA Replication Timing During Cancer Mutagenesis

An extensive sequence-based analysis of many different human tumors has revealed an increase in single-nucleotide variations (SNVs) and somatic copy number alterations (SCNAs) in late replicating regions of the genome [190, 245]. Interestingly, in these tumor samples, genomic deletions are enriched in late replicating regions whereas amplifications are enriched in early replicating regions, which is opposite of what is seen in the germline [190]. Other groups have found that genomic rearrangements in cancer correlate differently with replication timing depending on the type of tumor studied. For example, rearrangement breakpoints in breast cancer and neuroblastoma tend to lie within early-replicating regions, whereas breakpoints in colorectal cancer and melanoma tend to reside in late replicating regions [246, 247]. This variation

highlights the epigenetic heterogeneity of different tumor types and may be due to differential selective constraints imposed during tumor evolution or differential deregulation of DNA repair pathways.

Similar to what is seen in the germline, transition regions seem to be hotspots of copy number alterations in human cancers [190, 241]. As mentioned above, the complicated nature of DNA replication in these regions makes them especially susceptible to fork stalling and DNA damage [240]. Accordingly, many fusion genes and recurrent chromosome aberrations found in cancer lie within or near transition regions [27, 248]. Not only are these regions prone to mutagenesis in cancer, but also harbor a higher proportion of genes with oncogenic and tumor-suppressing functions [27, 190, 241, 248]. It has been proposed that the spatial proximity between regions of similar replication timing can influence translocation and rearrangement sites in the genome [190, 249]. This is seen in many cancer cells where regions that cluster next to one another in the nucleus are more likely to undergo translocation events than more distant regions [190, 250-252]. Not surprisingly, many recurrent and oncogenic translocations occur between regions of similar replication timing and nuclear proximity [250, 253, 254].

An additional unstable feature of most if not all mammalian chromosomes is the presence of chromosome fragile sites (CFSs) [255, 256]. Common CFSs are discrete regions of chromosomes that are prone to breakage during times of replication stress [257]. CSFs have been found to lie at the interface of R and G chromosome bands [258, 259], which is a hallmark of replication transition

regions [240]. This indicates that transition regions and CFSs may represent the same genomic feature. Accordingly, CFSs are common points of chromosomal breakage in tumors and CFS instability is often seen in the early stages of carcinogenesis [260]. Since some cancer-related genes lie within fragile sites, CFS instability can directly deregulate some oncogene/tumor-supressor functions [260]. Thus, not only does DNA replication timing influence the rate of mutagenesis, it can also bias the location of rearrangement breakpoints.

The above data compare replication-timing profiles in normal cells to the acquisition of mutations in cancer cells, with the observation that at least some of the mutagenesis observed in cancer is collateral damage of having a normal replication-timing program. However, there is accumulating evidence that there are numerous alterations to the normal replication-timing program during carcinogenesis. In fact, there is abundant evidence indicating that changes in replication timing often accompany cancer development. While the extent to which these replication-timing changes influence the transformation process is still largely unknown, the presence of these changes in many different types of tumors indicates that altered replication timing may be an important component in tumor development.

Aberrant DNA Replication Timing in Cancer

DNA replication is a highly regulated process. For most of the genome, homologous loci replicate at the same time during S phase in a highly coordinated manner. Exceptions to this rule are represented by loci that display a

mono-allelic gene expression pattern. Thus, mono-allelically expressed genes such as imprinted genes, allelically excluded genes and genes on female X chromosomes replicate asynchronously with one allele replicating before the other. This replication pattern is very stable in normal cells, and is independent of transcription [167].

One well-documented change in cancer cells is the aberrant asynchronous replication of loci that normally replicate synchronously [261-263]. The early studies that looked at individuals with cancer assayed individual loci and, therefore, gave no indication of how widespread this aberrant asynchronous replication was throughout the genome. In contrast, a recent whole-genome replication timing study indicated that 9-18% of the genome undergoes a change in replication timing in leukemia cells compared to normal controls ([225]; see Fig. A1a). Changes in replication timing were detected on all chromosomes and were evenly distributed throughout the genome. Although there were slight differences between different types of leukemias, many of the changes in replication timing were common to all samples, suggesting that altered replication at specific locations is an early epigenetic event in cancer development [225]. In addition, many but not all of the replication-timing changes occurred near sites of genomic rearrangement. Indeed, a replication-timing change was found at a common site of translocation in leukemia cells [225]. However, all leukemia cells studied displayed this replication-timing change, but only a few displayed the actual translocation, indicating that replication-timing changes may predispose the cell to certain translocation events [225]. In addition, instead of genomic

rearrangements correlating with small local changes in DNA replication timing, the changes were extensive and extended hundreds of kilobases beyond the site of rearrangement. Similar to the studies mentioned previously, the replication timing changes were generally from late-replicating regions replicating earlier, and fewer early-replicating regions replicating later [225]. This study was not calibrated to detect replication asynchrony, so it is unclear whether these loci that change replication timing replicate synchronously or asynchronously. However, it is likely that some of the site-specific replication asynchrony that was observed in the above reports is reflected in this genome-wide analysis. It will be interesting to determine if the replication-timing changes that occur genome-wide in different types of cancers are the result of a specific chromosomal feature (e.g. asynchronous replication) or if this deregulation represents a more heterogenous, nonspecific change.

Surprisingly, the asynchronous replication pattern observed in cancer patients is not restricted to tumor tissue but also occurs in noncancerous cells as well [264-268]. This is best exemplified by the presence of aberrant asynchronous replication between alleles in the peripheral lymphocytes of individuals with solid tumors [264, 265]. This replication asynchrony has been documented at multiple loci for many cancer-related genes and many other genomic locations indicating that this is not just the deregulation of a single locus or a single chromosome, but a widespread phenomenon [264, 266, 269]. Interestingly, this altered replication-timing pattern is present in pre-malignant cells, in individuals pre-disposed to cancer, and in individuals living in polluted

areas with a high likelihood of getting cancer, suggesting that this may be an early event during carcinogenesis [261, 263, 265, 270]. The replication asynchrony observed in cancerous tissue and normal cells in individuals with cancer is generally a result of the earlier replication of one of the alleles [265, 268, 270], however, in some cases the delayed replication of one allele has been detected [269]. This aberrant asynchronous replication is heritable (i.e. the earlier-replicating allele will be earlier replicating in all subsequent generations) but not dependent on parental origin and therefore resembling the process of human X chromosome inactivation and/or allelic exclusion [271]. Chemotherapy is not sufficient to correct the cancer-associated replication asynchrony detected in lymphocytes of cancer patients [268], but allogeneic stem cell transplantation, in vitro fusion to normal cells, or inhibiting DNA methylation by treating cells with 5-azacytidine can switch replication back to the normal state [267, 271-273].

The replication-timing abnormalities mentioned above effect loci present on many different chromosomes and are detected sporadically throughout the genome. They generally result in the advancement of the replication-timing program such that one or both alleles will replicate earlier than normal. In contrast, a functionally distinct replication timing aberration has also been observed in tumor-derived cells. In 1967, Dr. Harald zur Hausen documented a delay in replication timing of individual chromosomes in cultured leukemia cells ([68]; see Fig. A1b). This chromosome-wide replication delay has since been observed in many different tumor-derived cell lines and primary tumor samples [74, 82]. These studies indicated that some tumor cells contain individual

chromosomes that are delayed in initiation and completion of DNA replication by 2-3 hours along their entire length [74, 75]. This replication delay affects the entire chromosome but does not disrupt the replication timing of other chromosomes in the cell, and is therefore controlled by a cis acting mechanism. Whole chromosomes that are delayed in DNA replication timing also exhibit a delay in mitotic chromosome condensation and are associated with highly aneuploid karyotypes [74, 77, 80]. In addition, it has been demonstrated that this chromosome-wide replication delay increases the rate of secondary chromosomal rearrangement on the affected chromosome by 30-80 fold, indicating that this chromosome-wide replication delay causes genomic instability [77]. Unlike the genome-wide replication-timing changes mentioned previously, which tend to be a very stable feature of cancer cells during the progression of the disease, chromosome-wide replication delay is a more transient feature. The inherent instability of the delayed chromosomes makes them prone to extreme fragmentation occurring over a relatively few cell divisions, which eventually results in highly rearranged chromosomes that no longer display replication delay [74]. The transient existence of chromosome-wide replication delay in cancer cells makes it an underappreciated, yet potentially important force driving mutagenesis in cancer cells.

It is apparent that chromosome-wide replication delay can have a profound impact on the structural stability of individual chromosomes by increasing the rate of chromosome rearrangements [77]. In contrast, it is unclear whether the other sporadic genome-wide replication-timing changes that occur at

specific loci are contributing to instability or are merely correlative. Due to the early onset of genome-wide replication-timing changes in cancer development (in some cases preceding malignancy) it is likely that this deregulation is directly or indirectly linked to transformation. As discussed below, replication-timing changes can be generated by different mechanisms and can lead to genetic and epigenetic changes within the cell. In the right context, any of these changes has the potential to influence cell growth and survival.

Causes and Consequences of Aberrant DNA Replication Timing

Depending on the context, some examples of aberrant DNA replication timing appear to have direct consequences on mutagenesis; however, in many cases the temporal order of events has not been established. Consequently, we are left with indirect conclusions to establish a timeline of events. The interconnectedness of many cellular properties like chromatin modifications, replication timing, transcription, and nuclear positioning complicates a cause and effect analysis, because the experimental manipulation of one will, in many cases, affect all three. This has led to the proposal that these properties are interdependent, such that a change in any one of them will have an effect on the others [13, 274, 275]. While keeping this in mind, we have highlighted some examples that indicate a direct relationship between specific cellular events and changes in replication timing.

A) Gene Expression Changes

DNA replication is a complicated cellular process involving the coordinated action of many different gene products. The deregulated expression of certain genes involved in DNA synthesis can cause defects in replication timing. For example, mutations in ORC proteins [30], cyclins [276], CDKs [277], nucleotide reductases [278], and other proteins involved in DNA structure checkpoints [277, 279, 280] have all been shown to cause abnormal DNA replication-timing patterns. Consistent with a relationship between chromatin modifications and DNA replication timing, the deregulation of many chromatin-modifying enzymes can impact the temporal replication of loci throughout the genome [281-286]. Disruption of these genes has a *trans*-acting effect, meaning it impacts the replication timing of distant loci on multiple chromosomes. Accordingly, deregulated HP1 gene expression was found to change the replication timing of 5-10% of genomic loci, suggesting a widespread effect [285]. In some cases, the replication timing changes that result from the deregulation of *trans*-acting factors resembles the genome-wide replication-timing changes seen in some cancers [225], which indicates that deregulation of replication components and chromatin modifiers may be one cause of abnormal replication timing in cancer cells.

One unique aspect of changes in replication timing is that the consequences on transcription are only observed in *cis*. For example, it was found that the advanced replication timing of loci on the inactive X chromosome was critical for their escape from gene inactivation [287]. Other studies have indicated that replication timing changes at specific loci can occur upstream of

gene transcription changes [20, 288]. However, it should be noted that a change in the replication timing of a particular gene does not always cause a change in transcription [269]. This has led to the idea that a change in replication timing is not sufficient to cause a change in transcription, and that other factors, such as the presence and activity of transcriptional activators, are required. Thus, replication timing may not affect transcription directly but, rather, affect transcriptional competence [31].

B) Epigenetic Changes

Although DNA replication timing is strongly associated with gene expression, the association is even stronger with some epigenetic modifications. Certain chromatin and DNA methylation states appear to have a substantial impact on replication timing. Tethering a histone acetylase to a late replicating origin is sufficient to convert it to early replicating, and the opposite is true when a histone deacetylase is brought to an early replicating origin [289]. Many different groups have observed that histone acetylation changes can precede changes in replication timing [290-293], which is consistent with histone deacetylation occurring at the G1/S transition prior to late-origin replication [294]. Changes in histone methylation have also been implicated in causing changes in DNA replication timing [282, 286], bolstering the concept that chromatin accessibility and DNA replication timing go hand-in-hand. Similar to many histone modifications, changes in DNA methylation have also been found to precede changes in replication timing [283, 295], and manipulation of DNA methylation

can even reverse aberrant replication timing under certain conditions [267, 271, 272].

The studies mentioned above indicate that changes in histone modifications can switch the replication timing of an origin, but some chromatin modifiers can also inhibit the firing of an origin altogether [284]. Indeed, the addition or subtraction of origins can affect the replication timing of adjacent regions by decreasing or increasing the time it takes the replication fork to reach them [19, 21, 288, 296]. Since most loci are replicated by clusters of origins, the addition or deletion of one origin generally won't have much of an impact on replication timing. However, if a change in origin usage occurs in a region that is devoid of origins, like a transition region, then the impact can be large. This differential origin usage model has been used to describe why transition regions are so prone to drastic changes in replication timing [240]. In support of this model, the addition of an origin at the Igh locus, which lies in a transition region, was found to coincide with a shift from late to early replication [288].

In keeping with the interdependent theme of this section, changes in replication timing can also precede changes in epigenetic modifications. Studies in Dr. Howard Cedar's lab found that plasmid DNA injected into a cell in early S-phase will be packaged into acetylated chromatin while DNA injected into cells in late S phase will be associate with hypoacetylated chromatin [297]. This was followed by the demonstration that a plasmid containing hypoacetylated histones injected into cells in early S phase will become remodeled with acetylated histones during replication and vice versa [298]. This suggests that the time of

DNA replication within S phase can dictate the acetylation state of histones that are loaded onto DNA. This has led some to speculate that the differential association of various chromatin modifiers with DNA throughout S phase is responsible for the close association between DNA replication timing and chromatin modifications [275]. This notion is supported by the demonstration that some repressive histone modifiers and transcriptional repressors only localize to replication foci in mid-late S phase [299-301]. Therefore, a change in replication timing can change the chromatin landscape, and transcriptional competence, of a particular region by dictating which chromatin modifiers can associate during replication. DNA replication-timing changes have also been observed to occur before DNA methylation changes, indicating that the temporal order of replication can also affect the methylation status of DNA [302, 303].

Although the cause and effect relationship between DNA replication timing and chromatin/DNA modifications has been studied extensively, there is still much more to learn. Furthermore, we still lack a good understanding of how DNA replication timing and 3-dimensional nuclear structure affect one another. Some studies have indicated that a change in nuclear position is not sufficient to change DNA replication timing [14, 292, 304], which would suggest that DNA replication timing determines nuclear position. However, other studies have suggested that the presence of replication foci in specific nuclear compartments dictates the temporal order of replication [305]. It is likely that both models are correct in certain contexts. As is the case with transcription and epigenetic

modifications, DNA replication timing seems to be controlled by as many cellular events as it controls.

The effects of aberrant DNA replication timing on chromatin structure can extend beyond S phase as well. It has been observed that chromosome-wide delayed DNA replication can lead to abnormal mitotic chromosome condensation in early mitosis [30, 74]. This delay in condensation coincides with a delay in the recruitment of Aurora B kinase resulting in a delay in the mitosis-specific phosphorylation of histone H3 [74, 75]. Therefore, delayed replication can lead to chromosomes that are in an “interphase state” of condensation during mitosis [74].

C) Genetic Changes

Perhaps it is no surprise that genetic changes can be the cause of aberrant DNA replication timing, but the varied types of genetic damage and replication-timing changes discussed in this section indicate that this relationship is far from straightforward. Treating cells with various DNA damaging agents, such as ionizing radiation, hydrogen peroxide, and mitomycin C can lead to aberrant replication timing [270, 306]. Furthermore, double-strand breaks caused by site-specific recombination, ionizing radiation and endonuclease digestion can induce a chromosome-wide delay in replication [77, 80], indicating that DNA damage can cause different types of replication-timing defects.

Different types of DNA damaging agents can produce different types of mutations, and different types of mutations have been implicated in replication-

timing changes. For example, nucleotide substitution at CTCF binding sites can deregulate allele-specific replication in imprinted regions [307]. In addition, telomere shortening can advance the replication timing of telomeric origins [308]. It has been known for some time that the juxtaposition of genetic regions to non-native loci can cause replication-timing changes [24, 309, 310]. One common type of genomic rearrangement in cancer is inter-chromosomal translocation, which brings together two regions from two different chromosomes. It has been observed that chromosomal translocations often accompany replication-timing changes [18, 311, 312]. Many of these replication-timing defects result from the newly acquired replication of homologous loci due to the juxtaposition of one locus to another [311, 312]. In fact, most translocations that juxtapose regions of differential replication timing result in the earlier or later replication of at least one of the translocated alleles [225]. Because these abrupt replication-timing changes occur by juxtaposing an early-replicating region with a late-replicating region, it should be noted that a translocation involving regions of similar replication timing would not be expected to result in a replication-timing change [225]. These studies indicate that the majority of translocation events cause replication-timing changes that are relatively minor, and only affect local sequences or domains (Fig. A1a).

A recent phenomenon of localized hypermutation, termed ‘‘kataegis,’’ was observed in some cancer cells [116]. Kataegis is characterized by an increase in the frequency of SNVs in a particular region of the genome. Regions of kataegis differed between cancers, but usually colocalized with somatic rearrangements.

While the mechanisms responsible for kataegis remain unknown, we propose that the localized replication-timing changes that occur near or prior to translocation breakpoints [225] could be responsible for the localized mutagenesis observed in the localized regions with kataegis (Fig. A2a).

In addition, genomic rearrangements can have larger effects than just the change in temporal replication of a specific locus or domain. Unlike juxtaposition-induced replication asynchrony, where the rearranged locus is the only site of aberrant replication timing, certain chromosomal rearrangements cause a chromosome-wide delay in replication timing of the entire chromosome [74, 77, 79, 81, 122]. This chromosome-wide effect is a result of the disruption of cis-acting elements that normally act to ensure the proper replication-timing program of individual chromosomes. A recent series of “chromosome-engineering” studies led to the identification of a discrete cis-acting locus that controls chromosome-wide replication timing and structural stability of human chromosome 6 [122]. Molecular characterization of this chromosome 6 locus identified a large intergenic non-coding RNA gene, which was named asynchronous replication and autosomal RNA on chromosome 6 (ASAR6). Cre/loxP-mediated disruption of the ASAR6 gene results in extremely late replication, an under-condensed appearance during mitosis, and structural instability of human chromosome 6 [77, 122]. In a separate series of experiments, it was found that disruption of the large non-coding RNA gene Xist, results in extremely late replication, abnormal chromatin structure and instability of the X chromosome [79, 81]. The Xist gene resides within the X inactivation center, and is known to participate in the

silencing of genes during dosage compensation in female cells [184]. Interestingly, ASAR6 shares many characteristics with Xist, including random mono-allelic expression, asynchronous replication timing, and regulation of the expression of linked mono-allelic genes [122].

Furthermore, this chromosome-wide delayed replication timing phenotype has been detected on chromosome rearrangements involving many different human and mouse chromosomes [74, 77, 80, 122]. Therefore, it seems likely that all mammalian chromosomes contain loci that function to regulate chromosome-wide replication timing, mitotic condensation and stability of individual chromosomes. Given the similarities in structure and function of the two loci characterized to date, Xist and ASAR6, it was proposed that all mammalian chromosomes contain functional chromosome “inactivation/stability centers” that act to maintain proper replication timing and structural stability of individual chromosomes [122]. Under this scenario every mammalian chromosome contains four cis-acting elements, origins of replication, centromeres, telomeres, and ‘inactivation/stability centers’, all functioning to ensure proper replication, segregation and stability of individual chromosomes [85].

Chromosome-wide delay in replication timing results in at least two distinct types of genomic instability. The first is chromosome instability (CIN), which is characterized by an increase in the rate at which cells gain or lose entire chromosomes [83]. Thus, cells with chromosome-wide delayed replication timing of individual chromosomes display frequent gains or losses of entire chromosomes resulting in dramatic aneuploidy affecting the entire karyotype [74,

75]. In addition, cells containing chromosome-wide delayed replication contain abnormal mitotic spindles, abnormal centrosome number, and an increased frequency of endoreduplication [75]. It is unclear how chromosome-wide replication delay on one chromosome is causing these events, but these factors can certainly explain the CIN observed in cells with individual chromosomes with the delayed replication phenotype. The second type of instability observed in cells with chromosome-wide delayed replication is chromosome structure instability, which is characterized by an increase in the rate that new chromosomal rearrangements occur [77]. This structural instability is primarily observed on the affected chromosome, but other chromosomes can participate in inter-chromosomal translocations with the delayed chromosome, indicting that delayed replication on one chromosome can destabilize the structural integrity of all chromosomes within the cell [77].

This structural instability of individual chromosomes is reminiscent of the newly described phenomenon “chromothripsis”, which is present in some but not all cancers [115, 116]. Chromothripsis appears to be a cataclysmic event in which one or a few chromosomes or chromosome arms are fragmented and then reassembled in a haphazard manner. The sequences at the junctions showed either a lack of homology or microhomology between the joined segments, suggesting that the ends were joined by non-homologous end joining (NHEJ) pathway. In addition, the complex chromosome rearrangements associated with genomic disorders in humans were recently found to resemble chromothripsis [117, 118]. Sequencing the breakpoints at these complex rearrangements

identified characteristic features, including small templated insertions of nearby sequences and microhomologies, suggestive of replicative processes. These observations led the Lupski group to propose the term ‘‘chromoanasynthesis’’ as an alternative descriptor to chromothripsis for the shattering and reassembly of chromosomes via replicative mechanisms [117]. The Lupski group proposed a microhomology mediated break induced replication (MMBIR) and a related fork stalling and template switching (FoSTeS) model for the origin of these complex rearrangements [119]. The distinction between MMBIR/FoSTeS and NHEJ is that the microhomology junctions in MMBIR/FoSTeS are followed by stretches of DNA sequence derived from elsewhere, usually nearby. The MMBIR/FoSTeS models involve stalled DNA replication forks that are resolved by replication restart using short stretches of homology [119]. Thus, the stalled replication forks of the MMBIR/FoSTeS pathways could potentially be caused by the premature condensation of partially replicated chromosomes as they enter mitosis. Thus, our model for the instability of individual chromosomes includes: 1) delayed replication timing of individual chromosomes caused by genetic disruption of an “inactivation/stability center”, 2) delayed recruitment of Aurora B resulting in delayed mitotic chromosome condensation, 3) delayed mitotic spindle attachment leading to chromosome mis-segregation and the formation of micronuclei, 4) checkpoint adaptation and the onset of mitotic chromosome condensation prior to the completion of DNA synthesis leading to stalled replication forks, and 5) multiple rearrangements generated at the stalled replication forks via NHEJ and/or MMBIR/FoSTeS type mechanisms (Fig. A2b).

Concluding Remarks

Changes in the replication timing of individual loci throughout the genome in cancer cells can occur in two ways: either the advanced replication of individual loci or the delayed replication of individual loci. It currently appears that the advanced replication of individual loci is more common than delayed replication and it is unclear why this is. Furthermore, the change in replication timing of individual loci generally appears to result in aberrant asynchronous replication, but it is not known if this is always the case. Regardless, it is likely that the change in replication timing of individual loci sporadically throughout the genome can be caused by two different mechanisms. The first, which is highlighted by the Ryba et al. study, is genomic rearrangements resulting in a local change in replication timing [225]. In this scenario, regions of divergent replication timing are juxtaposed following a rearrangement and one of those regions changes its replication timing in response to its new environment. This can explain a switch in replication timing of an individual locus and aberrant asynchronous replication. However, this study also found replication-timing changes that did not coincide with rearrangement breakpoints [225]. Therefore, localized changes in replication timing may actually precede the rearrangement events at specific loci.

Delayed replication of individual chromosomes is a functionally distinct phenomenon from the aberrant asynchronous replication of individual loci. Although chromosome-wide replication delay does result in replication

asynchrony, the asynchrony is at the chromosome level rather than at specific loci scattered throughout the genome. This replication delay affects the entire chromosome but the replication timing of all the other chromosomes within the cell remains normal. It is currently unknown whether chromosome-wide replication delay has any effect on gene expression. However, it is likely that gene expression is affected by delaying the replication timing of an entire chromosome. Thus, this process closely resembles X chromosome inactivation in female mammalian cells, where the inactive X chromosome undergoes gene inactivation and a chromosome-wide delay in replication timing while the active X chromosome remains earlier replicating [313].

In summary, DNA replication timing has helped shape the genomic landscape of many, if not all, eukaryotic organisms. By separating the genome into regions prone to hypomutability (early-replicating) and hypermutability (late-replicating and transition regions), the replication-timing program dictates that the mutation rate is higher in some regions than in others. The reason for this is currently unclear, however, it most likely results from the predominant use of different DNA repair pathways in early versus late S phase [232, 234]. On one hand, it seems detrimental to have a high rate of mutagenesis anywhere in the genome, as most mutations either have no effect on fitness or are detrimental to the organism [314]. And if the cell uses error-free DNA repair pathways in early S phase, then why doesn't it also use these same pathways in late S phase? A closer look reveals that it may be slightly beneficial to keep a higher mutation rate in late-replicating regions. Thus, most repetitive sequences tend to be late

replicating and since many of these sequences are leftover viral or transposon integrations, this might be one way to ensure that non-native sequences mutate more frequently. This would result in a more rapid accumulation of mutations that inactivate viral or transposon gene products. Furthermore, tissue-specific genes are more likely to reside in late-replicating regions and since these genes are under greater pressure to adapt, a slightly higher mutation rate in those genes would provide the organism with a greater ability to respond to a changing environment.

Although there can be some benefits to having hypermutable regions of the genome in germ cells, it is hard to explain the presence of this phenomenon in somatic cells. Thus, the mutagenesis associated with the normal replication-timing program does appear to be contributing to the accumulation of mutations during cancer development. Although the normal replication-timing program may be responsible for an increased mutation rate in certain regions of the genome, the selective pressure to maintain early and late replicating regions ensures proper epigenetic regulation of gene expression and helps maintain genome stability.

Figures

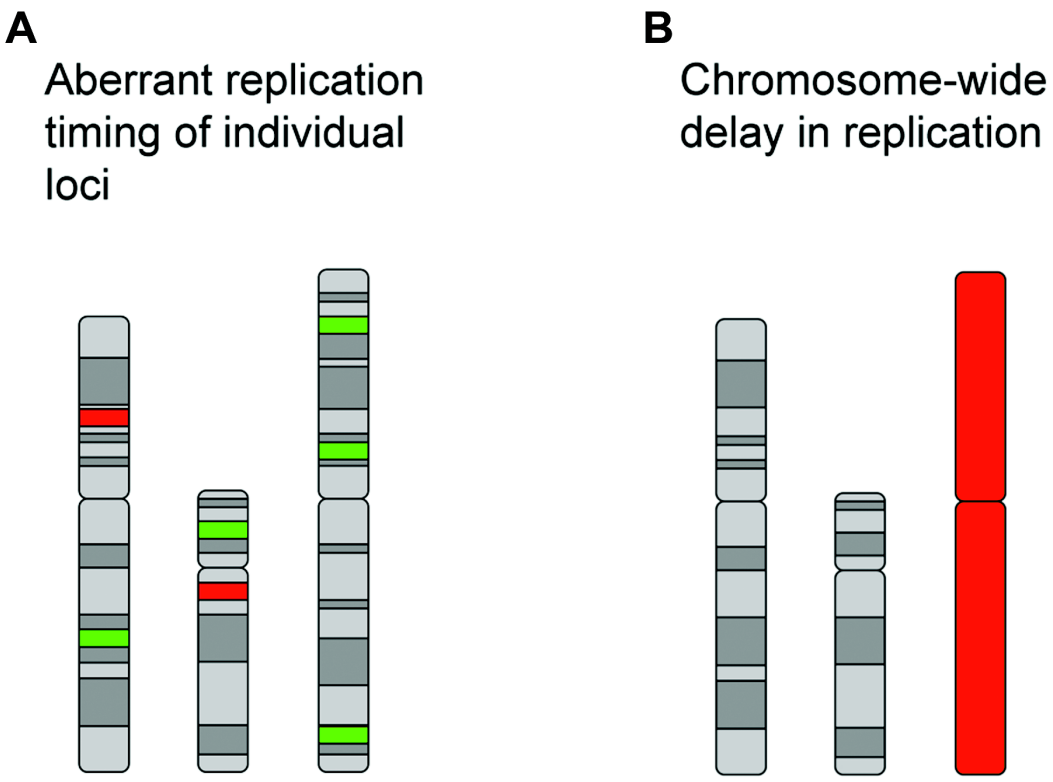


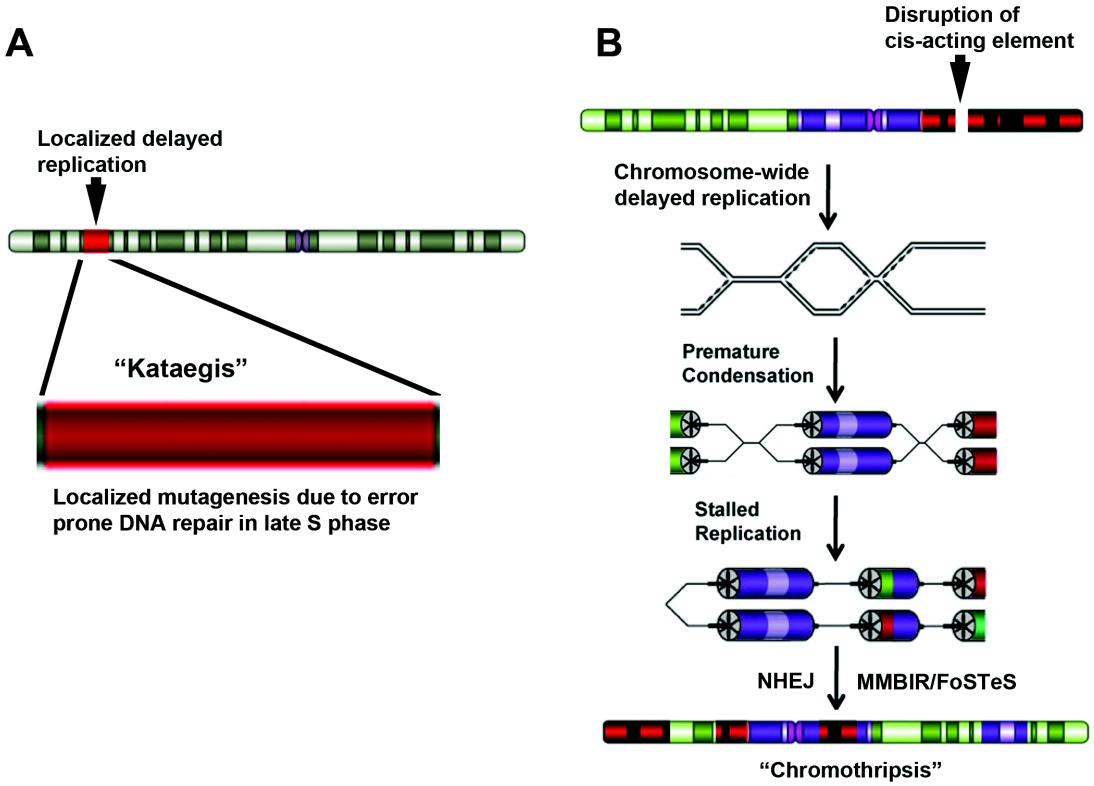
Figure A1: Acquired alterations in DNA replication timing in cancer cells.

A) Examples of individual loci that display a shift in replication timing. Loci that shift to an earlier time of replication are indicated in green, and regions that shift to a later time of replication are indicated in red. Three different chromosomes are shown. **B**) An example of an individual chromosome with a chromosome-wide delay in replication (red). Two chromosomes with normal replication timing are shown in gray.

Figure A2: Models for localized genomic instability in cancer cells.

A) Aberrant late replication model for kataegis. A localized region of a chromosome has acquired abnormally late replication (red) either as a result of chromosome rearrangement or as a result of a localized shift in the replication timing program [29]. Increased mutagenesis is induced in the late replicating region due to error prone repair mechanisms functioning during late replication. **B**) Aberrant late replication model for Chromothripsis. Disruption of discrete *cis*-acting loci result in a chromosome-wide delay in replication timing. Mitotic chromosome condensation initiates on the delayed chromosome prior to completion of DNA synthesis resulting in premature chromosome condensation, stalled replication forks, and rearrangement of the affected chromosomes via non-homologous end joining (NHEJ), microhomology mediated break induced replication (MMBIR) and fork stalling and template switching (FoSTeS) mechanisms. The resulting chromosome contains numerous structural alterations (translocations, deletions, inversions, and duplications). (Figure on next page).

Figure A2: Models for localized genomic instability in cancer cells.



REFERENCES

1. Robinson, N.P. and S.D. Bell, *Origins of DNA replication in the three domains of life.* FEBS J, 2005. **272**(15): p. 3757-66.
2. Huberman, J.A. and A.D. Riggs, *Autoradiography of chromosomal DNA fibers from Chinese hamster cells.* Proc Natl Acad Sci U S A, 1966. **55**(3): p. 599-606.
3. Diffley, J.F. and K. Labib, *The chromosome replication cycle.* J Cell Sci, 2002. **115**(Pt 5): p. 869-72.
4. Li, F., et al., *Spatial distribution and specification of mammalian replication origins during G1 phase.* J Cell Biol, 2003. **161**(2): p. 257-66.
5. Dimitrova, D.S., et al., *Mammalian nuclei become licensed for DNA replication during late telophase.* J Cell Sci, 2002. **115**(Pt 1): p. 51-9.
6. Wu, J.R. and D.M. Gilbert, *A distinct G1 step required to specify the Chinese hamster DHFR replication origin.* Science, 1996. **271**(5253): p. 1270-2.
7. Cvetic, C. and J.C. Walter, *Eukaryotic origins of DNA replication: could you please be more specific?* Semin Cell Dev Biol, 2005. **16**(3): p. 343-53.
8. Labib, K., *How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells?* Genes Dev, 2010. **24**(12): p. 1208-19.

9. Berezney, R., D.D. Dubey, and J.A. Huberman, *Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci.* Chromosoma, 2000. **108**(8): p. 471-84.

10. Hiratani, I., et al., *Replication timing and transcriptional control: beyond cause and effect--part II.* Curr Opin Genet Dev, 2009. **19**(2): p. 142-9.

11. Dimitrova, D.S. and D.M. Gilbert, *The spatial position and replication timing of chromosomal domains are both established in early G1 phase.* Mol Cell, 1999. **4**(6): p. 983-93.

12. Gilbert, D.M., *Cell fate transitions and the replication timing decision point.* J Cell Biol, 2010. **191**(5): p. 899-903.

13. Gilbert, D.M., et al., *Space and time in the nucleus: developmental control of replication timing and chromosome architecture.* Cold Spring Harb Symp Quant Biol, 2010. **75**: p. 143-53.

14. Heun, P., et al., *The positioning and dynamics of origins of replication in the budding yeast nucleus.* J Cell Biol, 2001. **152**(2): p. 385-400.

15. Ryba, T., et al., *Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types.* Genome Res, 2010. **20**(6): p. 761-70.

16. Goren, A. and H. Cedar, *Replicating by the clock.* Nat Rev Mol Cell Biol, 2003. **4**(1): p. 25-32.

17. Belyaev, N.D., A.M. Keohane, and B.M. Turner, *Histone H4 acetylation and replication timing in Chinese hamster chromosomes.* Exp Cell Res, 1996. **225**(2): p. 277-85.

18. Calza, R.E., et al., *Changes in gene position are accompanied by a change in time of replication.* Cell, 1984. **36**(3): p. 689-96.

19. Forrester, W.C., et al., *A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus.* Genes Dev, 1990. **4**(10): p. 1637-49.

20. Cimbora, D.M., et al., *Long-distance control of origin choice and replication timing in the human beta-globin locus are independent of the locus control region.* Mol Cell Biol, 2000. **20**(15): p. 5581-91.

21. Gilbert, D.M., *Replication timing and transcriptional control: beyond cause and effect.* Curr Opin Cell Biol, 2002. **14**(3): p. 377-83.

22. Simon, I., et al., *Developmental regulation of DNA replication timing at the human beta globin locus.* EMBO J, 2001. **20**(21): p. 6150-7.

23. Raghuraman, M.K., B.J. Brewer, and W.L. Fangman, *Cell cycle-dependent establishment of a late replication program.* Science, 1997. **276**(5313): p. 806-9.

24. Ofir, R., et al., *Position effect of human telomeric repeats on replication timing.* Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11434-9.

25. Barbosa, A.C., P.A. Otto, and A.M. Vianna-Morgante, *Replication timing of homologous alpha-satellite DNA in Roberts syndrome.* Chromosome Res, 2000. **8**(7): p. 645-50.

26. State, M.W., et al., *Epigenetic abnormalities associated with a chromosome 18(q21-q22) inversion and a Gilles de la Tourette syndrome phenotype.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4684-9.

27. Watanabe, Y., et al., *Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions.* Hum Mol Genet, 2002. **11**(1): p. 13-21.

28. Korenstein-Ilan, A., et al., *Allele-specific replication associated with aneuploidy in blood cells of patients with hematologic malignancies.* Cancer Genet Cytogenet, 2002. **139**(2): p. 97-103.

29. Ryba, T., et al., *Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia.* Genome Res, 2012.

30. Loupart, M.L., S.A. Krause, and M.S. Heck, *Aberrant replication timing induces defective chromosome condensation in Drosophila ORC2 mutants.* Curr Biol, 2000. **10**(24): p. 1547-56.

31. Hiratani, I. and D.M. Gilbert, *Replication timing as an epigenetic mark.* Epigenetics, 2009. **4**(2): p. 93-7.

32. Belmont, A.S., *Mitotic chromosome structure and condensation.* Curr Opin Cell Biol, 2006. **18**(6): p. 632-8.

33. Watrin, E. and V. Legagneux, *Introduction to chromosome dynamics in mitosis.* Biol Cell, 2003. **95**(8): p. 507-13.

34. Carmena, M. and W.C. Earnshaw, *The cellular geography of aurora kinases.* Nat Rev Mol Cell Biol, 2003. **4**(11): p. 842-54.

35. Porter, I.M., G.A. Khoudoli, and J.R. Swedlow, *Chromosome condensation: DNA compaction in real time.* Curr Biol, 2004. **14**(14): p. R554-6.

36. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., *Molecular Biology of the Cell*. 4th ed. 2002, New York: Garland Science.

37. Pflumm, M.F., *The role of DNA replication in chromosome condensation.* Bioessays, 2002. **24**(5): p. 411-8.

38. Johnson, R.T. and P.N. Rao, *Mammalian cell fusion: induction of premature chromosome condensation in interphase nuclei.* Nature, 1970. **226**(5247): p. 717-22.

39. Mullinger, A.M. and R.T. Johnson, *Units of chromosome replication and packing.* J Cell Sci, 1983. **64**: p. 179-93.

40. Christensen, T.W. and B.K. Tye, *Drosophila MCM10 interacts with members of the prereplication complex and is required for proper chromosome condensation.* Mol Biol Cell, 2003. **14**(6): p. 2206-15.

41. McHugh, B. and M.M. Heck, *Regulation of chromosome condensation and segregation.* Curr Opin Genet Dev, 2003. **13**(2): p. 185-90.

42. Gotoh, E., *Visualizing the dynamics of chromosome structure formation coupled with DNA replication.* Chromosoma, 2007. **116**(5): p. 453-62.

43. Li, J.J. and S.A. Li, *Mitotic kinases: the key to duplication, segregation, and cytokinesis errors, chromosomal instability, and oncogenesis.* Pharmacol Ther, 2006. **111**(3): p. 974-84.

44. Fatoba, S.T. and A.L. Okorokov, *Human SIRT1 associates with mitotic chromatin and contributes to chromosomal condensation.* Cell Cycle. **10**(14): p. 2317-22.

45. Chmielewski, J.P., et al., *Drosophila Psf2 has a role in chromosome condensation.* Chromosoma.

46. Kastan, M.B. and J. Bartek, *Cell-cycle checkpoints and cancer.* Nature, 2004. **432**(7015): p. 316-23.

47. Blagosklonny, M.V. and A.B. Pardee, *The restriction point of the cell cycle.* Cell Cycle, 2002. **1**(2): p. 103-10.

48. Bartek, J., C. Lukas, and J. Lukas, *Checking on DNA damage in S phase.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 792-804.

49. Vermeulen, K., D.R. Van Bockstaele, and Z.N. Berneman, *The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer.* Cell Prolif, 2003. **36**(3): p. 131-49.

50. Jones, R.M. and E. Petermann, *Replication fork dynamics and the DNA damage response.* Biochem J, 2012. **443**(1): p. 13-26.

51. Liu, Q., et al., *Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint.* Genes Dev, 2000. **14**(12): p. 1448-59.

52. Dimitrova, D.S. and D.M. Gilbert, *Temporally coordinated assembly and disassembly of replication factories in the absence of DNA synthesis.* Nat Cell Biol, 2000. **2**(10): p. 686-94.

53. Lopes, M., et al., *The DNA replication checkpoint response stabilizes stalled replication forks.* Nature, 2001. **412**(6846): p. 557-61.

54. Donzelli, M. and G.F. Draetta, *Regulating mammalian checkpoints through Cdc25 inactivation.* EMBO Rep, 2003. **4**(7): p. 671-7.

55. Costanzo, V., et al., *Reconstitution of an ATM-dependent checkpoint that inhibits chromosomal DNA replication following DNA damage.* Mol Cell, 2000. **6**(3): p. 649-59.

56. Sudakin, V., G.K. Chan, and T.J. Yen, *Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2.* J Cell Biol, 2001. **154**(5): p. 925-36.

57. Taylor, S.S., M.I. Scott, and A.J. Holland, *The spindle checkpoint: a quality control mechanism which ensures accurate chromosome segregation.* Chromosome Res, 2004. **12**(6): p. 599-616.

58. Fang, G., H. Yu, and M.W. Kirschner, *The checkpoint protein MAD2 and the mitotic regulator CDC20 form a ternary complex with the anaphase-promoting complex to control anaphase initiation.* Genes Dev, 1998. **12**(12): p. 1871-83.

59. Musacchio, A. and E.D. Salmon, *The spindle-assembly checkpoint in space and time.* Nat Rev Mol Cell Biol, 2007. **8**(5): p. 379-93.

60. Walczak, C.E., S. Cai, and A. Khodjakov, *Mechanisms of chromosome behaviour during mitosis.* Nat Rev Mol Cell Biol, 2010. **11**(2): p. 91-102.

61. Bouska, A., et al., *Mdm2 promotes genetic instability and transformation independent of p53.* Mol Cell Biol, 2008. **28**(15): p. 4862-74.

62. Liu, P.K., et al., *Analysis of genomic instability in Li-Fraumeni fibroblasts with germline p53 mutations.* Oncogene, 1996. **12**(11): p. 2267-78.

63. Hanks, S., et al., *Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B.* Nat Genet, 2004. **36**(11): p. 1159-61.

64. Sandell, L.L. and V.A. Zakian, *Loss of a yeast telomere: arrest, recovery, and chromosome loss.* Cell, 1993. **75**(4): p. 729-39.

65. Syljuasen, R.G., et al., *Adaptation to the ionizing radiation-induced G2 checkpoint occurs in human cells and depends on checkpoint kinase 1 and Polo-like kinase 1 kinases.* Cancer Res, 2006. **66**(21): p. 10253-7.

66. Syljuasen, R.G., *Checkpoint adaptation in human cells.* Oncogene, 2007. **26**(40): p. 5833-9.

67. Galgoczy, D.J. and D.P. Toczyski, *Checkpoint adaptation precedes spontaneous and damage-induced genomic instability in yeast.* Mol Cell Biol, 2001. **21**(5): p. 1710-8.

68. zur Hausen, H., *Chromosomal changes of similar nature in seven established cell lines derived from the peripheral blood of patients with leukemia.* J Natl Cancer Inst, 1967. **38**(5): p. 683-96.

69. Otto, P.G., P.A. Otto, and E. Therman, *The behavior of allocyclic chromosomes in Bloom's syndrome.* Chromosoma, 1981. **84**(3): p. 337-44.

70. Kuhn, E.M., E. Therman, and D.A. Buchler, *Do individual allocyclic chromosomes in metaphase reflect their interphase domains?* Hum Genet, 1987. **77**(3): p. 210-3.

71. Obe, G., et al., *Premature chromosome condensation in a case of Fanconi's anemia.* Humangenetik, 1975. **28**(2): p. 159-62.

72. Gripenberg, U., *The cytological behavior of a human ring-chromosome.* Chromosoma, 1967. **20**(3): p. 284-9.

73. Bobrow, M., L.F. Joness, and G. Clarke, *A complex chromosomal rearrangement with formation of a ring 4.* J Med Genet, 1971. **8**(2): p. 235-9.

74. Smith, L., A. Plug, and M. Thayer, *Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations.* Proc Natl Acad Sci U S A, 2001. **98**(23): p. 13300-5.

75. Chang, B.H., et al., *Chromosomes with delayed replication timing lead to checkpoint activation, delayed recruitment of Aurora B and chromosome instability.* Oncogene, 2007. **26**(13): p. 1852-61.

76. Hendzel, M.J., et al., *Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation.* Chromosoma, 1997. **106**(6): p. 348-60.

77. Breger, K.S., L. Smith, and M.J. Thayer, *Engineering translocations with delayed replication: evidence for cis control of chromosome replication timing.* Hum Mol Genet, 2005. **14**(19): p. 2813-27.

78. Kistenmacher, M.L. and H.H. Punnett, *Comparative behavior of ring chromosomes.* Am J Hum Genet, 1970. **22**(3): p. 304-18.

79. Diaz-Perez, S., et al., *The element(s) at the nontranscribed Xist locus of the active X chromosome controls chromosomal replication timing in the mouse.* Genetics, 2005. **171**(2): p. 663-72.

80. Breger, K.S., et al., *Ionizing radiation induces frequent translocations with delayed replication and condensation.* Cancer Res, 2004. **64**(22): p. 8231-8.

81. Diaz-Perez, S.V., et al., *A deletion at the mouse Xist gene exposes trans-effects that alter the heterochromatin of the inactive X chromosome and the*

replication time and DNA stability of both X chromosomes. Genetics, 2006. **174**(3): p. 1115-33.

82. Miles, C.P. and F. O'Neill, *3H labeling patterns of permanent cell line chromosomes showing pulverization or accentuated secondary constrictions.* J Cell Biol, 1969. **40**(2): p. 553-61.

83. Lengauer, C., K.W. Kinzler, and B. Vogelstein, *Genetic instabilities in human cancers.* Nature, 1998. **396**(6712): p. 643-9.

84. Crasta, K., et al., *DNA breaks and chromosome pulverization from errors in mitosis.* Nature, 2012. **482**(7383): p. 53-8.

85. Thayer, M.J., *Mammalian chromosomes contain cis-acting elements that control replication timing, mitotic condensation, and stability of entire chromosomes.* Bioessays, 2012. **34**(9): p. 760-70.

86. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.

87. Vogelstein, B. and K.W. Kinzler, *The multistep nature of cancer.* Trends Genet, 1993. **9**(4): p. 138-41.

88. Mitelman, F., F. Mertens, and B. Johansson, *A breakpoint map of recurrent chromosomal rearrangements in human neoplasia.* Nat Genet, 1997. **15 Spec No**: p. 417-74.

89. Mitelman, F., *Recurrent chromosome aberrations in cancer.* Mutat Res, 2000. **462**(2-3): p. 247-53.

90. Mitelman, F., B. Johansson, and F. Mertens. *Mitelman Database of Chromosome Aberrations in Cancer*. 2009; Available from: http://cgap.nci.nih.gov/Chromosomes/Mitelman.

91. Myllykangas, S., et al., *DNA copy number amplification profiling of human neoplasms.* Oncogene, 2006. **25**(55): p. 7324-32.

92. Gaasenbeek, M., et al., *Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers.* Cancer Res, 2006. **66**(7): p. 3471-9.

93. Vogelstein, B., et al., *Allelotype of colorectal carcinomas.* Science, 1989. **244**(4901): p. 207-11.

94. Boige, V., et al., *Concerted nonsyntenic allelic losses in hyperploid hepatocellular carcinoma as determined by a high-resolution allelotype.* Cancer Res, 1997. **57**(10): p. 1986-90.

95. Seymour, A.B., et al., *Allelotype of pancreatic adenocarcinoma.* Cancer Res, 1994. **54**(10): p. 2761-4.

96. Radford, D.M., et al., *Allelotyping of ductal carcinoma in situ of the breast: deletion of loci on 8p, 13q, 16q, 17p and 17q.* Cancer Res, 1995. **55**(15): p. 3399-405.

97. Stoler, D.L., et al., *The onset and extent of genomic instability in sporadic colorectal tumor progression.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 15121-6.

98. Rajagopalan, H. and C. Lengauer, *CIN-ful cancers.* Cancer Chemother Pharmacol, 2004. **54 Suppl 1**(1): p. S65-8.

99. Lengauer, C., K.W. Kinzler, and B. Vogelstein, *Genetic instabilities in human cancers.* Nature, 1998. **396**(6712): p. 643-9.

100. Nowak, M.A., et al., *The role of chromosomal instability in tumor initiation.* Proc Natl Acad Sci U S A, 2002. **99**(25): p. 16226-31.

101. Aubele, M., et al., *Intratumoral heterogeneity in breast carcinoma revealed by laser-microdissection and comparative genomic hybridization.* Cancer Genet Cytogenet, 1999. **110**(2): p. 94-102.

102. Sjoblom, T., et al., *The consensus coding sequences of human breast and colorectal cancers.* Science, 2006. **314**(5797): p. 268-74.

103. Cahill, D.P., et al., *Mutations of mitotic checkpoint genes in human cancers.* Nature, 1998. **392**(6673): p. 300-3.

104. Zhou, H., et al., *Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation.* Nat Genet, 1998. **20**(2): p. 189-93.

105. Shima, N., et al., *A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice.* Nat Genet, 2007. **39**(1): p. 93-8.

106. Wang, Y., et al., *Mutation in Rpa1 results in defective DNA double-strand break repair, chromosomal instability and cancer in mice.* Nat Genet, 2005. **37**(7): p. 750-5.

107. Spencer, F., et al., *Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae.* Genetics, 1990. **124**(2): p. 237-49.

108. Little, J.B., *Radiation-induced genomic instability.* Int J Radiat Biol, 1998. **74**(6): p. 663-71.

109. Kadhim, M.A., et al., *Long-term genomic instability in human lymphocytes induced by single-particle irradiation.* Radiat Res, 2001. **155**(1 Pt 1): p. 122-6.

110. Pongsaensook, P., et al., *Cis-acting transmission of genomic instability.* Mutat Res, 2004. **568**(1): p. 49-68.

111. Marder, B.A. and W.F. Morgan, *Delayed chromosomal instability induced by DNA damage.* Mol Cell Biol, 1993. **13**(11): p. 6667-77.

112. Moore, S.R., D. Papworth, and A.J. Grosovsky, *Non-random distribution of instability-associated chromosomal rearrangement breakpoints in human lymphoblastoid cells.* Mutat Res, 2006. **600**(1-2): p. 113-24.

113. Moore, S.R., et al., *Spontaneous and radiation-induced genomic instability in human cell lines differing in cellular TP53 status.* Radiat Res, 2005. **164**(4 Pt 1): p. 357-68.

114. Aguilera, A. and B. Gomez-Gonzalez, *Genome instability: a mechanistic view of its causes and consequences.* Nat Rev Genet, 2008. **9**(3): p. 204-17.

115. Stephens, P.J., et al., *Massive genomic rearrangement acquired in a single catastrophic event during cancer development.* Cell, 2011. **144**(1): p. 27-40.

116. Nik-Zainal, S., et al., *Mutational processes molding the genomes of 21 breast cancers.* Cell, 2012. **149**(5): p. 979-93.

117. Liu, P., et al., *Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements.* Cell, 2011. **146**(6): p. 889-903.

118. Kloosterman, W.P., et al., *Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline.* Hum Mol Genet, 2011. **20**(10): p. 1916-24.

119. Hastings, P.J., G. Ira, and J.R. Lupski, *A microhomology-mediated break-induced replication model for the origin of human copy number variation.* PLoS Genet, 2009. **5**(1): p. e1000327.

120. Allen, C., et al., *More forks on the road to replication stress recovery.* J Mol Cell Biol, 2011. **3**(1): p. 4-12.

121. Zhu, Y., et al., *Single-base deletion induced by benzo[a]pyrene diol epoxide at the adenine phosphoribosyltransferase locus in human fibrosarcoma cell lines.* Mutat Res, 1994. **321**(1-2): p. 73-9.

122. Stoffregen, E.P., et al., *An autosomal locus that controls chromosome-wide replication timing and mono-allelic expression.* Hum Mol Genet, 2011. **20**(12): p. 2366-78.

123. Donley, N., et al., *Asynchronous Replication, Mono-allelic Expression, and Long Range cis-effects of ASAR6.* PLoS Genet, 2013(In Press).

124. Chow, J.C., et al., *Silencing of the mammalian X chromosome.* Annu Rev Genomics Hum Genet, 2005. **6**: p. 69-92.

125. Panning, B., *X-chromosome inactivation: the molecular basis of silencing.* J Biol, 2008. **7**(8): p. 30.

126. Lucchesi, J.C., W.G. Kelly, and B. Panning, *Chromatin remodeling in dosage compensation.* Annu Rev Genet, 2005. **39**: p. 615-51.

127. Gomez, M. and N. Brockdorff, *Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage.* Proc Natl Acad Sci U S A, 2004. **101**(18): p. 6923-8.

128. Cohen, S.M., et al., *Same origins of DNA replication function on the active and inactive human X chromosomes.* J Cell Biochem, 2003. **88**(5): p. 923-31.

129. Morishima, A., M.M. Grumbach, and J.H. Taylor, *Asynchronous duplication of human chromosomes and the origin of sex chromatin.* Proc Natl Acad Sci U S A, 1962. **48**: p. 756-63.

130. Marahrens, Y., et al., *Xist-deficient mice are defective in dosage compensation but not spermatogenesis.* Genes Dev, 1997. **11**(2): p. 156-66.

131. Panning, B., J. Dausman, and R. Jaenisch, *X chromosome inactivation is mediated by Xist RNA stabilization.* Cell, 1997. **90**(5): p. 907-16.

132. Harnden, D.G., *Nuclear sex in triploid XXY human cells.* Lancet, 1961. **2**(7200): p. 488.

133. Wutz, A. and J. Gribnau, *X inactivation Xplained.* Curr Opin Genet Dev, 2007. **17**(5): p. 387-93.

134. Morey, C. and P. Avner, *Genetics and epigenetics of the X chromosome.* Ann N Y Acad Sci, 2010. **1214**: p. E18-33.

135. Chow, J. and E. Heard, *X inactivation and the complexities of silencing a sex chromosome.* Curr Opin Cell Biol, 2009. **21**(3): p. 359-66.

136. Brown, C.J., et al., *Localization of the X inactivation centre on the human X chromosome in Xq13.* Nature, 1991. **349**(6304): p. 82-4.

137. Lafreniere, R.G., et al., *Physical mapping of 60 DNA markers in the p21.1----q21.3 region of the human X chromosome.* Genomics, 1991. **11**(2): p. 352-63.

138. Lee, J.T., *The X as model for RNA's niche in epigenomic regulation.* Cold Spring Harb Perspect Biol, 2010. **2**(9): p. a003749.

139. Ng, K., et al., *Xist and the order of silencing.* EMBO Rep, 2007. **8**(1): p. 34-9.

140. Jeon, Y. and J.T. Lee, *YY1 tethers Xist RNA to the inactive X nucleation center.* Cell, 2011. **146**(1): p. 119-33.

141. Thorvaldsen, J.L., J.R. Weaver, and M.S. Bartolomei, *A YY1 bridge for X inactivation.* Cell, 2011. **146**(1): p. 11-3.

142. Senner, C.E. and N. Brockdorff, *Xist gene regulation at the onset of X inactivation.* Curr Opin Genet Dev, 2009. **19**(2): p. 122-6.

143. Carrel, L. and H.F. Willard, *X-inactivation profile reveals extensive variability in X-linked gene expression in females.* Nature, 2005. **434**(7031): p. 400-4.

144. Carrel, L., et al., *Genomic environment predicts expression patterns on the human inactive X chromosome.* PLoS Genet, 2006. **2**(9): p. e151.

145. Berletch, J.B., et al., *Genes that escape from X inactivation.* Hum Genet, 2011. **130**(2): p. 237-45.

146. Kazazian, H.H., Jr. and J.V. Moran, *The impact of L1 retrotransposons on the human genome.* Nat Genet, 1998. **19**(1): p. 19-24.

147. Bailey, J.A., et al., *Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6634-9.

148. Lyon, M.F., *X-chromosome inactivation: a repeat hypothesis.* Cytogenet Cell Genet, 1998. **80**(1-4): p. 133-7.

149. Clemson, C.M., et al., *XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure.* J Cell Biol, 1996. **132**(3): p. 259-75.

150. Zhang, L.F., K.D. Huynh, and J.T. Lee, *Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing.* Cell, 2007. **129**(4): p. 693-706.

151. Chaumeil, J., et al., *A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced.* Genes Dev, 2006. **20**(16): p. 2223-37.

152. Minks, J. and C.J. Brown, *Getting to the center of X-chromosome inactivation: the role of transgenes.* Biochem Cell Biol, 2009. **87**(5): p. 759-66.

153. Lee, J.T., et al., *A 450 kb transgene displays properties of the mammalian X-inactivation center.* Cell, 1996. **86**(1): p. 83-94.

154. Lee, J.T. and R. Jaenisch, *Long-range cis effects of ectopic X-inactivation centres on a mouse autosome.* Nature, 1997. **386**(6622): p. 275-9.

155. Herzing, L.B., et al., *Xist has properties of the X-chromosome inactivation centre.* Nature, 1997. **386**(6622): p. 272-5.

156. Heard, E., et al., *Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies.* Mol Cell Biol, 1999. **19**(4): p. 3156-66.

157. Hall, L.L., et al., *An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells.* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8677-82.

158. Chow, J.C., et al., *Ectopic XIST transcripts in human somatic cells show variable expression and localization.* Cytogenet Genome Res, 2002. **99**(1-4): p. 92-8.

159. Plath, K., et al., *Role of histone H3 lysine 27 methylation in X inactivation.* Science, 2003. **300**(5616): p. 131-5.

160. Chow, J.C., et al., *Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10104-9.

161. Zakharova, I.S., A.I. Shevchenko, and S.M. Zakian, *Monoallelic gene expression in mammals.* Chromosoma, 2009. **118**(3): p. 279-90.

162. Luedi, P.P., et al., *Computational and experimental identification of novel human imprinted genes.* Genome Res, 2007. **17**(12): p. 1723-30.

163. Gimelbrant, A., et al., *Widespread monoallelic expression on human autosomes.* Science, 2007. **318**(5853): p. 1136-40.

164. Chess, A., *Mechanisms and consequences of widespread random monoallelic expression.* Nat Rev Genet, 2012. **13**(6): p. 421-8.

165. Gimelbrant, A.A., et al., *Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells.* J Biol Chem, 2005. **280**(2): p. 1354-9.

166. Schlesinger, S., et al., *Allelic inactivation of rDNA loci.* Genes Dev, 2009. **23**(20): p. 2437-47.

167. Goldmit, M. and Y. Bergman, *Monoallelic gene expression: a repertoire of recurrent themes.* Immunol Rev, 2004. **200**: p. 197-214.

168. Allen, E., et al., *High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes.* Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9940-5.

169. Gimelbrant, A.A. and A. Chess, *An epigenetic state associated with areas of gene duplication.* Genome Res, 2006. **16**(6): p. 723-9.

170. Ensminger, A.W. and A. Chess, *Coordinated replication timing of monoallelically expressed genes along human autosomes.* Hum Mol Genet, 2004. **13**(6): p. 651-8.

171. Singh, N., et al., *Coordination of the random asynchronous replication of autosomal loci.* Nat Genet, 2003. **33**(3): p. 339-41.

172. Dutta, D., et al., *Asynchronous replication and autosome-pair non-equivalence in human embryonic stem cells.* PLoS One, 2009. **4**(3): p. e4970.

173. Bix, M. and R.M. Locksley, *Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells.* Science, 1998. **281**(5381): p. 1352-4.

174. Yeo, G.W., et al., *An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells.* Nat Struct Mol Biol, 2009. **16**(2): p. 130-7.

175. Brockdorff, N., et al., *The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus.* Cell, 1992. **71**(3): p. 515-26.

176. Gartler, S.M., et al., *The timing of XIST replication: dominance of the domain.* Hum Mol Genet, 1999. **8**(6): p. 1085-9.

177. Csankovszki, G., A. Nagy, and R. Jaenisch, *Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation.* J Cell Biol, 2001. **153**(4): p. 773-84.

178. Consortium, E., *A user's guide to the encyclopedia of DNA elements (ENCODE).* PLoS Biol, 2011. **9**(4): p. e1001046.

179. Valen, E., et al., *Genome-wide detection and analysis of hippocampus core promoters using DeepCAGE.* Genome Res, 2009. **19**(2): p. 255-65.

180. Hall, L.L. and J.B. Lawrence, *XIST RNA and architecture of the inactive X chromosome: implications for the repeat genome.* Cold Spring Harb Symp Quant Biol, 2010. **75**: p. 345-56.

181. Russell, L.B., *Mammalian X-chromosome action: inactivation limited in spread and region of origin.* Science, 1963. **140**(3570): p. 976-8.

182. Sharp, A.J., et al., *Molecular and cytogenetic analysis of the spreading of X inactivation in X;autosome translocations.* Hum Mol Genet, 2002. **11**(25): p. 3145-56.

183. White, W.M., et al., *The spreading of X inactivation into autosomal material of an x;autosome translocation: evidence for a difference between autosomal and X-chromosomal DNA.* Am J Hum Genet, 1998. **63**(1): p. 20-8.

184. Payer, B. and J.T. Lee, *X chromosome dosage compensation: how mammals keep the balance.* Annu Rev Genet, 2008. **42**: p. 733-72.

185. Rasmussen, T.P., et al., *Expression of Xist RNA is sufficient to initiate macrochromatin body formation.* Chromosoma, 2001. **110**(6): p. 411-20.

186. Marinelli, L.J., G.F. Hatfull, and M. Piuri, *Recombineering: A powerful tool for modification of bacteriophage genomes.* Bacteriophage, 2012. **2**(1): p. 5-14.

187. Yaffe, D. and O. Saxel, *Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle.* Nature, 1977. **270**(5639): p. 725-7.

188. Dostie, J. and W.A. Bickmore, *Chromosome organization in the nucleus - charting new territory across the Hi-Cs.* Curr Opin Genet Dev, 2012. **22**(2): p. 125-31.

189. Schwaiger, M. and D. Schubeler, *A question of timing: emerging links between transcription and replication.* Curr Opin Genet Dev, 2006. **16**(2): p. 177-83.

190. De, S. and F. Michor, *DNA replication timing and long-range DNA interactions predict mutational landscapes of cancer genomes.* Nat Biotechnol, 2011. **29**(12): p. 1103-8.

191. Warming, S., et al., *Simple and highly efficient BAC recombineering using galK selection.* Nucleic Acids Res, 2005. **33**(4): p. e36.

192. Helm, S., *Cancer Cytogenetics*. 1995, New York: Wiley-Liss.

193. Smith, L. and M. Thayer, *Chromosome replicating timing combined with fluorescent in situ hybridization.* J Vis Exp, 2012(70): p. e4400.

194. Keim, M., R.S. Williams, and A.J. Harwood, *An inverse PCR technique to rapidly isolate the flanking DNA of dictyostelium insertion mutants.* Mol Biotechnol, 2004. **26**(3): p. 221-4.

195. Harkey, M.A., et al., *Multiarm high-throughput integration site detection: limitations of LAM-PCR technology and optimization for clonal analysis.* Stem Cells Dev, 2007. **16**(3): p. 381-92.

196. Branda, C.S. and S.M. Dymecki, *Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice.* Dev Cell, 2004. **6**(1): p. 7-28.

197. *The data was obtained from the Wellcome Trust Sanger Institute Cancer Genome Project web site, http://www.sanger.ac.uk/genetics/CGP*.

198. Ross, M.T., et al., *The DNA sequence of the human X chromosome.* Nature, 2005. **434**(7031): p. 325-37.

199. Mostoslavsky, R., et al., *Asynchronous replication and allelic exclusion in the immune system.* Nature, 2001. **414**(6860): p. 221-5.

200. Goodwin, E. and J. Meyne, *Strand-specific FISH reveals orientation of chromosome 18 alphoid DNA.* Cytogenet Cell Genet, 1993. **63**(2): p. 126-7.

201. Zwemer, L.M., et al., *Autosomal monoallelic expression in the mouse.* Genome Biol, 2012. **13**(2): p. R10.

202. Meihlac, M., et al., *Amanitin binding to calf thymus RNA polymerase B.* FEBS Lett, 1970. **9**(5): p. 258-260.

203. Luedi, P.P., A.J. Hartemink, and R.L. Jirtle, *Genome-wide prediction of imprinted murine genes.* Genome Res, 2005. **15**(6): p. 875-84.

204. Hansen, R.S., et al., *Sequencing newly replicated DNA reveals widespread plasticity in human replication timing.* Proc Natl Acad Sci U S A, 2009. **107**(1): p. 139-44.

205. Breger, K.S., et al., *Ionizing radiation induces frequent translocations with delayed replication and condensation.* Cancer Research, 2004. **64**: p. 8231-8238.

206. Smith, L., A. Plug, and M. Thayer, *Delayed Replication Timing Leads to Delayed Mitotic Chromosome Condensation and Chromosomal Instability of Chromosome Translocations.* Proc Natl Acad Sci U S A, 2001. **98**: p. 13300-13305.

207. Stoffregen, E.P., et al., *An autosomal locus that controls chromosome-wide replication timing and mono-allelic expression.* Hum Mol Genet, 2011. **20**: p. 2366-2378.

208. Marshall, O.J., et al., *Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution.* Am J Hum Genet, 2008. **82**(2): p. 261-82.

209. Sullivan, B.A., M.D. Blower, and G.H. Karpen, *Determining centromere identity: cyclical stories and forking paths.* Nat Rev Genet, 2001. **2**(8): p. 584-96.

210. Shang, W.H., et al., *Chromosome Engineering Allows the Efficient Isolation of Vertebrate Neocentromeres.* Dev Cell, 2013.

211. Zhu, Y., et al., *Single-base deletion induced by benzo[a]pyrene diol epoxide at the adenine phosphoribosyltransferase locus in human fibrosarcoma cell lines.* Mutat Res, 1994. **321**(1-2): p. 73-9.

212. Trask, B. and D. Pinkel, *Flow cytometry*. Methods in Cell Biology, ed. H.A. Crissman and Z. Darzynkiewica. Vol. 33. 1990, New York: Academic Press.

213. Smith, L. and M. Thayer, *Chromosome replication timing combined with fluorescent in situ hybridization.* JOVE, 2012. **in press**.

214. Pruitt, K.D., T. Tatusova, and D.R. Maglott, *NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins.* Nucleic Acids Res, 2005. **33**(Database issue): p. D501-4.

215. Jurka, J., et al., *Repbase Update, a database of eukaryotic repetitive elements.* Cytogenet Genome Res, 2005. **110**(1-4): p. 462-7.

216. Stoffregen, E.P., *Genetic Analysis Of Chromosome Replication Timing: An Autosomal Locus That Controls Chromosome-Wide Replication Timing And Mono-Allelic Expression*, in *Molecular and Medical Genetics Dept.* 2011, Oregon Health & Science University: Portland. p. 253.

217. Vallot, C., et al., *XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells.* Nat Genet, 2013. **45**(3): p. 239-41.

218. Podlaha, O., et al., *Evolution of the cancer genome.* Trends Genet, 2012. **28**(4): p. 155-63.

219. Diaz Perez, S.V., et al., *Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin.* Hum Mol Genet, 2012. **21**(4): p. 751-64.

220. Lengner, C.J., et al., *Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations.* Cell, 2010. **141**(5): p. 872-83.

221. Sauer, B. and J. McDermott, *DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages.* Nucleic Acids Res, 2004. **32**(20): p. 6086-95.

222. Anastassiadis, K., et al., *Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice.* Dis Model Mech, 2009. **2**(9-10): p. 508-15.

223. Masai, H., et al., *Eukaryotic chromosome DNA replication: where, when, and how?* Annu Rev Biochem, 2010. **79**: p. 89-130.

224. Teer, J.K. and A. Dutta, *Regulation of S phase.* Results Probl Cell Differ, 2006. **42**: p. 31-63.

225. Ryba, T., et al., *Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia.* Genome Res, 2012. **22**(10): p. 1833-44.

226. Hiratani, I., et al., *Global reorganization of replication domains during embryonic stem cell differentiation.* PLoS Biol, 2008. **6**(10): p. e245.

227. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, 2009. **326**(5950): p. 289-93.

228. Prendergast, J.G., et al., *Chromatin structure and evolution in the human genome.* BMC Evol Biol, 2007. **7**: p. 72.

229. Hellmann, I., et al., *Why do human diversity levels vary at a megabase scale?* Genome Res, 2005. **15**(9): p. 1222-31.

230. Green, P., et al., *Transcription-associated mutational asymmetry in mammalian evolution.* Nat Genet, 2003. **33**(4): p. 514-7.

231. Wolfe, K.H., P.M. Sharp, and W.H. Li, *Mutation rates differ among regions of the mammalian genome.* Nature, 1989. **337**(6204): p. 283-5.

232. Lang, G.I. and A.W. Murray, *Mutation rates across budding yeast chromosome VI are correlated with replication timing.* Genome Biol Evol, 2011. **3**: p. 799-811.

233. Stamatoyannopoulos, J.A., et al., *Human mutation rate associated with DNA replication timing.* Nat Genet, 2009. **41**(4): p. 393-5.

234. Chen, C.L., et al., *Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes.* Genome Res, 2010. **20**(4): p. 447-57.

235. Cui, P., et al., *Distinct contributions of replication and transcription to mutation rate variation of human genomes.* Genomics Proteomics Bioinformatics, 2012. **10**(1): p. 4-10.

236. Cardoso-Moreira, M.M. and M. Long, *Mutational bias shaping fly copy number variation: implications for genome evolution.* Trends Genet, 2010. **26**(6): p. 243-7.

237. Cardoso-Moreira, M., et al., *Drosophila duplication hotspots are associated with late-replicating regions of the genome.* PLoS Genet, 2011. **7**(11): p. e1002340.

238. Chuang, J.H. and H. Li, *Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome.* PLoS Biol, 2004. **2**(2): p. E29.

239. Huang, S., et al., *More single-nucleotide mutations surround small insertions than small deletions in primates.* Hum Mutat, 2012. **33**(7): p. 1099-106.

240. Watanabe, Y. and M. Maekawa, *Spatiotemporal regulation of DNA replication in the human genome and its association with genomic instability and disease.* Curr Med Chem, 2010. **17**(3): p. 222-33.

241. Watanabe, Y., T. Ikemura, and H. Sugimura, *Amplicons on human chromosome 11q are located in the early/late-switch regions of replication timing.* Genomics, 2004. **84**(5): p. 796-805.

242. Taylor, J.S. and J. Raes, *Duplication and divergence: the evolution of new genes and old ideas.* Annu Rev Genet, 2004. **38**: p. 615-43.

243. Farkash-Amar, S., et al., *Global organization of replication time zones of the mouse genome.* Genome Res, 2008. **18**(10): p. 1562-70.

244. Zhang, L. and W.H. Li, *Mammalian housekeeping genes evolve more slowly than tissue-specific genes.* Mol Biol Evol, 2004. **21**(2): p. 236-9.

245. Woo, Y.H. and W.H. Li, *DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes.* Nat Commun, 2012. **3**: p. 1004.

246. Janoueix-Lerosey, I., et al., *Preferential occurrence of chromosome breakpoints within early replicating regions in neuroblastoma.* Cell Cycle, 2005. **4**(12): p. 1842-6.

247. Drier, Y., et al., *Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability.* Genome Res, 2012.

248. Watanabe, Y., et al., *Relationships between replication timing and GC content of cancer-related genes on human chromosomes 11q and 21q.* Gene, 2009. **433**(1-2): p. 26-31.

249. Soutoglou, E. and T. Misteli, *On the contribution of spatial genome organization to cancerous chromosome translocations.* J Natl Cancer Inst Monogr, 2008(39): p. 16-9.

250. Engreitz, J.M., V. Agarwala, and L.A. Mirny, *Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease.* PLoS One, 2012. **7**(9): p. e44196.

251. Fudenberg, G., et al., *High order chromatin architecture shapes the landscape of chromosomal alterations in cancer.* Nat Biotechnol, 2011. **29**(12): p. 1109-13.

252. Zhang, Y., et al., *Spatial organization of the mouse genome and its role in recurrent chromosomal translocations.* Cell, 2012. **148**(5): p. 908-21.

253. Meaburn, K.J., T. Misteli, and E. Soutoglou, *Spatial genome organization in the formation of chromosomal translocations.* Semin Cancer Biol, 2007. **17**(1): p. 80-90.

254. Lukasova, E., et al., *Localisation and distance between ABL and BCR genes in interphase nuclei of bone marrow cells of control donors and patients with chronic myeloid leukaemia.* Hum Genet, 1997. **100**(5-6): p. 525-35.

255. Hecht, F. and G.R. Sutherland, *Detection of fragile sites on human chromosomes.* Clin Genet, 1985. **28**(1): p. 95-6.

256. Laird, C., et al., *Fragile sites in human chromosomes as regions of late-replicating DNA.* Trends in Genetics, 1987. **3**(10): p. 274-281.

257. Arlt, M.F., A.M. Casper, and T.W. Glover, *Common fragile sites.* Cytogenet Genome Res, 2003. **100**(1-4): p. 92-100.

258. Debatisse, M., E. El Achkar, and B. Dutrillaux, *Common fragile sites nested at the interfaces of early and late-replicating chromosome bands: cis acting components of the G2/M checkpoint?* Cell Cycle, 2006. **5**(6): p. 578-81.

259. El Achkar, E., et al., *Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18069-74.

260. Ozeri-Galai, E., A.C. Bester, and B. Kerem, *The complex basis underlying common fragile site instability in cancer.* Trends Genet, 2012. **28**(6): p. 295-302.

261. Amiel, A., et al., *Replication pattern of the p53 and 21q22 loci in the premalignant and malignant stages of carcinoma of the cervix.* Cancer, 1998. **83**(9): p. 1966-71.

262. Amiel, A., et al., *Temporal differences in replication timing of homologous loci in malignant cells derived from CML and lymphoma patients.* Genes Chromosomes Cancer, 1998. **22**(3): p. 225-31.

263. Amiel, A., et al., *Replication pattern in cancer: asynchronous replication in multiple myeloma and in monoclonal gammopathy.* Cancer Genet Cytogenet, 1999. **108**(1): p. 32-7.

264. Cytron, S., et al., *Epigenetic analyses in blood cells of men suspected of prostate cancer predict the outcome of biopsy better than serum PSA levels.* Clin Epigenetics, 2011. **2**(2): p. 383-388.

265. Reish, O., et al., *Modified allelic replication in lymphocytes of patients with neurofibromatosis type 1.* Cancer Genet Cytogenet, 2003. **143**(2): p. 133-9.

266. Litmanovitch, T., et al., *Asynchronous replication of homologous alpha-satellite DNA loci in man is associated with nondisjunction.* Cytogenet Cell Genet, 1998. **81**(1): p. 26-35.

267. Dotan, Z.A., et al., *Altered mode of allelic replication accompanied by aneuploidy in peripheral blood lymphocytes of prostate cancer patients.* Int J Cancer, 2004. **111**(1): p. 60-6.

268. Grinberg-Rashi, H., et al., *Replication timing aberrations and aneuploidy in peripheral blood lymphocytes of breast cancer patients.* Neoplasia, 2010. **12**(8): p. 668-74.

269. Fritz, A.J., et al., *Alterations in replication timing of cancer related genes in malignant human breast cancer cells.* J Cell Biochem, 2012.

270. Bras, A., et al., *Asynchronous DNA replication detected by fluorescence in situ hybridisation as a possible indicator of genetic damage in human lymphocytes.* Oncol Rep, 2008. **19**(2): p. 369-75.

271. Dotan, Z.A., et al., *Aberrant allele-specific replication, independent of parental origin, in blood cells of cancer patients.* BMC Cancer, 2008. **8**: p. 390.

272. Nagler, A., et al., *The aberrant asynchronous replication - characterizing lymphocytes of cancer patients - is erased following stem cell transplantation.* BMC Cancer, 2010. **10**: p. 230.

273. Adolph, S., H. Hameister, and C.L. Schildkraut, *Molecular analysis of the aberrant replication banding pattern on chromosome 15 in murine T-cell lymphomas.* Chromosoma, 1992. **101**(7): p. 388-98.

274. Farkash-Amar, S., et al., *Systematic Determination of Replication Activity Type Highlights Interconnections between Replication, Chromatin Structure and Nuclear Localization.* PLoS One, 2012. **7**(11): p. e48986.

275. Gondor, A. and R. Ohlsson, *Replication timing and epigenetic reprogramming of gene expression: a two-way relationship?* Nat Rev Genet, 2009. **10**(4): p. 269-76.

276. Donaldson, A.D., et al., *CLB5-dependent activation of late replication origins in S. cerevisiae.* Mol Cell, 1998. **2**(2): p. 173-82.

277. Shechter, D., V. Costanzo, and J. Gautier, *ATR and ATM regulate the timing of DNA replication origin firing.* Nat Cell Biol, 2004. **6**(7): p. 648-55.

278. Zhou, J., A.R. Snyder, and P.M. Lieberman, *Epstein-Barr virus episome stability is coupled to a delay in replication timing.* J Virol, 2009. **83**(5): p. 2154-62.

279. Cosgrove, A.J., C.A. Nieduszynski, and A.D. Donaldson, *Ku complex controls the replication time of DNA in telomere regions.* Genes Dev, 2002. **16**(19): p. 2485-90.

280. Watanabe, Y., et al., *p53-dependent change in replication timing of the human genome.* Biochem Biophys Res Commun, 2007. **364**(2): p. 289-93.

281. Hayashi, M.T., et al., *The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus.* Nat Cell Biol, 2009. **11**(3): p. 357-62.

282. Black, J.C., et al., *Conserved antagonism between JMJD2A/KDM4A and HP1gamma during cell cycle progression.* Mol Cell, 2010. **40**(5): p. 736-48.

283. Jorgensen, H.F., et al., *The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells.* Genome Biol, 2007. **8**(8): p. R169.

284. Stevenson, J.B. and D.E. Gottschling, *Telomeric chromatin modulates replication timing near chromosome ends.* Genes Dev, 1999. **13**(2): p. 146-51.

285. Schwaiger, M., et al., *Heterochromatin protein 1 (HP1) modulates replication timing of the Drosophila genome.* Genome Res, 2010. **20**(6): p. 771-80.

286. Wu, R., P.B. Singh, and D.M. Gilbert, *Uncoupling global and fine-tuning replication timing determinants for mouse pericentric heterochromatin.* J Cell Biol, 2006. **174**(2): p. 185-94.

287. Hansen, R.S., et al., *Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant.* Hum Mol Genet, 2000. **9**(18): p. 2575-87.

288. Zhou, J., et al., *Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells.* Mol Cell Biol, 2002. **22**(13): p. 4876-89.

289. Goren, A., et al., *DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin.* Genes Dev, 2008. **22**(10): p. 1319-24.

290. Aparicio, J.G., et al., *The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae.* Mol Cell Biol, 2004. **24**(11): p. 4769-80.

291. Vogelauer, M., et al., *Histone acetylation regulates the time of replication origin firing.* Mol Cell, 2002. **10**(5): p. 1223-33.

292. Zappulla, D.C., R. Sternglanz, and J. Leatherwood, *Control of replication timing by a transcriptional silencer.* Curr Biol, 2002. **12**(11): p. 869-75.

293. Aggarwal, B.D. and B.R. Calvi, *Chromatin regulates origin activity in Drosophila follicle cells.* Nature, 2004. **430**(6997): p. 372-6.

294. Zhou, J., et al., *Cell cycle regulation of chromatin at an origin of DNA replication.* EMBO J, 2005. **24**(7): p. 1406-17.

295. Jablonka, E., et al., *DNA hypomethylation causes an increase in DNase-I sensitivity and an advance in the time of replication of the entire inactive X chromosome.* Chromosoma, 1985. **93**(2): p. 152-6.

296. Aladjem, M.I., et al., *Participation of the human beta-globin locus control region in initiation of DNA replication.* Science, 1995. **270**(5237): p. 815-9.

297. Zhang, J., et al., *Establishment of transcriptional competence in early and late S phase.* Nature, 2002. **420**(6912): p. 198-202.

298. Lande-Diner, L., J. Zhang, and H. Cedar, *Shifts in replication timing actively affect histone acetylation during nucleosome reassembly.* Mol Cell, 2009. **34**(6): p. 767-74.

299. Zhong, Q., et al., *Perturbation of TSG101 protein affects cell cycle progression.* Cancer Res, 1998. **58**(13): p. 2699-702.

300. Rountree, M.R., K.E. Bachman, and S.B. Baylin, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.* Nat Genet, 2000. **25**(3): p. 269-77.

301. Tatematsu, K.I., T. Yamazaki, and F. Ishikawa, *MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase.* Genes Cells, 2000. **5**(8): p. 677-88.

302. Gribnau, J., et al., *Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization.* Genes Dev, 2003. **17**(6): p. 759-73.

303. Lock, L.F., N. Takagi, and G.R. Martin, *Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation.* Cell, 1987. **48**(1): p. 39-46.

304. Bridger, J.M., et al., *Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts.* Curr Biol, 2000. **10**(3): p. 149-52.

305. Sadoni, N., et al., *Stable chromosomal units determine the spatial and temporal organization of DNA replication.* J Cell Sci, 2004. **117**(Pt 22): p. 5353-65.

306. Korenstein-Ilan, A., et al., *Terahertz radiation increases genomic instability in human lymphocytes.* Radiat Res, 2008. **170**(2): p. 224-34.

307. Bergstrom, R., et al., *CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain.* Cell Cycle, 2007. **6**(4): p. 450-4.

308. Bianchi, A. and D. Shore, *Early replication of short telomeres in budding yeast.* Cell, 2007. **128**(6): p. 1051-62.

309. Lin, C.M., et al., *Dynamic alterations of replication timing in mammalian cells.* Curr Biol, 2003. **13**(12): p. 1019-28.

310. Hayashi, M., et al., *Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast.* EMBO J, 2007. **26**(5): p. 1327-39.

311. Amiel, A., et al., *Replication status as a possible marker for genomic instability in cells originating from genotypes with balanced rearrangements.* Chromosome Res, 2001. **9**(8): p. 611-6.

312. Sun, Y., et al., *Expression and replication timing patterns of wildtype and translocated BCL2 genes.* Genomics, 2001. **73**(2): p. 161-70.

313. Boumil, R.M. and J.T. Lee, *Forty years of decoding the silence in X-chromosome inactivation.* Hum Mol Genet, 2001. **10**(20): p. 2225-32.

314. Eyre-Walker, A. and P.D. Keightley, *The distribution of fitness effects of new mutations.* Nat Rev Genet, 2007. **8**(8): p. 610-8.