# GENETIC ANALYSIS OF CHROMOSOME REPLICATION TIMING:

## AN AUTOSOMAL LOCUS THAT CONTROLS CHROMOSOME-WIDE REPLICATION TIMING AND MONO-ALLELIC EXPRESSION

by

Eric Paul Stoffregen

A DISSERTATION

Presented to the Molecular and Medical Genetics Department and

the Oregon Health & Science University School of Medicine

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

July 2011

School of Medicine

Oregon Health & Science University

---

CERTIFICATE OF APPROVAL

---

This is to certify that the Ph.D. dissertation of

Eric Paul Stoffregen

has been approved

---

Mathew Thayer, Ph.D., Mentor

---

Rosalie Sears, Ph.D., Committee member

---

R. Michael Liskay, Ph.D., Committee member

---

Soren Impey, Ph.D., Committee member

---

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

# TABLE OF CONTENTS

# LIST OF FIGURES AND TABLES

# LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| AAV | – | adeno-associated virus |
| Air | – | antisense to Igf2r |
| APRT | – | adenine phosphoribosyltransferase |
| *ASAR6* | – | asynchronous replication and autosomal RNA on 6 |
| ATM | – | ataxia-telangiectasia mutated |
| ATR | – | ataxia-telangiectasia and RAD3-related |
| BAC | – | bacterial artificial chromosome |
| Blas | – | blasticidin |
| bp | – | base pair |
| BrdU | – | bromodeoxyuridine |
| C | – | Celsius |
| Cdk | – | cyclin-dependent kinase |
| cDNA | – | complementary DNA |
| ChIP | – | chromatin immunoprecipitation |
| CIN | – | chromosome instability |
| $CO_2$ | – | carbon dioxide |
| Cre | – | causes recombination |
| DAPI | – | 4'-6-Diamidino-2-phenylindole |
| DD | – | double-double |
| DMC | – | delay in mitotic chromosome condensation |
| DNA | – | deoxyribonucleic acid |

| | | |
|---|---|---|
| dNTP | – | deoxyribonucleotide triphosphate |
| Dre | – | D6 recombinase |
| DRT | – | delay in replication timing |
| DSB | – | double strand break |
| EZH2 | – | enhancer of zeste homolog 2 |
| *FHL5* | – | four and a half LIM domains 5 |
| *FHL5OST* | – | *FHL5* opposite strand transcript |
| FISH | – | fluorescence *in situ* hybridization |
| FITC | – | fluorescein isothiocyanate |
| Flp | – | flippase |
| FRT | – | flippase recognition target |
| *FUT9* | – | fucosyltransferase 9 |
| G-band | – | Giemsa band |
| h | – | hour |
| HR | – | homologous recombination |
| Hyg | – | hygromycin |
| ICE | – | imprint control element |
| Igf2 | – | insulin-like growth factor 2 |
| Igf2r | – | insulin-like growth factor 2 receptor |
| IR | – | ionizing radiation |
| ITR | – | inverted terminal repeats |
| Kb | – | kilobase |
| LAM-PCR | – | linear amplification PCR |
| LCR | – | locus control region |
| LINE-1 | – | long interspersed nuclear element-1 |

| | | |
|---|---|---|
| LOH | – | loss of heterozygosity |
| *lox*P | – | locus of cross-over (x-over) P1 |
| LTR | – | long terminal repeat |
| M | – | molar |
| *MANEA* | – | mannosidase, endo-alpha |
| Mb | – | mega-base |
| mg | – | milligram |
| MIN | – | microsatellite instability |
| min | – | minute |
| ml | – | milliliter |
| mM | – | millimolar |
| MMR | – | mismatch repair |
| MOI | – | multiplicity of infection |
| NCBI | – | National Center for Biotechnology Information |
| ncRNA | – | non-coding RNA |
| Neo | – | neomycin |
| NER | – | nucleotide excision repair |
| ng | – | nanogram |
| NHEJ | – | non-homologous end joining |
| ORC | – | origin recognition complex |
| P-clone | – | parental clone |
| PcG | – | Polycomb group |
| PCR | – | polymerase chain reaction |
| PRC2 | – | Polycomb repressive complex 2 |
| pre-RC | – | pre-replication complex |

| | | |
|---|---|---|
| rAAV | – | recombinant adeno-associated virus |
| R-clone | – | recombinant clone |
| RIP | – | RNA immunoprecipitation |
| RNA | – | ribonucleic acid |
| RNAi | – | RNA interference |
| rox | – | region of cross-over (x-over) |
| RT-PCR | – | reverse transcription-PCR |
| s | – | second |
| SD | – | single-double |
| SNP | – | single nucleotide polymorphism |
| SS | – | single-single |
| SUZ12 | – | suppressor of zeste 12 homolog |
| TIN | – | translocation instability |
| U | – | unit |
| µg | – | microgram |
| µl | – | microliter |
| µM | – | micromolar |
| UV | – | ultraviolet |
| Xa | – | active X chromosome |
| XCI | – | X-chromosome inactivation |
| Xi | – | inactive X chromosome |
| Xic | – | X inactivation center |
| *Xist* | – | X-inactive specific transcript |
| *Xite* | – | X-inactivation intergenic transcription elements |

# ACKNOWLEDGEMENTS

I owe many people a debt of gratitude for the very existence of this dissertation; without their help and support along the way it would not be my name on the cover. First and foremost, I have to thank my mentor, Dr. Mathew Thayer. Matt is an exceptional scientist and I could not have asked for a better mentor. Thank you for always being available, helpful, and patient. Thank you for cultivating a great laboratory working environment. This made six years of graduate school go by much faster and more pleasantly than I could have expected.

None of the work here would have been possible without Leslie Smith's unparalleled cytogenetics expertise. Thank you Leslie for the invaluable work you do on the microscope, for keeping the laboratory running smoothly, and for being such a great person to work with. I would like to thank my fellow graduate student in the lab, Nathan Donley, for his assistance with experiments for this thesis and for giving me someone to bounce ideas off of.

I would like to Dr. Dan Stauffer, Dr. Dan Pankratz, Dr. Bill Chang, and Dr. Sarah Smolik who were all extremely helpful in showing me the ropes when I joined the laboratory and in providing guidance and advice along the way. Thank you to Dr. Kevin Breger and other past Thayer laboratory members who got this project started and made my work possible.

Thank you to the members of my dissertation advisory and exam committees, Dr. Mike Liskay, Dr. Rosie Sears, Dr. Soren Impey, and Dr. Stephen Lloyd, for their helpful suggestions and advice about my research and about this written dissertation. And thank you especially for giving up your time to do all of those things.

Thank you to Dr. David Marcey at California Lutheran University who was a great undergraduate mentor and was one of the first to tell me that I should go on to get a Ph.D. Thank you to Dr. Brian Druker for giving me my start at OHSU. I could not have dreamed up a better place to land for my first science job. Thank you to Dr. Thomas O'Hare for taking me under his wing in the Druker lab, for teaching me a lot about science and more, and for being a good friend.

Last, but certainly not least, I would like to thank my friends and family. Graduate school can be a difficult endeavor, but having good friends makes it much easier. Thank you to the many good friends on whom I have been able to rely during graduate school. Having a very supportive family was the most important thing of all. I'm fortunate to have great and supportive in-laws. Thank you to Rick and Judy Ham for the amazing amount of support you have given me over the last decade. Thank you to my brothers-in-law Tom and Dan for their support as well.

Thank you to my older brother, Kevin. I'm glad you made a more financially lucrative decision when choosing your career path so you could be so generous with us! Thank you to Rachel for being the best sister-in-law one could ask for.

Thank you to my wonderful nieces Megan and Katelyn, who were always a welcome distraction and stress-relief during graduate school.

Thank you to my parents Kim and Larry Stoffregen. Parents are generally expected to provide for and support their children, but you certainly never settled for the minimum in those regards. Thank you for being such wonderful and supportive parents and for giving me such a great example to follow. Having you just ten minutes away during graduate school was a great gift and a life-saver!

To my wife, Cindy, thank you is hardly a sufficient expression. Although I'm practically a saint (just ask me), I know it hasn't always been easy having a husband in graduate school for the past six years. Thank you for your love, support, and encouragement through it all. Even if I could have accomplished this without you (and I'm not saying I could have), I wouldn't have wanted to. Thank you to my daughter Ellie. You may have made it more difficult to stay in the laboratory and get work done, or to get enough rest when I was at home, but seeing you smile at me when I came through the door made it all worth it!

# ABSTRACT

A subset of tumor-derived chromosomal rearrangements exhibit a significant delay in replication timing (DRT) and a subsequent delay in mitotic chromosome condensation (DMC). Chromosomes with DRT/DMC are unstable; they have a 30-80-fold increased rate of new gross chromosomal rearrangements. We have used site-specific recombination-mediated chromosome engineering strategies to investigate the genetic basis of the DRT/DMC phenotype. Using these strategies, I have identified a locus on human chromosome 6 that controls the replication timing program of the entire chromosome in *cis*. Rearrangements at this locus result in a chromosome-wide delay in replication timing as well as re-activation of the previously silent alleles of linked mono-allelically expressed genes, also in *cis*. Characterization of this locus revealed the presence of a large, intergenic, mono-allelically expressed non-coding RNA, which we have named asynchronous replication and autosomal RNA on chromosome 6, or *ASAR6*. Homologous alleles typically replicate synchronously; however, mono-allelically expressed genes such as imprinted genes, allelically excluded genes and genes on the female X chromosome replicate asynchronously. The ~1 Mb region surrounding, and including, *ASAR6* replicates asynchronously in a coordinated fashion with other mono-allelically expressed genes on chromosome 6. The early replicating locus of *ASAR6* is coordinated with the late replicating alleles of other

mono-allelically expressed genes, and represents a novel finding for coordination of asynchronous replication timing. These data indicate that human autosomes contain discrete *cis*-acting loci that control chromosome-wide replication timing, mono-allelic expression, and the stability of entire chromosomes. Disruptions of these loci could contribute to the genomic instability associated with cancer and with cells exposed to ionizing radiation.

# CHAPTER ONE

## Introduction

## Preface

In the 1960s and early 1970s several publications observed morphological differences between chromosomes within the same cell. The first of these publications described the particular consequences of viral infection on chromosomes [1-2]. The chromosomes in these reports were extremely fragmented, and thus the phenotype was described as ‘pulverization’. It was later found by other groups that chromosomes with this pulverization phenotype also showed ‘incomplete condensation’ and were late in replicating their DNA [3-5]. However, little research on this chromosomal phenotype occurred over the next few decades, despite the fact that chromosome abnormalities, particularly those related to the instability and fragmentation of chromosomes, are a common occurrence in cancer cells [6-7].

After repeatedly observing an abnormal phenotype associated with certain tumor-derived chromosomal rearrangements, the Thayer laboratory discovered that these rearranged chromosomes had problems in their replication and condensation. The laboratory has extensively characterized these abnormal chromosomes, and we think this work bridges the gap between those initial observations of pulverized chromosomes and the common observation of chromosomal aberrations in cancer. The chromosomes described by the Thayer laboratory display a significant delay in replication timing (DRT), characterized by a 2-3 hour delay in the onset and completion of DNA synthesis, and a subsequent delay in mitotic chromosome condensation (DMC) [8]. These DRT/DMC chromosomes also display significant instability [8-10], which is

intriguing because DRT/DMC may provide insight into one potential model for the genomic instability seen in cancer cells and cells exposed to ionizing radiation (IR) [9].

Since the genetic basis for this phenotype was unknown, the Thayer laboratory developed a novel chromosome engineering strategy that allows us to study the initiating events of DRT/DMC by generating a set of balanced translocations that display the phenotype [10]. Accordingly, the central aim of my dissertation was to explore the genetic basis of DRT/DMC using one of the cell lines that emerged from the chromosome engineering project.

In the remaining part of Chapter 1, I will provide a review of the literature relevant to understanding the DRT/DMC phenotype and the progress we have made in describing the genetic aspects of the phenotype. In Chapter 2 and Appendix 1, which comprise the bulk of my dissertation, I will detail the identification of an autosomal locus on chromosome 6 that controls chromosome-wide replication timing and mono-allelic expression. Deletions or translocations at this locus cause the DRT/DMC phenotype. We hypothesize that these novel findings on chromosome 6 will not be unique to this chromosome; rather, every autosome has one or more similar loci that have yet to be identified. Included in Chapter 3 is the result of a proof-of-principle study to show the feasibility of using recombinant adeno-associated virus (rAAV) to engineer defined rearrangements in human chromosomes. In Chapter 4, I will offer conclusions and future directions for this project, highlighting some of the important questions that this work has led to and how those questions are being addressed.

## Genomic Instability in Cancer

One hallmark of cancer cells is their accumulation of a large number of mutations. The presence of so many genetic changes is usually attributed to an underlying genetic instability in cancer cells [6]. This instability is seen in the form of a variety of genetic insults that are common in tumors. These include simple genetic changes involving base substitution, insertion or deletion of one or a few nucleotide base pairs. Additionally, gene amplification can occur, leading to increased copy number and expression of growth-promoting oncogenes. Aneuploidy (losses or gains of whole chromosomes) is another type of genetic abnormality seen in most major cancers. Chromosome rearrangements, such as translocations, insertions, deletions and inversions, are also common genetic alterations seen in cancer [6, 11]. Stoler et al. found ~11,000 genetic alterations in carcinomas compared to normal tissues [11], which highlights the vast number of changes that can occur in cancer cells. However, large numbers of mutations in a tumor cell do not a priori define the tumor as being unstable. For that designation, one would have to determine the *rate* of mutation accumulation [6].

Several experiments have shown that cancer cells do have an increase in the rate of mutation accumulation, and can thus be defined as having genomic instability [12-13, 6]. Cancer cells with defects in DNA repair, either in nucleotide excision repair (NER) or in mismatch repair (MMR), show an increased rate of mutations involving nucleotide changes such as base substitutions or small deletions or insertions. Instability caused by defects in MMR is usually referred to as microsatellite instability (MIN) due to the common occurrence of these

alterations at microsatellite sequences [6]. An increased rate of change in the number of chromosomes, termed chromosomal instability (CIN) has also been shown in cancer cells by more than one methodology [12-13]. Dr. Bert Vogelstein's laboratory showed that colon cancer cell lines can display MIN or CIN and that they are two separate phenomena of genetic instability [13]. In a separate study, CIN was measured at a 10-100-fold increased rate in cancer cells versus normal cells [13]. While CIN describes the rate of change in the number of chromosomes, cancer cells also experience an increase in both the rate and number of structural chromosomal rearrangements (translocations, deletions, inversions) (reviewed in [14]). This increased rate of accumulation of chromosome rearrangements can be termed translocation instability (TIN) [8].

Genomic instability is also a common feature of cells exposed to IR. This phenomenon occurs *in vitro* or *in vivo* following a time delay, and these changes are persistent [15-16]. These radiation-induced genetic insults include chromosome rearrangements, gene amplifications, and gene mutations [15, 17]. IR-induced genomic instability appears analogous to the instability seen so often in tumor cells, and they may share a common mechanism [17, 9].

A number of mechanisms have been proposed to explain the acquisition of CIN (reviewed in [18, 14], but only a few of these offer any explanation for TIN. Not much is known about the process by which a cancer cell can rapidly accumulate a large number of chromosomal aberrations. One recently proposed model, termed chromothripsis, posits that one single event causes the shattering of a chromosome into tens or hundreds of pieces, which are then repaired in a

patchwork manner, leaving pieces of that chromosome strewn throughout the genome [19]. The authors propose that this model is the only way by which a chromosome could undergo such a vast number of rearrangements in so short a time. Their model is based on computer simulations and lacks any supporting data. However, we think previously published data from our laboratory provides a model that can explain the observations of a single chromosome with tens or hundreds of rearrangements [8-10]. The data from our laboratory indicates that the TIN commonly seen in cancer cells and in cells exposed to IR can begin with a single genetic change [8-10]. This single event results in a chromosome that is susceptible to DNA damage followed by chromosome rearrangement. The chromosome displays such a drastically increased rate of genetic instability that hundreds of rearrangements are possible in a very short period of time

## DNA Replication and Chromosome Condensation

Cell division is a highly regulated process in normal cells. This process, known as the cell cycle, is divided into four phases: G1, when the cell grows and prepares for DNA synthesis; S-phase, when DNA replication occurs; G2, when the cell prepares for mitosis, including proper condensation of the chromosomes; and M-phase, or mitosis, when the replicated chromosomes are separated into two daughter cells [20-21].

In eukaryotes, the faithful replication of the DNA content of each cell must occur once, and only once, each cell cycle, prior to cell division. This process has been studied extensively (for reviews see [21-23]), but for the purposes of this

dissertation introduction, I will only touch on the basics. DNA replication takes place during S-phase and begins at replication origins scattered along each chromosome. The number of active origins during each cell cycle has been estimated at 30,000-50,000 in mammalian cells [24]. At the end of mitosis, origins are recognized and bound by the origin-recognition complex (ORC). In early G1, two proteins necessary for replication initiation, Cdc6 and Cdt1, are recruited to origins and function to load the Mcm2-7 complex, which has ATPase-dependent DNA helicase activity and is thus responsible for unwinding the DNA ahead of the replication complex, onto the replication origins. This forms what is known as the pre-replication complex (pre-RC) [21].

At this point, origins are “licensed” and ready to fire, but must still be activated before DNA synthesis begins. The activation of the pre-RC occurs following the binding of the well-conserved and essential replication factor, Mcm10. Subsequent to this activation of replication, Cdc6 is released upon phosphorylation by the cyclin-dependent kinase, Cdk2/Cyclin A complex. Cdc6 is replaced by Cdc45 following its own modification by Cdk2/Cyclin E. Cdc45 associates with DNA polymerase-α and is necessary for the loading of the polymerase and DNA primase at the replication complex [21]. S-phase begins at this point; with DNA replication proceeding bi-directionally until replication of all DNA associated with that origin, termed a replicon, is complete. As DNA synthesis is beginning, Cdt1 is inactivated via direct binding by the protein geminin [25-26]. Re-licensing of the origins, and re-replication of DNA, is blocked

by the combined inactivation of Cdc6 and Cdt1, which are no longer available to load the Mcm2-7 helicase onto the DNA at the origins [21-22].

Not all of the origins fire at the same time during S-phase. Origins can fire during early, middle, or late S-phase. Clusters of origins firing at the same time comprise replication domains containing several hundred kilobases (kb) of DNA [27]. These replication domains replicate their DNA content according to a specific temporal program. This temporal program can be visualized by labeling cells with bromodeoxyuridine (BrdU), which incorporates into replicating DNA. Cells are labeled for varying amounts of time, mitotic cells harvested, and the incorporated BrdU detected with an anti-BrdU antibody and a UV fluorescence microscope [28]. Using this process, it is possible to observe the banded pattern of replication on the chromosomes.

Little is known about what regulates the temporal program of replication-timing, but several correlations have been observed (reviewed in [27]). Early replicating regions tend to be more GC-rich, contain more actively transcribed regions, and contain more euchromatin than late replicating regions, which are more AT-rich, contain less active regions of transcription, and contain more heterochromatin [29-30]. The early and late replicating regions correspond to the light and dark Giemsa-staining bands (G-bands) of chromosomes, respectively [28]. The latest replicating regions are the highly heterochromatic centromeric regions of each chromosome [31].

A better-defined correlation between early and late replicating chromosome domains is the nuclear space occupied by a replication domain. Late replicating domains tend to cluster at the periphery of the nucleus or in the perinucleolar regions, while early replicating domains occupy the area in the interior of the nucleus [32-33, 31]. This is perhaps best exemplified by the inactive X (Xi) chromosome in female cells. While the active X chromosome (Xa) replicates early and is localized to the interior of the nucleus, the Xi is late-replicating, localized to the nuclear periphery [27], and is found in a compact or condensed state called a Barr body [34]. Although no mechanism has been uncovered to explain the control of the replication timing program that exists for each chromosome, some experimental data indicate that *cis*-acting elements may control the replication timing of local replication origins (reviewed in [28]). For instance, telomere sequences appear to influence origins to fire late in S-phase. In yeast, a late-replicating origin was moved away from its location near the telomere; the result was a switch to replicating in early, rather than late, S-phase [35-36]. Conversely, when an early replicating origin was moved near the telomere, it switched to replicating in late S-phase [35]. This phenomenon of shifting replication timing by changing the genomic context of a sequence was also observed in human cells; when a deletion on chromosome 22 brought telomeric repeat sequences near origins that normally replicate in middle S-phase, they became late-replicating [37]. Replication timing of DNA sequences, then, can be modulated in *cis* by the chromosome position and sequence context of the origins. There is evidence, discussed later in this introduction, and in

Chapter 2, for *cis*-acting sequences influencing replication timing on a whole-chromosome level, not just at nearby origins.

Following DNA replication, chromosomes must condense prior to mitosis. The process of condensation compacts each chromosome by more than 10,000-fold in length compared to the B-form DNA structure (the most common form of DNA under physiologic conditions and when not wrapped around histones), while the ratio of compaction between chromosomes prior to mitosis and during mitosis is ~50-fold [20, 38]. This level of compaction is needed to prevent chromosomes from breaking apart during the process of mitosis, when chromosomes are moved first to the center of the cell along the spindle equator, and then pulled to opposite poles of the cell to form two daughter cells (for review of this process see [39]). Replicating chromosomes are also much longer than the dividing cell, so they must condense in order to avoid being cleaved during cell division [40]. Therefore, the proper condensation of chromosomes is a very important process, as under-condensed chromosomes are highly susceptible to damage, which could lead to structural rearrangements of the chromosome.

There is evidence to suggest that DNA replication itself is a crucial step in the proper condensation of chromosomes [20]. Delayed replication, then, would also result in a delay in condensation for a given chromosome or chromosomal region. Cell-fusion studies done by Johnson and Rao demonstrated that progression through S-phase is an important first step in chromosome condensation [41]. Fusions between two nuclei, one in S-phase and one in mitosis, show decondensed gaps in the chromosomes; as S-phase proceeds,

these gaps are replaced by properly condensed regions of DNA based on comparing fusions from later S-M time-points ([41] and reviewed in [20]). Therefore, Dr. Michelle Pflumm proposes that DNA replication is directly tied to the lengthwise compaction of chromosomes prior to mitosis and that this process is dependent upon the density of replication origins [20]. Her model posits that as DNA is being replicated at fixed replication centers in the nucleus, each completed replicon is pulled through and looped away. A higher replication origin density means shorter replicons and thus more compact DNA loops. The more compact DNA loops result in a more highly condensed chromosome. Ultimately, whether by direct or indirect means, replication and condensation are linked, and a defect or delay in replication has consequences for the subsequent condensation of DNA.

One example of the consequences that can arise when DNA replication and chromosome condensation are compromised is seen with chromosome fragile sites (for reviews see [42-43]). Fragile sites are observed as gaps in the condensed chromosomes; these gaps are areas of the chromosome that failed to complete replication and properly condense prior to mitosis. Not coincidentally, fragile sites are areas of very late replication during S-phase and are particularly sensitive to replication stress. As a consequence of this late replication and failure to completely condense prior to mitosis, fragile sites are prone to a high rate of chromosome breakage, resulting in deletions and translocations [43].

## Cell Cycle Checkpoints

The cell has developed several ways to ensure faithful DNA replication and cell division without sustaining permanent deleterious damage, including mechanisms to survey for DNA damage and to delay the cell cycle until that damage is repaired. These mechanisms are collectively termed cell cycle checkpoints [see [44-46] for reviews]. Traditionally, checkpoints are described as signaling cascades responsible for initiating a delay in the cell cycle, typically by affecting the activity of the cyclin-dependent kinases (Cdks) that regulate the transitions through the cell cycle. This makes it possible for the cell to respond to DNA damage or to arrest further cell cycle progression if crucial processes in the cycle have not been properly completed [44, 46]. The proteins involved in checkpoint sensing and activation also play roles in activation of DNA repair, chromatin remodeling, transcriptional regulation, and promoting permanent cell-cycle arrest or apoptosis [44].

Mammalian cells have checkpoints that can respond to environmental signals or genotoxic stress and arrest the cycle in G1 (the G1/S checkpoint) or G2 (the G2/M checkpoint), thus preventing the cell from entering S-phase or M-phase, respectively, until the environment is suitable or until the DNA damage is repaired [47, 44]. There is also a checkpoint during mitosis, the mitotic spindle checkpoint, which ensures that all chromosomes are aligned at the metaphase plate and that each chromosome is correctly attached to the mitotic spindles prior to chromosome segregation during anaphase [46]. The focus of this introduction, however, is on the S-phase checkpoints. I will not go in-depth into the proteins

involved in these checkpoints (for more comprehensive reviews, see [44-46]), but rather, I will focus more on discussing the points at which the cells must determine whether to pause S-phase, continue S-phase, or proceed to mitosis because that decision has implications for the DRT/DMC phenotype and the work discussed in this dissertation.

The cell has a replication-dependent intra-S-phase checkpoint (the replication checkpoint) that is activated when replication forks are stalled due to collisions with irregular secondary DNA structures or to DNA damage, inhibition of DNA polymerases, or a lack of deoxyribonucleotides (dNTPs), which are required for synthesis of new DNA. There is also a replication-independent intra-S-phase checkpoint that is activated in response to double strand breaks (DSBs) in the DNA, including those induced by exposure to IR [44]. Activation of these checkpoints leads to a delay or arrest of S-phase via signaling through the phosphatidyl-inositol-3-OH kinase-like [PI(3)K-like] proteins ataxia-telangiectasia mutated (ATM) and ataxia-telangiectasia and RAD3-related (ATR). ATM seems to be the predominant player in the response to DNA damage, while ATR is more critical in the response to stalled replication forks [45]. ATM and ATR activate the "effector kinases" Chk1 and Chk2, which are involved in phosphorylating and inhibiting Cdc25a, the phosphatase required for the activation of the cyclin E(A)/Cdk2 complexes. The downstream consequence of this kinase cascade is to inhibit Cdk2 from loading Cdc45 onto the DNA. Since Cdc45 normally recruits DNA polymerase α to the origins, the outcome of this Cdk2 inhibition is to prevent the remaining, non-replicated origins from firing due to the absence of the DNA

polymerase, and thus preventing the continuation of the cell cycle until the DNA damage is repaired [44-46].

The last S-phase checkpoint is the S-M checkpoint. This checkpoint serves to ensure that the entire DNA content of the cell has fully replicated prior to entry into mitosis. The mechanistic details of this checkpoint are similar to the other S-phase checkpoints in that the signaling cascade initiates with ATR [48, 44]. In this case, however, the ultimate target of the checkpoint is inhibition of the cyclin-dependent kinase Cdk1/Cyclin B complex necessary for entry into mitosis [44]. Further evidence for the role of ATR in this checkpoint comes from its reported involvement in regulating the stability of fragile sites in the genome [48]. Without ATR, fragile site expression occurs even in the absence of replication inhibition, indicating that ATR normally functions to suppress fragile site instability. Thus, ATR is important for preventing the cell from entering mitosis when late-replicating regions, such as fragile sites, have not completed DNA replication and the subsequent condensation of those regions of DNA.

Occasionally, cells continue through the cell-cycle despite the persistent activation of checkpoint responses, in a phenomenon known as checkpoint adaptation (reviewed in [49]. This was first described in the yeast *Saccharomyces cerevisiae* [50-52] and later in the frog *Xenopus* [52] , and was only more recently discovered in human cells as well [53]. The main evidence for adaptation in human cells comes from the study of cells exposed to IR. Despite sustaining DNA damage from the IR treatments, cells can continue to replicate and divide for several generations before undergoing cell death [53]. It has also

been reported that a proportion of cells with chromosome abnormalities can continue into mitosis despite activation of the S-M-phase checkpoint [54].

## DRT/DMC: a mechanism for genomic instability

In the late 1990s, the Thayer laboratory noticed a recurring chromosomal abnormality associated with certain tumor-derived chromosome rearrangements. These rearranged chromosomes showed an incomplete condensation phenotype during mitosis, termed a delay in mitotic chromosome condensation [8]. The DMC phenotype was characterized by having at least two of the following characteristics: the chromosome was at least twice as long as any other chromosome in the mitotic spread, less than half as wide as any other chromosome, or contained a bend of greater than 180°. Delayed condensation of chromosomes was also characterized by a delay in the mitosis-specific phosphorylation of serine 10 on histone H3 on these chromosomes [8, 54]. This delayed phosphorylation is likely due to an observed delay in the recruitment of Aurora B [54], which has been shown to phosphorylate histone H3 on serine 10 and plays a role in proper chromosome condensation [55].

DMC is likely a consequence of the fact that these chromosomes display DRT along the entire chromosome. These chromosomes initiate and complete replication 2-3 hours later than the other chromosomes in the same cell and can continue replication even into mitosis [8, 54]. Failure to complete replication prior to mitosis should trigger the S-M checkpoint and prevent the cell from entering mitosis [48, 44]. The chromosomes with DRT/DMC do appear to activate the S-M

checkpoint, as evidenced by the fact that they contain phosphorylated Chk1; however, DNA synthesis can still be detected at a low frequency in some cells during mitosis, indicating adaptation to the checkpoint in a proportion of the cells [54].

DRT/DMC does not appear to be a rare phenotype, as chromosomes with this phenotype were found in 5 of 7 tumor-derived cell lines and in 5 of 13 primary tumor samples that were assayed [8]. In all cases, DRT/DMC occurred only on chromosomes with rearrangements, it was never found on any nonrearranged chromosomes in any cell type. It was not known whether DRT/DMC was restricted to certain types of chromosomal rearrangements nor whether the phenomenon could only be seen in tumor cells. Therefore, the Thayer laboratory tested whether DRT/DMC chromosomes could be found in normal cells and which types of rearrangements the phenotype would be associated with following IR-induced DNA damage [9]. After exposing mammalian cells to increasing doses of IR, DRT/DMC was detected on one or more chromosomes in >25% of surviving cells, and these cells persisted in culture [9]. Primary blood lymphocytes exposed to IR produced chromosomes with DRT/DMC. IR also induced DRT/DMC *in vivo*; mice exposed to IR contained cells with DRT/DMC chromosomes for up to 2 years following exposure [9]. Chromosomes displaying DRT/DMC were predominantly derivatives of inter-chromosomal translocations. Overall, the data from the IR experiments indicate that DRT/DMC can occur in normal cells, both *in vitro* and *in vivo*, and that ~5%

of all inter-chromosomal translocations induced by IR display the DRT/DMC phenotype [9].

Chromosomes with DRT/DMC are also unstable [8]. Cells containing one or more chromosomes with DRT/DMC have a 30-80-fold increased rate of new gross chromosomal rearrangements [10]. It was proposed that these chromosomes display TIN due to their failure to complete DNA replication prior to mitosis [8]. DRT/DMC chromosomes also display abnormal centrosome amplification and mitotic spindles, an increase in endoreduplication (DNA replication without cell division, resulting in polyploid karyotypes), and significant CIN [54]. Taken together, the data on the instability associated with DRT/DMC chromosomes suggest that DRT/DMC may contribute to the TIN and CIN phenotypes commonly seen in cancer cells and in cells exposed to IR.

Due to the relatively high frequency at which DRT/DMC was found following IR (~5%), it was hypothesized that a chromosome engineering strategy designed to generate inter-chromosomal translocations could produce translocation derivatives that display the DRT/DMC phenotype [9-10]. This would allow for the systematic analysis of DRT/DMC. A Cre/*lox*P site-specific recombinase system [see [56] for a review of Cre/*lox*P] was developed for the generation of cell lines containing randomly targeted chromosomes that undergo balanced translocations [9-10] (see Appendix 1 text and Fig. A1.1). Briefly, *lox*P sites were randomly inserted into the genome of the adenine phosphoribosyltransferase- (*APRT*-) null cell line, HTD114 [57]. Cre recombinase, which generates recombination between the 34 bp Cre-recognition sequences (*lox*P sites), was

subsequently transfected into the cells. Cre works in a distant-dependent manner [58], so inter-chromosomal translocations occur at a low frequency. In order to isolate such events, recombination between the *lox*P sites reconstitutes a selectable marker, the *Aprt* gene.

Approximately 5% of the engineered translocations result in at least one derivative with DRT/DMC [10]. Four cell lines from this experiment have balanced translocations in which only one derivative displays DRT/DMC. In addition to site-specific recombination with Cre recombinase, translocations were also engineered by non-homologous end joining (NHEJ) following double-stranded breaks induced by the restriction enzyme I-Sce1. There is no difference in phenotypic outcome between these two methods, indicating that DRT/DMC is not the consequence of a specific type of DNA damage or of a specific DNA repair process [10]. Taken together, the data from the chromosome engineering experiments indicate that DRT/DMC is controlled by a *cis*-acting mechanism [9-10]. Engineered chromosomes with DRT/DMC recapitulate all of the aspects of the phenotype previously described, the delay in replication timing, delay in mitotic condensation, and the genomic instability [8-10], collectively making them an ideal model with which to study the initiating events of DRT/DMC.

## X-chromosome inactivation as a model for DRT/DMC

Many organisms have had to evolve a mechanism to deal with expression differences between females with two X chromosomes and males with only one X chromosome (reviewed in [59]. In *Drosophila*, transcription from the one male X chromosome is hyperactivated so that it equals the level of transcription from the two female X chromosomes. In the worm *Caenorhabditis elegans*, transcription from the two female X chromosomes is halved so that it equals the transcriptional level of the one male X chromosome. Mammals have evolved a third path for dosage compensation, which is to inactivate one of the two X chromosomes from female cells. In placental mammals, this process of X-chromosome inactivation (XCI) involves the random inactivation of one of the X chromosomes early in development. Once inactivated, the silenced state of the inactive X (Xi) chromosome persists and is stably inherited by each successive generation [59].

In order to always maintain this dosage compensation, XCI follows an ‘n-1 rule’ with respect to the number of inactive X chromosomes in a cell (with n being the total number of X chromosomes in a cell, and ‘n-1’ being the number of X chromosomes that become inactive), so that only one X chromosome is chosen to be active per cell and all remaining X chromosomes are inactivated [60]. This ‘n-1 rule’ ensures that even in cases of trisomy of the X chromosome, only one X will be active and the remaining two will become inactive. The end result of XCI is the mono-allelic expression of over a thousand genes from the active X (Xa) chromosome, and thus equalized gene expression between XY males and XX

females [61]. Some genes escape inactivation; these are genes with homologs on the Y chromosome, genes that may not be as sensitive to expression levels, or genes that exhibit tissue-specific escape due to different gene dosage requirements in different tissues [62].

XCI is a well studied process. It was first described in the mouse in 1961 [63], and since that time, researchers have shown that the Xi takes on many of the hallmarks of facultative heterochromatin (reviewed in [61, 59, 64, 62]), the type of heterochromatin that is able to form compact chromatin structures similar to constitutive heterochromatin, but can also adopt an open conformation in certain temporal or spatial circumstances [65]. The Xi predominantly localizes to the nuclear periphery or the nucleolus in a visibly condensed state, known as the Barr body [34, 62]. The process of XCI results in asynchronous replication timing between the two X-chromosomes; the Xa replicates early, and the Xi replicates late [66, 62]. The Xi and Xa are also differentiated by their respective global chromatin modifications. The Xi is enriched in histone modifications that are consistent with heterochromatin, such as H3K27me3, H3K9me2, H3K9me3, and marcroH2A; the Xa, meanwhile, is enriched in euchromatic-associated histone marks, such as H3K4me2/3 and acetylated lysines on H3 and H4 [62].

The process by which XCI is established involves several key players residing at a locus on the X chromosome known as the X-inactivation center (Xic). In this 100-500 kb region, there are at least seven non-coding transcripts [67], several of which will be discussed here. *Xist* (X-inactivation specific transcript) is a 17 kb non-coding RNA (ncRNA) that is expressed only from the Xi. *Xist* RNA spreads

along the Xi in *cis*, coating the chromosome, and is thought to recruit chromatin modifiers and silencing complexes along the way [68, 67]. *Xist* spreading leads to the creation of a compartment that excludes RNA polymerase II and transcription factors, followed by silencing of the genes on the Xi that reside in this compartment. The genes that escape XCI remain outside of this compartment [69]. *Xist* is negatively regulated in *cis* by *Tsix*, a 40 kb ncRNA that is transcribed in the antisense orientation, beginning ~12 kb downstream of *Xist* [61, 67]. Prior to XCI, *Tsix* is transcribed from both X chromosomes; but upon differentiation, *Tsix* expression is lost from the Xi as *Xist* expression becomes up-regulated. The continued expression of *Tsix* from the Xa appears to prevent the up-regulation of *Xist* from that chromosome, thus preventing the inactivation of the Xa. Upstream of *Tsix*, an enhancer element, the ncRNA *Xite*, maintains the expression of *Tsix* on the Xa [67]. XCI, at least in part, seems to hinge on the balance of expression between *Xist* and *Tsix* (see Fig 1.1).

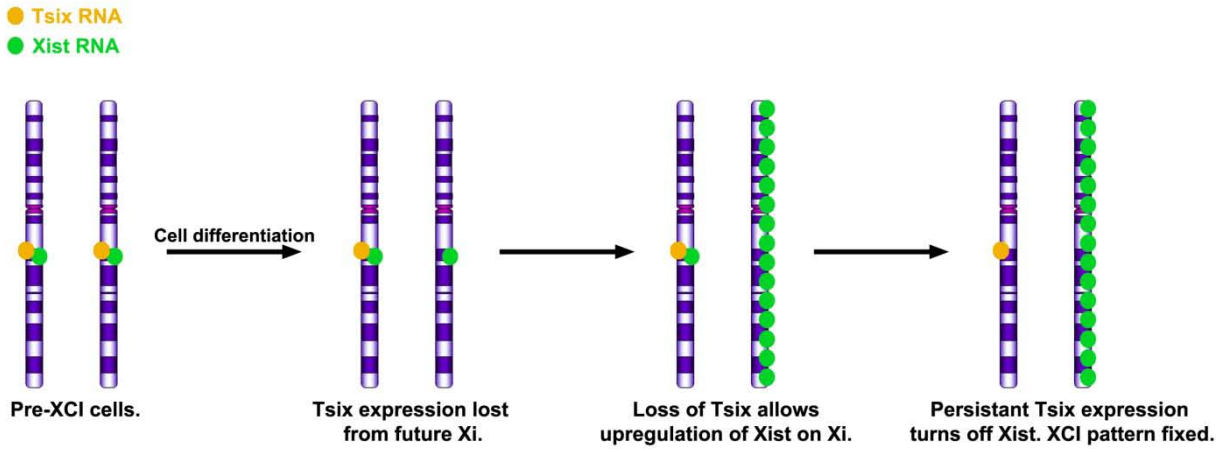



**Figure 1.1. Relationship between Tsix and Xist RNAs.** The relationship between *Tsix* and *Xist* expression helps establish the pattern of XCI. Prior to cell

differentiation and XCI, both X chromosomes express high levels of *Tsix* and low basal levels of *Xist*. Following cell differentiation, *Tsix* expression is restricted to the chromosome chosen to remain active. *Tsix* expression on the Xa prevents up-regulation of *Xist* on the Xa. Conversely, the loss of *Tsix* from the Xi allows for the up-regulation and spreading of *Xist* on the Xi to initiate silencing. (Figure adapted from [67]).

Counting the number of X chromosomes in a given cell, and the choice of which X chromosome will remain active appears to involve *trans* communication between the X chromosomes; and this communication involves regions within the Xic [70-72]. These processes depend upon transient homologous pairing between the two X chromosomes at the beginning of XCI; this pairing occurs between the Xics of each chromosome, requiring specific sequences within the *Xite* and *Tsix* genes [70-72].

Another of the ncRNAs that has been identified and characterized at the Xic locus is a small 1.6 kb transcript called RepA. This ncRNA is transcribed from the 5' end of the Xist locus; specifically, from the region known as the Repeat A motif, which is known to be essential for the function of *Xist*. The RepA transcript has been shown to bind Ezh2, a subunit of Polycomb repressive complex 2 (PRC2) [73]. It is thought that *RepA* recruits PRC2 to the Xi where PRC2 then lays down the repressive chromatin mark, H3K27me3.

The maintenance of the repressive state of the Xi seems to involve a specific aspect of the Xi nuclear localization. Following pairing of the two X

chromosomes, the Xi localizes to the perinucleolar region [74], a subnuclear structure located near the nuclear periphery that is associated with, but distinct from, the nucleolus [75]. It has been proposed that the heterochromatic state of the Xi is replicated and maintained at this perinucleolar compartment during its visit there during mid-to-late S-phase [74]. This perinucleolar localization is dependent on *Xist* and the Xic.

One possible mechanism for XCI involves the high density of long interspersed nuclear element-1 (LINE-1) elements on the X chromosome. LINE-1 elements are ancient retrotransposons specific to the mammalian genome. Most of these repeats are transcriptionally silent, although some of the younger elements retain some transcriptional and retrotransposon activity [76]. The human genome is ~17% LINE-1 repeats, but the X chromosome has nearly 30% LINE-1 repeats [61], with the highest density clustering around the Xic [77]. This over-representation of LINE-1 repeats, along with data from X;autosome translocations showing that the spread of inactivation along the autosome correlates with the density of LINE-1 repeats present on the autosome, led to the Lyon repeat hypothesis [78]. This hypothesis posits that the LINE-1 repeat serves as the “booster element” first proposed by Gartler and Riggs [79] that spreads the X inactivation signal from the Xic along the length of the chromosome. Later studies have furthered the notion that LINE-1 elements do indeed play a role in XCI [77, 80-81, 76].

In 2005 and 2006, Dr. York Marahrens’ laboratory published two papers that make XCI an intriguing model for DRT/DMC [82-83]. In the first set of

experiments, a 21 kb region of the *Xist* locus was deleted from the Xa, the non-transcribed allele of *Xist*, in adult somatic cells. The consequence of this deletion was to alter the replication timing of the entire chromosome, causing it to replicate significantly later in S-phase than the wild-type Xa without the *Xist* deletion replicates [82]. The implications from this study are that elements at this locus are responsible for maintaining the early replication timing of the entire Xa, in *cis*.

In the second set of experiments, Diaz-Perez et al. noticed that while this same 21 kb deletion of the *Xist* locus on the late-replicating Xi caused the loss of certain chromatin modifications specific to the Xi, it did not abolish the late replication; instead, it caused a shift to even later replication in S-phase [83]. Additionally, when the *Xist* deletion was made on both X chromosomes, it caused extreme late replication of certain chromosome regions, specifically those associated with a high density of LINE-1 elements. This very late replication resulted in fragile sites on the chromosomes, indicating condensation defects, which subsequently led to deletions and translocations involving both X chromosomes [83]. These two reports by Diaz-Perez et al. indicate that this 21 kb region of the *Xist* locus on both the Xa and Xi has a role in the replication timing of the entire chromosome, in *cis*. These studies also suggest that this locus plays a role in the communication between the two chromosomes. This communication between the homologs also appears to be involved in the replication timing of the chromosomes, suggesting a *trans* effect for this locus. The observations published in these two studies closely resemble the DRT/DMC

phenotype, indicating that perhaps there are similarities between the *Xist* locus, which has now been shown to be involved in the replication timing of the X chromosome, and the loci that are involved in the DRT/DMC phenotype.

## XCI properties on autosomes

Since deletions at the Xic recapitulate properties of the DRT/DMC phenotype, an important question is whether the properties associated with the Xic and with XCI also exist on autosomes. The Xic is known to be involved in counting the number of X chromosomes, in mono-allelic expression of the genes on the X chromosome, and in the asynchronous replication timing of the X chromosomes [one X replicates early (Xa), and one replicates late (Xi)] [64, 62]. It turns out that these properties are not restricted to the X chromosome; they can be found on all mammalian autosomes as well.

X-chromosome inactivation shares many properties with the phenomena of genomic imprinting and random mono-allelic inactivation (reviewed in [84-85]). Genomic imprinting results in the parent-of-origin-specific expression of only one allele of a gene. It is always either the maternal allele or the paternal allele that is expressed, while the opposite allele is inactivated. Imprinted genes tend to reside in clusters with other imprinted genes. In the mouse, there are ~90 known imprinted autosomal genes [85]. In random mono-allelic expression, similar to XCI, the choice of which allele to express is random; therefore, in any given cell only one allele will be expressed, but in a population of cells both alleles can be detected.

Random mono-allelic expression was first studied in large multigene families like the olfactory receptor genes and the immunoglobulins are often found arranged in clusters, similar to imprinted genes [85]. In order to ensure expression of only a single antigen receptor, the immune system undergoes random mono-allelic inactivation of the antigen-specific receptors; this process is therefore described as allelic exclusion [84]. However, random mono-allelic expression may be much more prevalent. Gimelbrant et al. showed that mono-allelic expression is a wide-spread phenomenon that might occur in 5-10% of all genes [86]. The patterns of mono-allelic expression could change based on developmental stage and tissue type, further increasing diversity in the regulation of expression.

In addition to sharing the trait of mono-allelic expression, the regulatory mechanisms of these processes have many similarities. All of these processes involve differential chromatin modifications, *cis*- and *trans*-acting factor involvement in regulation and allele choice, nuclear positioning of the alleles, and asynchronous replication timing; frequently, many of these features involve ncRNAs [84-85]. Additionally, imprinted gene clusters and random mono-allelically expressed genes, like the X chromosome, tend to reside in regions with a high LINE-1 density [87].

Clusters of imprinted genes contain an ncRNA, which can be >100 kb in length, and a *cis*-acting imprint control element (ICE) [85]. The ICE is differentially methylated between the parental alleles. The ncRNA is expressed from the chromosome with the unmethylated ICE, and expression of this ncRNA

correlates with repression of the imprinted protein-coding genes in the cluster, in *cis* [85]. Studies on the functional consequences of deleting these ncRNAs have only been done for three imprinted loci: the *Air* ncRNA at the *Igf2r* imprinted cluster, the *Kcnq1ot1* ncRNA at the *Kcnq1* cluster, and the *H19* ncRNA at the *Igf2* cluster. Of these ncRNAs, *Air* and *Kcnq1ot1* had a clear role in the silencing of the imprinted alleles, while *H19* did not [88].

Various models have been proposed for how these ncRNAs might function in the silencing of genes at the imprinted clusters [88]. They could operate in a manner similar to *Xist*, which coats the Xi, recruits silencing factors and creates a silencing compartment [68, 67]. However, this may not be the case, as imprinted clusters involve only a few genes, not the entire chromosome. Additionally, there is an imprinted form of XCI in marsupials and in the extra-embryonic tissue of mice; the imprinted XCI in marsupials is unlikely to require an *Xist* ortholog [88], and the initiation of the imprinted XCI in the mouse extra-embryonic tissue is independent of *Xist* (though *Xist* is necessary for stabilizing the silencing) [89]. Therefore, it is unlikely that the ncRNAs at these imprinted loci would function exactly like *Xist*. Since imprinted and random XCI appear to be regulated in a different manner, it follows that autosomal imprinted clusters would have a unique, though potentially overlapping, regulatory system.

Another proposed model for the function of the ncRNAs at imprinted loci involves RNA interference (RNAi), a gene-silencing mechanism brought about by the formation of double-stranded RNA [90]. Many of the ncRNAs at the imprinted loci are transcribed in the antisense orientation from one of the imprinted genes

and thus share the sequence homology necessary for RNAi; however, this would not explain the silencing of other genes located farther away in the imprinted cluster [88]. It is still possible, though, that repeat elements such as LINE-1 elements, which are common in these imprinted clusters, could serve as substrates for the RNAi-based silencing [91]. Other models for the function of these ncRNAs include the idea that it is transcription through the ncRNA region, and not the ncRNA itself that is responsible for silencing, or that the ncRNA is involved in forming higher-order chromatin structures or subnuclear silencing compartments [88]. One common and important principle of many of these models is that they involve a *cis*-acting component in the regulation of these loci.

Progress in clarifying the mechanism of regulation by the *Kcnq1ot1* ncRNA at the *Kcnq1* locus has been made recently [92-93]. The *Kcnq1* imprinted cluster is an ~1 Mb region that contains 8-10 imprinted genes and the 91 kb ncRNA *Kcnq1ot1* [91]. The *Kcnq1ot1* ncRNA is expressed only from the paternal allele, which correlates with *cis*-mediated silencing of the protein coding genes located on the same chromosome, resulting in maternal-specific expression of those genes [91]. *Kcnq1ot1* has an important domain at its 5' end, composed of conserved repeats, that is involved in the transcriptional silencing of the imprinted genes [92]. This 5' domain is also required for targeting the imprinted cluster to the perinucleolar compartment, which is enriched in factors involved in replication and heterochromatin formation, during S-phase [92]. The *Kcnq1ot1* ncRNA has been shown to have a role in the silencing of genes, rather than merely the act of transcription at this locus. *Kcnq1ot1* is a nuclear-only, moderately stable RNA

polymerase II transcript that interacts with chromatin to mediate gene silencing [93]. It binds to both the G9a methyltransferase and components of the Polycomb repressive complex 2 (PRC2), such as SUZ12 and EZH2 [93]. It is thought that *Kcnq1ot1* recruits and directs these proteins, which are involved in generating repressive chromatin modifications (mono- and dimethylation on H3K9 by G9a [94] and H3K27 trimethylation by PRC2 [95]) to the genes that are silenced, in a lineage-specific manner [93]. Kcnq1ot1 also contains a high proportion of LINE-1 elements that could play a role in transcriptional silencing at this locus [91].

Imprinted regions are also known to replicate asynchronously, with the active allele replicating early in S-phase and the inactive allele replicating late [84-85]. The asynchronous replication timing is established in the gametes and maintained throughout development, and it is independent of DNA methylation [96-97]. However, the late replicating alleles do correspond to a more peripheral nuclear location than the early replicating allele [97]. The ICE region of the Igfr-*H19* imprinted locus appears to be required for its asynchronous replication timing, indicating regulation of this process by a *cis*-acting mechanism [98].

In allelic exclusion and olfactory receptor regulation, various *cis*-acting regulatory elements have been identified [84-85]. Often, these regulatory elements are ncRNAs [84]. Differential methylation of the alleles has been observed, as has asynchronous replication timing, with the demethylated allele replicating early and the methylated allele replicating late [84]. Much remains unknown about what regulates the widespread random mono-allelic expression that has been identified on human autosomes [86], but it would not be surprising

to find *cis*-acting sequences, ncRNAs, asynchronous replication timing, or all three elements playing a role.

The asynchronous replication timing present on the X chromosome encompasses the entire chromosome, whereas the asynchronous replication timing of random mono-allelically expressed genes on autosomes occurs only at distinct loci scattered along the chromosome. However, the autosomal asynchronous replication timing is coordinated at the whole-chromosome level [99-101]. It has been shown in both mouse [99] and human [100-101] cells that the early replicating alleles are always on the same chromosome and the late replicating alleles are on the homologous chromosome. This coordination crosses the centromere, indicating that it is a chromosome-wide coordination [100].

Like the X chromosomes, human autosomes also show an ‘n-1 rule’ [100]. When three alleles for an asynchronously replicating gene are present in a cell, one allele replicates early in S-phase and the other two alleles replicate late [100]. In XCI, only one X chromosome is active and early replicating, while all others are inactivated and replicate late [84]. Therefore, human autosomes, like the X chromosomes, show chromosome-pair non-equivalence along the entire chromosome, meaning that though they are homologous, the two chromosomes behave differently. This non-equivalence is established early in development, as it is also observed in human embryonic stem cells [101].

Although the replication timing of autosomal mono-allelically expressed genes is coordinated on the whole-chromosome level, the expression of mono-allelic genes does not appear to be coordinated [86]. For example, human B cell clones analyzed by Gimelbrant et al. could mono-allelically express the maternal allele or the paternal allele, while other clones show bi-allelic expression [86]. Sometimes the expressed allele of one gene is from the maternal chromosome, while the expression of another gene is from the paternal chromosome, meaning that expression of the mono-allelic genes is not coordinated in *cis* in the same way that the asynchronous replication timing of mono-allelic genes is [99-100]. It has also been shown that the mouse *p120* catenin gene can be mono-allelic in some cell types, but bi-allelic in others [102]. However, the replication timing at the *p120* catenin locus is still asynchronous, whether expression is mono-allelic or bi-allelic. The same holds for the olfactory receptor genes in both mouse and human cells; these mono-allelically expressed genes replicate asynchronously even in cell types in which they are not expressed [99-100]. In total, these studies indicate that expression and asynchronous replication timing are not directly correlated in every cell; the mono-allelic expression of transcripts could arise via a distinct mechanism from the asynchronous replication timing of the corresponding DNA, such as allele-specific DNA methylation or allele-specific transcription factor binding [103]. However, the asynchronous replication timing may indicate that a gene is mono-allelically expressed in at least one cell type or at one particular developmental stage [102].

The direct comparison between coordination of replication timing and 'non-coordination' of expression, however, has not been done at the same time in the same system. A more detailed analysis is required in which the coordination of asynchronous replication timing and allele-specific expression are done in parallel to see whether the non-coordination of expression is a feature of all random mono-allelically expressed genes or whether it occurs only for a subset of those genes.

In the data presented in this dissertation, in which I explore the genetic basis for DRT/DMC, I will provide evidence to support our hypothesis that human autosomes have *cis*-acting loci that, similar to the Xic, control chromosome-wide replication timing. The locus we identified on chromosome 6 that controls chromosome-wide replication timing also controls the mono-allelic expression of nearby transcripts. In addition to displaying mono-allelic expression, this locus also displays asynchronous replication timing. Although this asynchronous replication timing is coordinated in *cis* with nearby asynchronously replicating genes, it is coordinated in *trans* with other asynchronously replicating genes on chromosome 6. In *trans* coordination, the early replicating allele of one gene is coordinated with the late replicating allele of another gene on the same homolog. This is the first time *trans* coordination has been observed for asynchronously replicating genes.

# CHAPTER TWO

## An Autosomal Locus that Controls Chromosome-wide Replication Timing and Mono-Allelic Expression

Eric P. Stoffregen, Nathan Donley, Daniel Stauffer, Leslie Smith, and Mathew J. Thayer*

Department of Biochemistry and Molecular Biology

Oregon Health & Science University

3181 S. W. Sam Jackson Park Road

Portland, Oregon 97239, USA

***Corresponding author.**

TEL: (503) 494-2447

FAX: (503) 494-7368

Email:thayerm@ohsu.edu

**Human Molecular Genetics** (2011) **20** (12), 2366-2378.

## Abstract

Mammalian DNA replication initiates at multiple sites along chromosomes at different times, following a temporal replication program. Homologous alleles typically replicate synchronously; however, mono-allelically expressed genes such as imprinted genes, allelically excluded genes and genes on the female X chromosome replicate asynchronously. We have used a chromosome engineering strategy to identify a human autosomal locus that controls this replication-timing program in *cis*. We show that Cre/*lox*P-mediated rearrangements at a discrete locus at 6q16.1 result in delayed replication of the entire chromosome. This locus displays asynchronous replication timing that is coordinated with other mono-allelically expressed genes on chromosome 6. Characterization of this locus revealed mono-allelic expression of a large intergenic non-coding RNA, which we have named asynchronous replication and autosomal RNA on chromosome 6, *ASAR6*. Finally, disruption of this locus results in the activation of the previously silent alleles of linked mono-allelically expressed genes. We previously found that chromosome rearrangements involving eight different autosomes display delayed replication timing, and that cells containing chromosomes with delayed replication timing have a 30-80-fold increase in the rate at which new gross chromosomal rearrangements occurred. Taken together, these observations indicate that human autosomes contain discrete *cis*-acting loci that control chromosome-wide replication timing, mono-allelic expression and the stability of entire chromosomes.

## Introduction

Morphological differences between chromosomes within the same cell were first observed in mammalian cells over 40 years ago [1-2]. These alterations, referred to as ‘incomplete condensation’ or ‘pulverization’ of chromosomes, could occur on one or a few chromosomes during mitosis. In addition, these abnormally condensed chromosomes synthesized DNA after the normally condensed chromosomes had ceased replication [3-4]. However, whether or not the chromosomes displaying these abnormal phenotypes contained structural abnormalities was not determined in these earlier studies. Subsequently, we found that certain tumor-derived chromosome rearrangements exhibit a significant delay in replication timing (DRT), which is characterized by a >2 h delay in both the initiation as well as the completion of DNA synthesis along the entire length of the chromosome [8]. Chromosomes with DRT also display a significant delay in mitotic chromosome condensation (DMC) that is characterized by an under-condensed appearance during mitosis and a concomitant delay in the phosphorylation of serine 10 of histone H3 [8, 54]. Chromosomes with DRT/DMC were detected in 5 of 7 tumor-derived cell lines and 5 of 13 primary tumor samples [8]. DRT/DMC was also detected on ~5% of inter-chromosomal translocations induced by exposure to ionizing radiation [9]. Finally, we found that cells containing chromosomes with DRT/DMC have a 30-80-fold increase in the rate at which new gross chromosomal rearrangements occur, indicating that DRT/DMC causes genomic instability [10]. Taken together,

these observations indicate that DRT/DMC is a common phenotype in cancer cells and in cells exposed to ionizing radiation.

We have developed a chromosome engineering strategy that allows for the systematic analysis of chromosomes with DRT/DMC [9-10]. This strategy is based on the Cre/*lox*P site-specific recombinase system to generate precise chromosomal rearrangements (Appendix 1, Fig. A1.1). Using this system, we previously identified four balanced translocations, each displaying DRT/DMC on only one of the derivative chromosomes [10]. In this report, we show that Cre/*lox*P-mediated translocations or deletions at a discrete locus on human chromosome 6 result in DRT/DMC. In addition, we show that the rearrangements that cause DRT/DMC disrupt a mono-allelically expressed, asynchronously replicating, long intergenic non-coding RNA, which we have named asynchronous replication and autosomal RNA on chromosome 6, *ASAR6*. Finally, we show that the genetic alterations that cause DRT/DMC also result in bi-allelic expression of two linked mono-allelically expressed genes.

## Results

### ‘Alternative partner’ analysis

Our models for the generation of DRT/DMC predict that either disruption of discrete chromosomal loci that promote normal replication timing of entire chromosomes, or that juxtaposition of incompatible chromosome domains at certain translocation breakpoints, results in delayed replication that subsequently results in delayed mitotic condensation [10, 54]. In either case, the information that is responsible for the phenotype of the engineered translocations with DRT/DMC must be linked in *cis* to one or both of the *lox*P cassettes. Therefore, to determine whether one or both *lox*P integration sites are necessary for generating the phenotype, we created ‘alternative partner’ translocations with an existing pair of *lox*P-tagged chromosomes, which we had previously shown to generate a Cre-dependent balanced translocation displaying DRT/DMC [10] (Appendix 1, Figs A1.2 and A1.3]. Thus, generating different translocations with these existing *lox*P-tagged chromosomes would allow us to determine whether DRT/DMC segregates with one or both of the integration sites, and would allow us to determine whether DRT/DMC occurs only with certain chromosomal exchanges and not with others. Fig. 2.1A shows a schematic representation of this alternative partner analysis for chromosomes 6 and 10 in the P175 cell line [10] (see Appendix 1 for a more detailed description of this alternative partner analysis). In total, we have characterized four new balanced translocations

involving chromosome 6 and three new balanced translocations involving chromosome 10 (Appendix 1, Figs A1.4-12).

We next used a BrdU ‘terminal label’ protocol to measure the chromosome replication timing of these new translocations (Fig. 2.1B). This protocol allows us to visualize the latest replicating regions of chromosomes. Accordingly, the banded pattern of BrdU incorporation allows us to detect actively replicating regions of chromosomes, and differences in replication timing between chromosome pairs are seen as differences in this banding pattern. In addition, measuring the total amount of BrdU incorporation in individual chromosomes allows us to quantify any differences in the normally synchronous replication timing of homologous chromosome pairs. First, to characterize the replication timing of the chromosomes involved in the alternative partner translocations, we carried out an extensive analysis of chromosome replication timing in the parental P175 cells prior to the generation of any Cre-mediated rearrangements (Appendix 1, Figs A1.13-16). Analysis of the banding pattern of BrdU incorporation at multiple time points indicated that the replication timing of each pair of chromosomes, 1, 4, 5, 6, 7, 8, 9, 10, and 17, was consistent with the known replication timing maps for these chromosomes [104-105]. In addition, analysis of the banding pattern and quantification of the BrdU incorporation indicated that each pair of chromosomes replicated synchronously (Appendix 1, Figs A1.13-16). Therefore, the alterations in replication timing of the rearranged chromosomes, described below, are due to Cre-mediated events and not to pre-existing replication timing differences in this set of chromosomes.

Figure 2.1C and D shows an example of BrdU incorporation into the chromosomes of a mitotic cell containing a chromosome 6 alternative partner, t(6;9)(q15;p21) (Appendix 1, Fig. A1.7). The only chromosome showing detectable BrdU incorporation in this mitotic spread was the chromosome 9 derivative of the t(6;9), indicating that it displays DRT. Comparing the banded pattern of BrdU incorporation of the t(6;9) with the non-rearranged chromosomes 6 and 9 in P175 cells (Appendix 1, Figs A1.13 and A1.14) indicated that the chromosome 9 derivative was delayed in replication by at least 4 h. Note that the chromosome 9 derivative of the t(6;9) displays DRT, but does not display DMC in this mitotic cell. Analysis of additional mitotic spreads indicated that the chromosome 9 derivative does indeed display the DMC phenotype (Appendix 1, Fig. A1.7c). A second example of DRT without DMC on this t(6;9) is shown in Appendix 1, Fig. A1.8. Given these inconsistencies in detecting DMC, we have concentrated on the replication timing of the chromosome rearrangements described below.

DRT was also detected on two other chromosome 6 alternative partner translocations, a t(6;17)(q15;q25) and a t(6;7)(q15;q36) (Appendix 1, Figs A1.5 and A1.6, respectively). However, analysis of a fourth translocation involving chromosome 6, t(6;8)(q15;q24.1), indicated that it did not display DRT (Appendix 1, Fig. A1.9). Therefore, three of four alternative partner translocations involving this chromosome 6 *lox*P cassette integration site display DRT. The reason for the apparent ‘normal’ replication timing of the t(6;8) translocation is currently unknown. Regardless, we found that four of five translocations at this *lox*P

integration site, including the original t(6;10), display DRT. Furthermore, we found that only one derivative chromosome of these balanced translocations displays DRT, and that DRT segregated with the distal portion of chromosome 6. In contrast, none of the new alternative partner translocations involving chromosome 10, a t(1:10)(p22.3;q11.2), a t(4;10)(q25;q11.2) and a t(5;10)(q35.1;q11.2), displays DRT (Appendix 1, Figs A1.9, A1.10, and A1.11, respectively). These observations suggest that the chromosome 6 integration site is required for generating DRT on this set of inter-chromosomal translocations, and that the chromosome 10 integration site plays only a passive role in generating DRT/DMC on the original t(6;10).

## Chromosome 6 deletions display DRT

We next determined whether deletions in chromosome 6, nested at the original *lox*P-3'RT integration site in P175 cells, also display DRT. For this analysis, we used a lentiviral vector to generate pools of clones containing new AP-*lox*P cassette integrations in P175 cells. The rationale behind this set of experiments was based on previous studies showing that intra-chromosomal Cre events are markedly more efficient than inter-chromosomal Cre events [106]. Thus, expression of Cre in pools of lentiviral infected cells would reconstitute the selectable marker (*Aprt*) at a higher frequency in cells containing lentiviral integrations near the original *lox*P-3'RT cassette. A schematic diagram of the lentiviral construct is shown in Fig. 2.2A (also see Appendix 1, Figs A1.1 and A1.17A). In addition, because this scheme can generate deletions in only one

direction with respect to the original *lox*P-3'RT integration site (i.e. proximal), we also modified the original *lox*P-3'RT integration site in P175 cells so that deletions could be generated distal to the original *lox*P cassette (Appendix 1, Figs A1.18 and A1.19). Using this random lentiviral integration approach, we have generated more than 50 independent deletions in chromosome 6, and these nested deletions extend either proximal or distal from the original *lox*P-3'RT cassette integration site. To characterize these deletions in molecular detail, we first cloned and sequenced the *lox*P-3'RT integration site in P175 genomic DNA [96,386,321 base pairs (NCBI Build 36/hg18); see Appendix 1]. A schematic illustration of this integration site, showing the location of the two closest protein-coding genes *MANEA* and *FUT9*, is shown in Figure 2.2A. We next characterized the deletions using multiple independent assays, including: Southern blot hybridizations (Appendix 1, Fig. A1.17B); junction PCR designed to span the genome-*lox*P-3'RT cassette (Appendix 1, Fig. A1.17C and A1.17D); LAM-PCR [107] to identify the lentiviral 5'-LTR integration sites (Appendix 1, Table A1.1); Affymetrix SNP micro-array and PCR-based assays for loss of heterozygosity (LOH) and copy number changes (Appendix 1, Figs A1.20 and A1.21); and fluorescence *in situ* hybridization (FISH) using BACs (Bacterial Artificial Chromosomes) or Fosmids (F-factor Cosmid cloning vectors) located within the deleted regions (Figs 2.2A-C, 2.3G; Appendix 1, Figs A1.22, A1.23 and A1.25). A partial set of the lentiviral integration sites for the proximal deletions is shown in Figure 2.2A (Appendix 1, Table A1.1). The largest deletion proximal to the *lox*P-3'RT integration site is ~30 Mb and extends to near the centromere, and the

smallest proximal deletion is ~76 kb. The largest deletion distal to the original *lox*P cassette is ~26 Mb, and the smallest distal deletion is only ~18 kb (Appendix 1, Fig. A1.19 and Table A1.1).

We next assayed the replication timing of the chromosome 6s in a subset of these deletion clones. Figure 2.2C shows examples of the BrdU incorporation pattern in cells containing an ~16.6 Mb proximal deletion. Cultures were incubated with BrdU for 4.5 h, and mitotic cells were harvested, processed for BrdU incorporation and subjected to FISH, using a chromosome 6 centromeric probe plus a BAC from the deleted region (Fig. 2.2A). The FISH signal from the chromosome 6 centromeric probe allowed us to identify both chromosome 6s, and the presence or absence of the BAC allowed us to distinguish between the non-deleted or deleted chromosomes, respectively. Comparing the BrdU incorporation patterns between chromosome 6s in multiple cells indicated that the deleted 6s were delayed in replication timing by >2 h. Similarly, DRT was detected on all of the proximal deletions assayed, including deletions of: ~21.8 Mb (Appendix 1, Fig. A1.22), ~2.1 Mb (not shown), ~233 kb (not shown), ~231 kb (Appendix 1, Fig. A1.23), ~94 kb (Appendix 1, Fig. A1.24), and ~76 kb (Fig. 2.3). Importantly, delayed replication was detected on both arms of the deleted chromosomes (Fig. 2.3A-D; Appendix 1, Figs A1.22-24), indicating that the effects of these genetic alterations cross the centromere. In contrast, the deletions distal to the original *lox*P-3'RT integration site, including deletions of ~25.9 Mb (Appendix 1, Fig. A1.25A-C), ~1.4 Mb (Appendix 1, Fig. A1.25D), and ~18 kb (Appendix 1, Fig. A1.25E), did not display DRT. Although the extent of

DRT in any given cell can be quite variable, we have not detected any significant differences between the different deletion mutations in chromosome 6. Taken together, these observations indicate that DRT occurs with deletions as small as ~76 kb proximal to the original *lox*P-3'RT integration site, and does not occur with deletions distal to this integration site, even with deletions as large as ~26 Mb.

## Mono-allelic Expression and DRT

Previous reports have indicated that Cre/*lox*P-mediated deletion of the *Xist* gene from adult somatic cells causes a delay in the replication timing of the entire X chromosome [82-83]. This delayed replication phenotype appears to be similar to the DRT phenotype observed on the autosomes described here and previously [8-10, 54]. Therefore, given the similarities in replication timing defects associated with *Xist* gene deletions and the deletions in chromosome 6, we next determined whether the region of chromosome 6 responsible for DRT is transcribed. Our first indication of transcription from the intergenic region between *MANEA* and *FUT9* came from a report that identified RNAs expressed in human embryonic stem cells that were associated with the RNA-binding protein FOX2 [108]. To confirm expression of this region in our cell culture system, we used RT-PCR to assay numerous locations within the intergenic region between *MANEA* and *FUT9*. First, we used primers designed to amplify unique sequences from RNA extracted from HTD114 (the parent cells of P175), P175 and two chromosome 6 mono-chromosomal hybrids (mouse L cells, each containing one of the two chromosome 6s from HTD114). We detected RNA

expression from HTD114, P175, and one of the mono-chromosomal hybrids (Appendix 1, Table A1.2; Fig. 2.4D). Consistent with this observation, allele-specific RT-PCR of a heterozygous single-nucleotide polymorphism (SNP) indicated that the RNA transcripts expressed in P175 cells were mono-allelically expressed (Fig. 2.4A). In addition, Figure 2.4B and C shows sequencing traces from two different SNPs that are heterozygous in genomic DNA, but homozygous in cDNA from P175 cells, again indicating that this intergenic region is mono-allelically expressed. In total, we have detected mono-allelic expression of 4 heterozygous SNPs and 29 additional non-polymorphic locations between *MANEA* and the *lox*P-3'RT integration site in P175 cells (Appendix 1, Table A1.2). Additional PCR and sequencing reactions from HTD114 and the chromosome 6 mono-chromosomal hybrids indicated that the chromosome that contains the *lox*P-3'RT integration site in P175 contains the expressed allele for this entire region, spanning positions 96,171,771 to 96,373,551 of chromosome 6, representing >200 kb of genomic DNA (Fig. 2.4E; Appendix 1, Table A1.2). In addition, cells containing the smallest deletion that causes DRT, ~76 kb, lack expression from the deleted region as expected, but retain expression of the mono-allelic transcripts proximal to the deletion (Fig. 2.4D), indicating the existence of transcriptional start sites within the non-deleted region. Expression of this intergenic region was also detected in R175 cells (Fig. 2.4D; Appendix 1, Table A1.2), indicating that the translocation event that generated the t(6;10) does not interfere with expression of this region. As expected, the chromosome 6 deletions that do not result in DRT, i.e. distal to the *lox*P-3'RT integration site,

also do not interfere with expression of these mono-allelic transcripts (Fig. 2.4D). Furthermore, RNA-DNA FISH analysis showed mono-allelic expression of this intergenic region (Fig. 2.4E-G; Appendix 1, Fig. A1.27). Because this intergenic region also displays asynchronous replication (see what follows), we have named these non-coding transcripts asynchronous replication and autosomal RNA on chromosome 6, *ASAR6*.

Consistent with our observations of mono-allelic expression at this locus, a previous study, using the density of LINE-1 elements to identify mono-allelically expressed genes, predicted that the *FUT9* and *FHL5* genes (located ~180 kb and ~730 kb distal to the *lox*P integration site, respectively; Fig. 2.4E) would be mono-allelically expressed [87]. Therefore, we next determined whether *FUT9* and *FHL5* were also mono-allelically expressed in our cell system. For this analysis we used allele-specific RT-PCR of heterozygous SNPs (Fig. 2.5A), RT-PCR followed by sequencing of heterozygous SNPs located within introns (Fig. 2.5B; Appendix 1, Table A1.2), RT-PCR using primers spanning exons in mono-chromosomal hybrids (Fig. 2.5C), and RNA-DNA FISH for expression of primary transcripts (Fig. 2.5E and F; Appendix 1, Figs A1.28 and A1.30). This analysis indicated that the *FUT9* gene is indeed mono-allelically expressed. In addition, we detected mono-allelic expression from within and distal to the *FHL5* gene (Appendix 1, Table A1.2 and Fig. A1.26A). However, exon-specific primers, which spanned large introns, failed to generate RT-PCR products for *FHL5* (Appendix 1, Fig. A1.26B), indicating that properly spliced *FHL5* protein-coding transcripts were not expressed. In addition, the transcripts detected from the

*FHL5* region were generated from the antisense orientation (Appendix 1, Fig. A1.26C). In addition, both *FUT9* and these *FHL5* opposite strand transcripts (*FHL5OST*) were not expressed from the chromosome 6 that contains the *lox*P-3'RT cassette (Appendix 1, Table A1.2). Therefore, the chromosome 6 rearrangements that cause DRT/DMC were generated on the chromosome 6 that expresses *ASAR6*, but is silent for both *FUT9* and *FHL5OST*. In addition, analysis of the parental HTD114 cells indicated that *FUT9* is also mono-allelic, and that HTD114 and P175 cells express the same allele, indicating that integration of the original *lox*P-3'RT cassette was not responsible for mono-allelic expression of this region in P175 cells (Appendix 1, Table A1.2). In contrast, the *MANEA* and *KIAA0776* genes are bi-allelically expressed in HTD114 and P175 cells (Appendix 1, Table A1.2 and Fig. A1.32). A similar analysis of other known or predicted mono-allelically expressed genes present on chromosome 6 indicated either no or bi-allelic expression in P175 cells (Appendix 1, Table A1.5).

We next determined whether the mono-allelic expression from this region of chromosome 6 was due to imprinting or random mono-allelic inactivation. For this analysis, we examined non-clonal and clonal cell lines derived from lymphoblast cultures for expression of SNPs located within *ASAR6*. Figure 2.5G-I shows that both alleles of *ASAR6* were detected in the RNA of the non-clonal line, and that either one or the other allele was detected in the clonal lines. This analysis indicated that *ASAR6* is subject to random mono-allelic expression in these lymphoblastoid cells.

## Asynchronous replication on Chromosome 6

Because mono-allelically expressed genes display asynchronous replication (reviewed in [84]), we next tested whether the region of chromosome 6 between *MANEA* and *FHL5* also displays asynchronous replication. Loci can be identified as replicating either synchronously or asynchronously using a FISH-based assay [109]. FISH analysis of interphase nuclei pulse-labeled with BrdU allows selective examination of cells in S-phase. This assay also utilizes a methanol/acetic acid fixation, which destroys the nuclear structure and allows for a relatively accurate analysis of replication timing [110-111]. Using a probe to a particular chromosomal site, some cells display two single-hybridization dots, indicating that neither allele has replicated (an SS pattern); other cells display two double dots, indicating that both alleles have replicated (a DD pattern); and a third class of cells contains one single dot and one double dot, indicating that only one of the two alleles (an SD pattern) has replicated (Appendix 1, Fig. A1.36). Asynchronously replicating loci show the SD pattern in 30–50% of S-phase cells, whereas synchronously replicating genes typically present the SD pattern in only 10–20% of S-phase cells [110, 99-100]. We used primary human skin fibroblasts and this FISH assay to quantify the number of hybridization signals per locus present in S-phase nuclei. We found that five different probes from this region of chromosome 6, spanning ~1 Mb of genomic DNA including *MANEA, ASAR6, FUT9*, BAC959I6 and *FHL5*, as well as four additional chromosome 6 mono-allelically expressed genes (*HTR1E, KCNQ5, ME1*, and *FRK*) located ~10-20 Mb from *ASAR6*, all show asynchronous replication (Fig. 2.6A-C; Table 2.1).

Previous studies have shown that the random mono-allelically expressed genes present on a given autosome are coordinated in their asynchronous replication so that the alleles on one homolog replicate earlier than the alleles on the other [99-100]. Therefore, we next tested whether the asynchronously replicating loci described above also display coordination in their asynchronous replication. The level of coordination was examined using a two-color FISH assay and scoring cells that simultaneously displayed the SD signal for both loci [99-100]. For this analysis we used a BAC probe representing *ASAR6* in combination with probes representing *MANEA, FUT9*, BAC959I6 and *FHL5* (Fig. 2.6A; Table 2.1). We found that the asynchronous replication of *ASAR6* was coordinated in *cis* with *FUT9* (78/100 cells, $P < 1 \times 10\text{-}5$; Fig. 2.6D-F), *FHL5* (74/100 cells $P < 1 \times 10\text{-}5$; Fig. 2.6G-I), BAC959I6 (70/100 cells, $P < 1 \times 10\text{-}4$; Appendix 1, Fig. A1.37A-C), and *MANEA* (78/100 cells, $P < 1 \times 10\text{-}5$; Appendix 1, Fig. A1.37D-F).

Because asynchronous replication on human autosomes is coordinated at the whole chromosome level [100], we also assayed other mono-allelically expressed genes located either centromeric or telomeric to *ASAR6* (Fig. 2.6A). Interestingly, we found that the asynchronous replication of *ASAR6* was also coordinated with these other chromosome 6 asynchronous genes; however the coordination was in *trans*. Thus, the earlier replicating allele of *ASAR6* was linked to the later replicating alleles of *ME1* (86/100 cells, $P < 1 \times 10\text{-}5$; Fig. 2.6J-L), *HTR1E* (76/100 cells, $P < 1 \times 10\text{-}5$; Fig. 2.6M-O), *KCNQ5* (74/100 cells, $P < 1 \times 10\text{-}5$; Appendix 1, Fig. A1.37G-I), and *FRK* (74/100 cells, $P < 1 \times 10\text{-}5$; not shown).

## Disruption of mono-allelic expression

We next tested whether the genetic alterations in chromosome 6 that cause DRT also affect mono-allelic expression of *FUT9* and *FHL5OST*. For this analysis, we used RNA-DNA FISH to assay their expression in R175 and Δ175-1i [containing the original t(6;10) or an ~230 kb proximal deletion, respectively]. Figure 2.7A-D shows that these genetic alterations result in bi-allelic expression of both *FUT9* and *FHL5OST* (Appendix 1, Figs A1.28-35). Taken together, our data indicate that mono-allelic transcription of *ASAR6* is occurring only on the chromosome 6 that is 'inactive' for *FUT9* and *FHL5OST*, and that the Cre-mediated rearrangements that result in DRT also result in expression of the previously silent *FUT9* and *FHL5OST* alleles in *cis*. A schematic illustration of the location of the *ASAR6* transcripts, the location of the smallest deletion that causes DRT, and the activation of the previously silent *FUT9* allele are shown in Figure 2.7E and F.

## Discussion

The observations described here are consistent with previous reports showing that, like X chromosomes, autosome pairs display coordination in their asynchronous replication timing of mono-allelically expressed genes [99-100, 112]. Interestingly, deletion of the *Xist* gene from adult somatic cells results in a delayed replication phenotype [82-83] that appears to be similar to the DRT phenotype described for certain autosomal rearrangements [8-10]. In this report, we show that inter-chromosomal translocations or deletions at a specific location on human chromosome 6 result in delayed replication of the entire chromosome. These data define a *cis*-acting locus responsible for chromosome-wide replication timing of a human autosome. Furthermore, because the vast majority of chromosome 6 genes are not subject to mono-allelic expression and asynchronous replication, our data indicate that disruption of the *ASAR6* locus delays replication timing of both asynchronously and synchronously replicating genes along the entire chromosome.

Although parallels between X-inactivation and the coordinated replication asynchrony of autosomal mono-allelically expressed genes have been made [99-100, 84, 113, 112, 85], a previous report showed that not all random mono-allelically expressed alleles are expressed from the same homolog [86]. These observations suggest that although autosomal asynchronous replication timing is coordinated in *cis* [99-100, 112], mono-allelic expression is not [86, 114]. However, it should be pointed out that the asynchronous replication timing of these ‘non-coordinated’ mono-allelically expressed genes was not assayed. In

this report, we show that *ASAR6* shows asynchronous replication that is coordinated in *cis* with the mono-allelically expressed genes *FUT9* and *FHL5OST*, even though *ASAR6* is expressed from the opposite homolog. In addition, we found that *ASAR6* shows coordination in asynchronous replication with other mono-allelically expressed genes located at a distance along chromosome 6, but this coordination was in *trans*. Moreover, the apparent mono-allelic expression observed with certain transcripts can arise via mechanisms that may not be associated with asynchronous replication timing (reviewed in [115]). Therefore, a detailed analysis of the coordinated replication timing in combination with allele-specific expression assays is required to determine whether the 'non-coordinated' expression pattern is a characteristic of certain genes or is a characteristic of all random mono-allelically expressed genes located on autosomes.

The deletions that cause DRT disrupt the *ASAR6* RNA and result in bi-allelic expression of the previously silent alleles of the mono-allelically expressed genes *FUT9* and *FHL5OST*. However, in contrast to *XIST* RNA, *ASAR6* RNA does not appear to coat the entire chromosome. Interestingly, *ASAR6* transcripts can be detected in human embryonic stem cells (E.P.S. and M.J.T., unpublished data; also see reference [108]), but are not expressed in all adult tissues (Appendix 1, Fig. A1.38). In contrast, *Xist* remains expressed on the inactive X in most if not all adult tissues (reviewed in [116]. However, *Xist* is apparently not absolutely required for the maintenance of X chromosome silencing, as *Xist* deletion after X-inactivation does not automatically result in global X-reactivation [117-118].

Therefore, it will be interesting to determine whether disruption of the *ASAR6* locus in cells that do not express *ASAR6* RNA, also results in delayed replication. We believe that this is a distinct possibility, as disruption of the silent *Xist* gene causes a delay in the replication timing of the active X chromosome [82].

Previously, we used ionizing radiation to generate chromosome rearrangements in mouse and human cells, and found that ~5% of inter-chromosomal translocations involving autosomes display DRT/DMC [9]. Furthermore, our original 'chromosome-engineering' screen for DRT/DMC identified five balanced translocations, involving eight different autosomes [10]. In more recent studies, we have generated a set of nested deletions in chromosome 15, anchored at the original *lox*P integration site in the P268 cell line (see Appendix 1, Fig. A1.2 and reference [10]), and found that the deleted 15s also display DRT (N.D., L.S. and M.J.T., unpublished data). On the basis of these observations, we propose that all mammalian autosomes contain discrete *cis*-acting loci that function to regulate chromosome-wide replication timing.

## Materials and Methods

### Cell culture

HTD114 cells are a human *APRT*-deficient cell line derived from the HT1080 fibrosarcoma [119]. This cell line and all of its subclones, including P175, the alternative partner clones (R-lines), and the deletion clones (Δ-lines), were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (Hyclone). HTD114 derivatives were grown as above with the addition of 500 mg/ml Geneticin (Gibco), 200 mg/ml Hygromycin B (Calbiochem), and/or 10 µg/ml Blasticidin S HCl (Invitrogen). Lentiviral infections were carried out, using a multiplicity of infection of <0.1, according to standard procedures (Invitrogen). The HTD114 R-line and Δ-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 for Aprt-expressing cells. APRT counter-selection was carried out in 10µg/ml diaminopurine (Sigma). Low-passage primary human skin fibroblasts were obtained from ATCC and cultured in DMEM plus 10% fetal bovine serum (Hyclone). The lymphoblast cell line GM13130 was obtained from the Coriell Institute and cultured in RPMI (GIBCO) media supplemented with 15% fetal bovine serum (Hyclone). Clones of GM13130 were obtained by limiting dilution in 96-well dishes, expanded, and harvested for DNA and RNA. Cells were grown in a humidified incubator at 37°C in a 5% carbon dioxide atmosphere.

## DNA Fluorescence *in situ* hybridization (FISH)

Trypsinized cells were centrifuged at 1000 r.p.m. for 10 min in a swinging bucket rotor. The cell pellet was re-suspended in 75 mM potassium chloride for 15-30 min at 37°C, re-centrifuged at 1000 r.p.m. for 10 min 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 [120]. Slides with mitotic spreads were baked at 85°C for 20 min and then treated with 0.1 mg/ml RNAase for 1 h at 37°C. After RNAase treatment, the slides were washed in 2x SSC (1x SSC is 150 mM NaCl and 15 mM sodium citrate) with three changes for 3 min each and dehydrated in 70%, 90%, and 100% ethanol for 3 min each. The slides were denatured in 70% formamide in 2x SSC 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 100 µg/ml for

1h at 37°C and washed in 2x SSC and dehydrated in 70%, 90% and 100% ethanol. Chromosomal DNA was denatured at 75°C for 3 min in 70% formamide/2x SSC, 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/µl. Centromeric probe cocktails (Vysis) plus BAC or Fosmid DNAs were denatured at 75°C for 10 min and prehybridized at 37°C for 30 min. Probes were applied to slides and incubated overnight at 37°C. Post-hybridization washes consisted of three 3 min rinses in 50% formamide/2x SSC, three 3 min rinses in 2x SSC, and finally three 3 min rinses in PN buffer (0.1 M $Na_2HPO_4$ + 0.1 M $NaH_2PO_4$, pH 8.0, +2.5% Nonidet NP-40), all at 45°C. Signal detection was carried out as described [121]. 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 (2.5 µg/ml) or DAPI (15 µg/ml) and viewed under UV fluorescence (Olympus).

## RNA-DNA FISH

Cells were plated on microscope slides at ~45% confluence and incubated overnight in complete media in a 37°C humidified $CO_2$ incubator. Slides were rinsed one time with sterile RNase free PBS. Slides were incubated for 30 s in CSK buffer (100 mM NaCl, 300 mM Sucrose, 3 mM MgCl2,10 mM Pipes, pH 6.8), 5 min in CSK buffer plus 0.1% Triton X-100, and then for an additional 30 s in CSK buffer at room temperature. Cells were fixed in 4% paraformaldehyde in PBS for 10 min 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 30 min. Slides were hybridized at 37°C for 14-16 h. Slides were washed as follows: three times in 50% formamide/2x SSC at 42°C for 5 min, three times in 2x SSC at 42°C for 5 min, three times in 4x SSC/0.1% Tween-20 at room temperature for 3 min. Slides were then fixed in 4% paraformaldehyde in PBS for 5 min at room temperature, and briefly rinsed in 2x SSC 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 (2.5 µg/ml) or DAPI (15 µg/ml) and viewed under UV fluorescence (Olympus). Z-stack images were generated using the Cytovision workstation.

## Replication timing assay

The BrdU replication timing assay was performed on exponentially dividing cultures as follows: asynchronously growing cells were exposed to 20 µg/ml of BrdU (Sigma) for 4.5, 5, 6, 7, 8 or 9 h. Mitotic cells were harvested in the absence of colcemid, treated with 75 mM KCl for 15-30 min 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 2x SSC at 70°C for 3 min and processed for DNA FISH, as described above. The incorporated BrdU was then detected using an 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.

### PCR and expression analysis

Genomic DNA and total RNA were isolated from HTD114 cells and its sub-clones using TRIZOL Reagent (Invitrogen). Total human RNA samples were from the Ambion FirstChoice® Human Total RNA Survey Panel. cDNA was prepared using the SuperScript™ III First-Strand Synthesis System (Invitrogen). Reverse transcriptase reactions were performed in the presence or absence of reverse transcriptase on 5 µg of total RNA. PCR (genomic and RT-PCR) was performed in a 25-50 µl volume using 50-100 ng of genomic DNA or 1-2 µl of cDNA (50-100 ng of input RNA equivalent), 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 35-45 cycles of 95°C for 30 s, 55-62°C for 45 s, 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. Sequencing of PCR products was carried out at the Vollum Institute DNA Sequencing Core facility.

## Acknowledgements

We thank Dr. Michael Harkey, at the Fred Hutchinson Cancer Research Center's Clonal Analysis Core, for the LAM-PCR analysis and Dr. Phil Soriano for the Flp expression construct. We are grateful to Dr. Mark Groudine, Dr. David Kabat and Dr. Sarah Smolik for critical comments on the manuscript. We are also

grateful to Dr. Andrew Chess for providing the genotypes of the GM13130 lymphoblast cell line. We are grateful to DNASTAR Inc. for the gift of the Lasergene DNA sequence analysis software. E.P.S. was supported by a National Science Foundation Graduate Research Fellowship, an ARCS Foundation Scholarship, a Tartar Trust Fellowship, and a Vertex Fellowship. N.D. was supported by the Pre Doctoral Training Grant in Molecular Hematology, T32 HL00781. This work was supported by grants from the National Cancer Institute, CA104693 and CA131967, to M.J.T.

## Erratum

This chapter is slightly modified from the version published in Human Molecular Genetics. The term Bacterial Artificial Chromosome was added to explain the abbreviation BAC and F-factor cosmid cloning vectors was added to explain the term Fosmid. The text that accompanies Table 2.1 was added for this dissertation.

## Figures and Tables

**Figure 2.1. Alternative partner analysis.** (**A**) Illustration of the *lox*P integration sites, chromosomes 6 (red) and 10 (green), and balanced translocation, t(6;10), in P175/R175. The random integration of a new *lox*P cassette is expected to integrate into a third chromosome (purple) and generate a new Cre-dependent translocation. (**B**) BrdU terminal labeling procedure. BrdU is added to the media for increasing times (green arrows) and cultures are harvested for mitotic spreads and processed for FISH and BrdU incorporation. (**C** and **D**) Cells containing a t(6;9) were incubated with BrdU for 4.5 h and processed for BrdU incorporation (green) and for FISH using probes for chromosome 6 (CHR 6 paint, red) and chromosome 9 centromere (CHR 9 cen, pseudo-colored purple) (**D**). The inset shows an enlarged image of the chromosome 9 derivative. Chromosomes were stained with DAPI. (Figure on next page).

**Figure 2.1. Alternative partner analysis.**

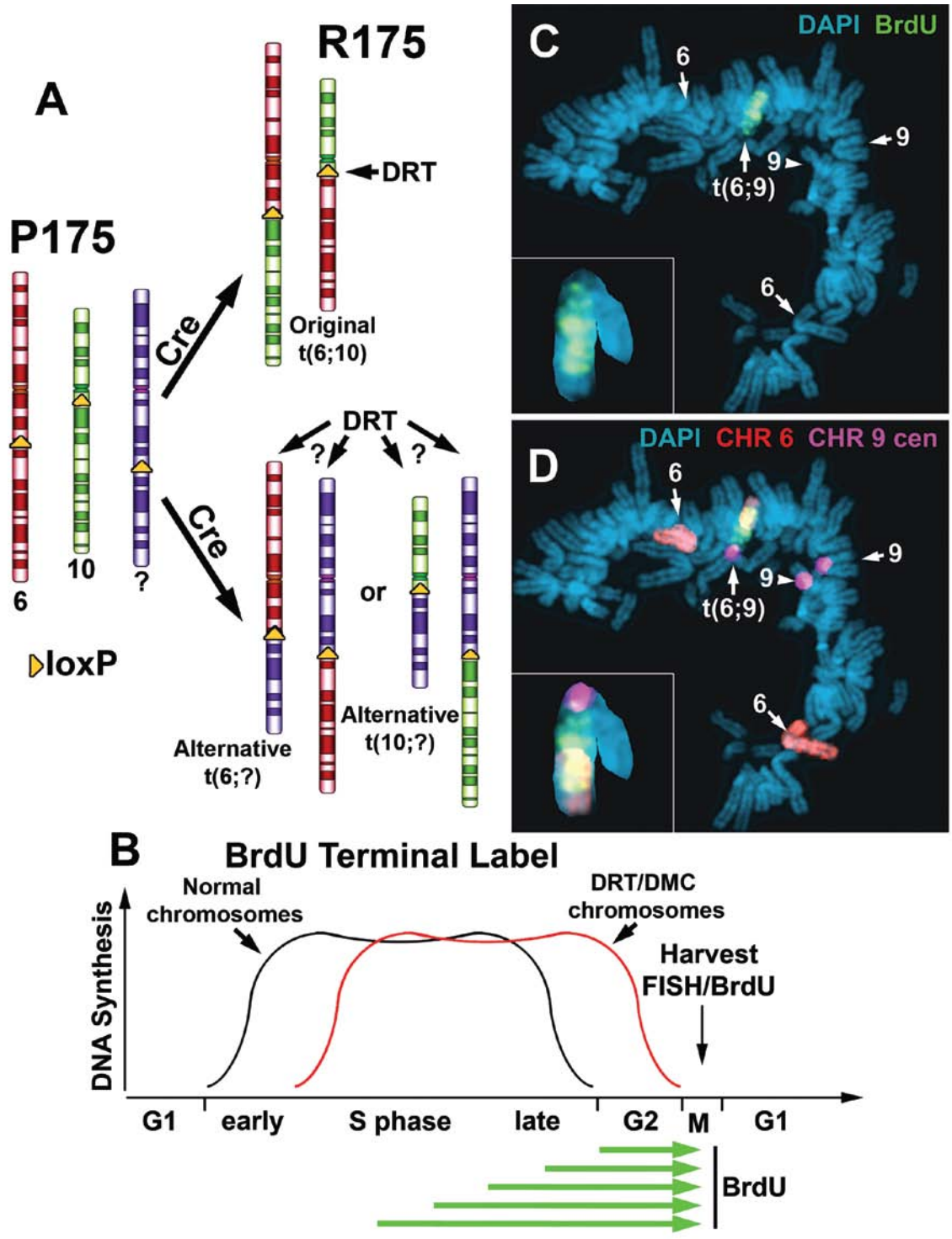


**Figure 2.2. Chromosome 6 proximal deletions.** (**A**) Diagram of chromosome 6 showing the orientation and integration site (in megabases) of the original *lox*P-3'RT cassette, structure of the AP-*lox*P lentivirus, and the locations of the 5'LTR-genome junctions for 15 lentiviral integration sites (arrows), BACs RP11-374I15 (374I15) and CTD-84D21 (84D21), Fosmid G248P86031E7 (G248E7), and two protein-coding genes, *MANEA* and *FUT9*. (**B**) Deletion of BAC 374I15 from one copy of chromosome 6 (~2.1 Mb deletion). Mitotic cells were processed for FISH using probes for the chromosome 6 centromere (green), BAC 374I15 (red), and a second BAC 84D21 (green). (**C**) Replication timing assay on an ~16.58 Mb deletion. Cells were incubated with BrdU for 4.5 h, harvested and processed for BrdU incorporation (green) and FISH, using probes for the chromosome 6 centromere (red) and BAC 374I15 (red). Chromosomes from three independent mitotic spreads are shown. Chromosomes were stained with DAPI (blue). (Figure on next page).

**Figure 2.2. Chromosome 6 proximal deletions.**

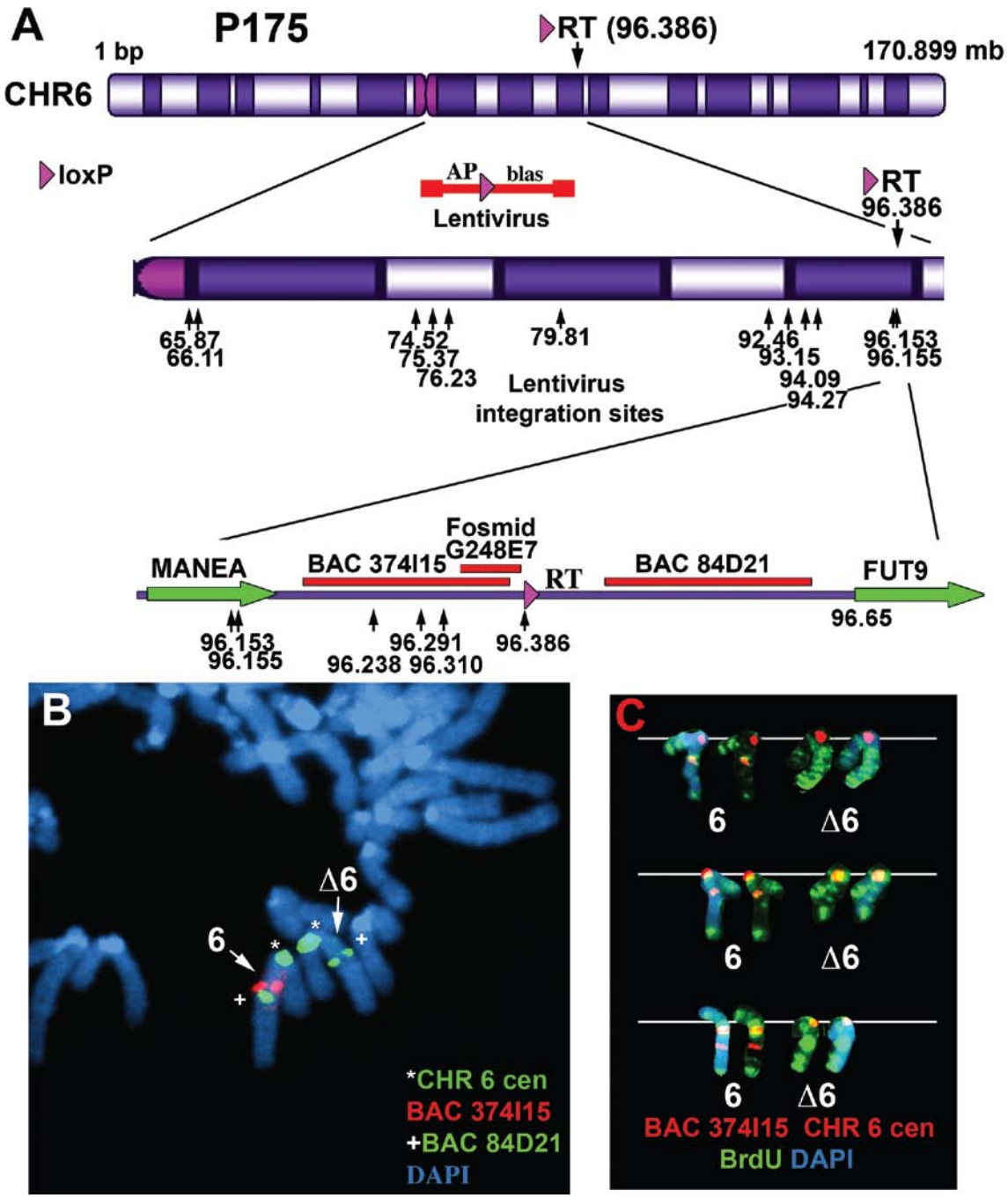
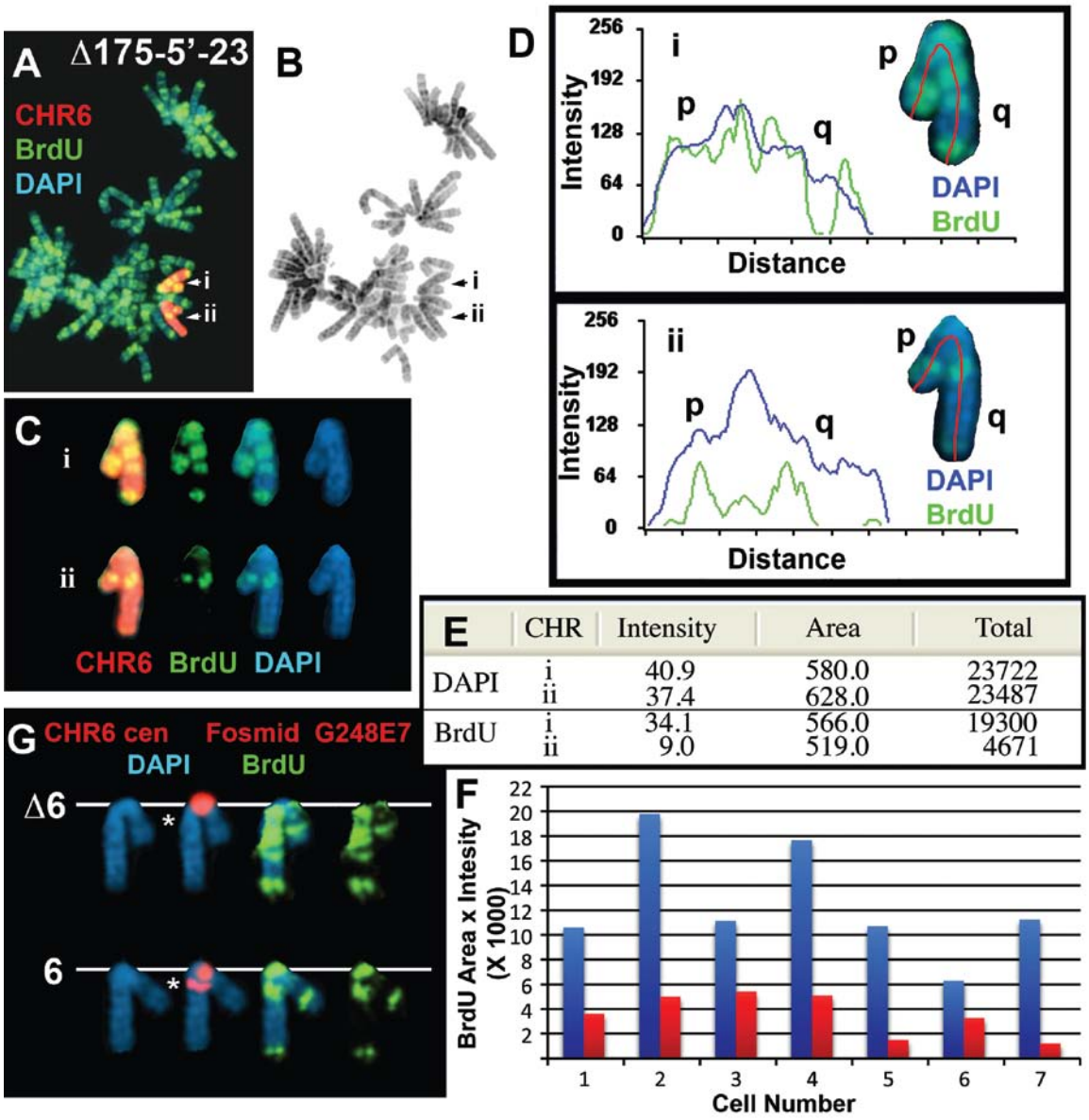


**Figure 2.3. Delayed replication timing on an ~76 kb deletion.** Cells were incubated with BrdU for 4.5 h, harvested, and processed for FISH, using either a chromosome 6 paint (**A**-**E**; red) or a chromosome 6 centromeric probe (red) plus Fosmid G248E7 (red), and DAPI (blue) (**G**). (**B**) Reverse DAPI banding of the chromosomes shown in (**A**). (**C**) The two chromosome 6s from (**A**), were cut out and aligned. (**D**) Pixel-intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the two chromosome 6s from (**A**). (**E**) The pixel-intensity (average intensity x area) for each chromosome showing the total amount of BrdU incorporation or DAPI staining. (**F**) Quantification of the BrdU incorporation in seven different pairs of chromosome 6s. The red and blue bars represent the two chromosomes identified by the chromosome 6 paint. (Figure on next page).

**Figure 2.3. Delayed replication timing on an ~76 kb deletion.**




| E | CHR | Intensity | Area | Total |
|---|---|---|---|---|
| DAPI | i | 40.9 | 580.0 | 23722 |
| | ii | 37.4 | 628.0 | 23487 |
| BrdU | i | 34.1 | 566.0 | 19300 |
| | ii | 9.0 | 519.0 | 4671 |

**Figure 2.4. Mono-allelic expression at chromosome 6q16.1.** (**A**) Allele-specific RT-PCR from the intergenic region between *MANEA* and the *lox*P-3'RT integration site. PCRs, using forward primers containing single-base differences (allele G or A; see Appendix 1, Table A1.3) at their 3' ends, plus a common reverse primer, for SNP rs4645429. PCRs were carried out on genomic DNAs isolated from P175, two mono-chromosomal hybrids containing the two different chromosome 6s from HTD114 [L(Hyg)-1 contains the *lox*P-3'RT integrated chromosome 6, and L(Neo)-38 contains the other chromosome 6], and either with (+RT) or without (-RT) reverse transcriptase on P175 RNA. (**B** and **C**) DNA-sequencing traces from PCRs designed to detect SNPs rs4645429 and rs6934550. The top panels show the traces from P175 DNA, and the bottom panels show the traces from P175 cDNA (RNA). The arrows mark the location of the SNPs. (**D**) Mono-allelic expression of the intergenic region in mono-chromosomal hybrids and in the chromosome 6 deletions. RT-PCRs were carried out on RNA isolated from P175, L(Hyg)-1, L(Neo)-38, ∆175-23a (~76 kb proximal deletion), ∆175F-11c (~25.9 Mb distal deletion) and R175 [t(6;10)]. Genomic DNAs from each cell line served as positive controls. PCRs were carried out using primers designed to detect SNP rs4629688 (located within the smallest deletion), rs4840081 (located proximal to the smallest deletion), and rs9374586 (located within *MANEA*). Sequencing of the PCR products confirmed that all primers amplified the correct chromosome position (not shown). (**E**) Schematic diagram of the region of chromosome 6 between *MANEA* and *FHL5* showing the location of the smallest proximal deletions (arrows), the three Fosmids used as

probes in the RNA FISH (green), and BAC959I6 used in the DNA FISH (red) in (**F**) and (**G**). (**F**) RNA-DNA FISH using Fosmid G248P86031A6 (6031A6). (**G**) The number of single (blue) and double (red) sites of RNA expression. Two hundred cells were scored for RNA FISH signals that were associated with the DNA FISH signal for three Fosmid probes [see (**E**) and Appendix 1, Fig. A1.27). (Figure on next page).

**Figure 2.4. Mono-allelic expression at chromosome 6q16.1.**

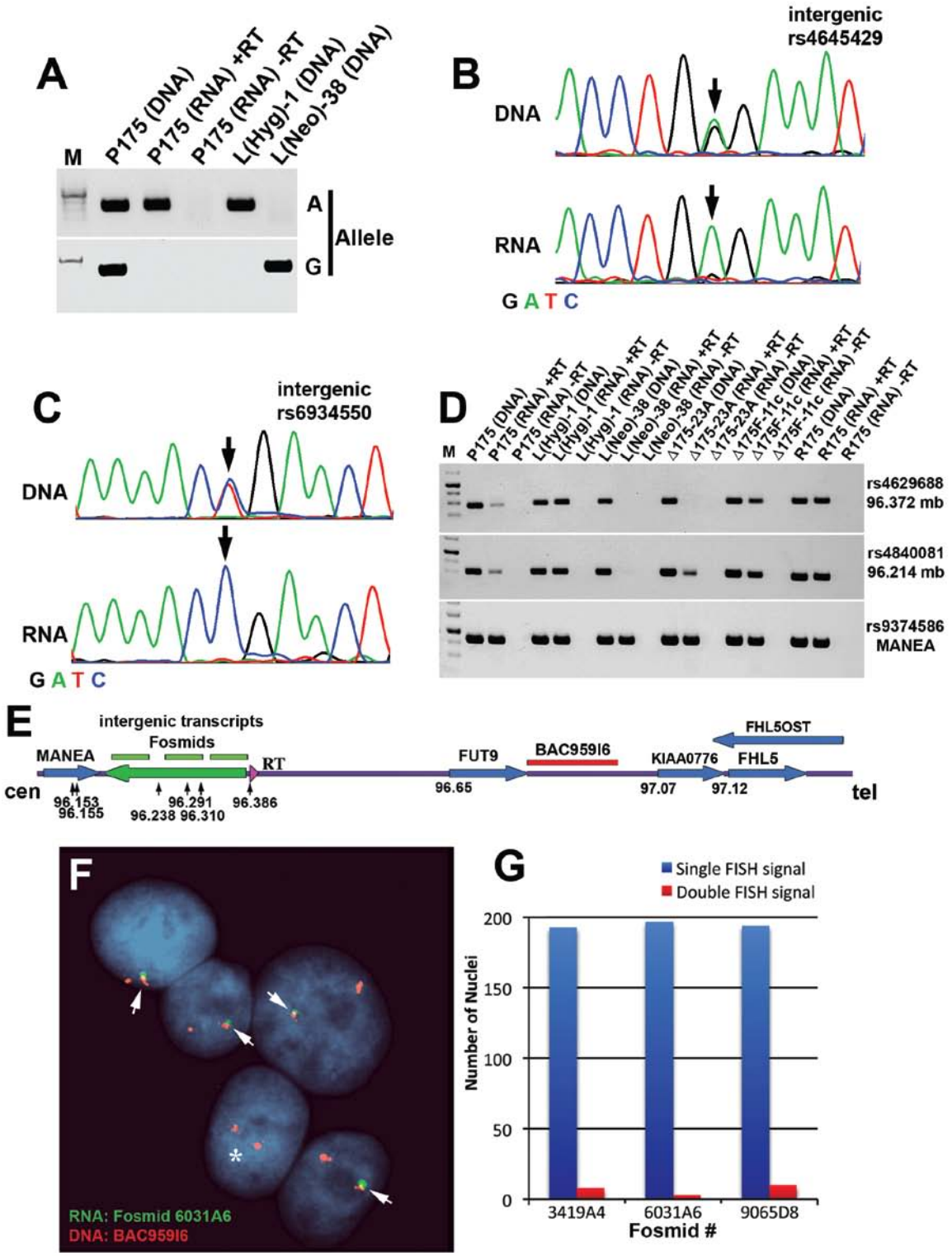


**Figure 2.5. Mono-allelic expression of FUT9 and random mono-allelic expression of ASAR6.** (**A**) Allele-specific RT-PCR of a SNP located in the first intron of *FUT9*. PCR using forward primers containing single-base differences (allele G or C; see Appendix 1, Table A1.3) at their 3' ends, plus a common reverse primer, for SNP rs4002794. PCRs were carried out on genomic DNAs from P175, the two mono-chromosomal hybrids [L(Hyg)-1 and L(Neo)-38], either with (+RT) or without (-RT) reverse transcriptase on P175 RNA. (**B**) DNA-sequencing traces from PCRs designed to detect SNP rs4380761, located in the second intron of *FUT9*. The top panel shows the trace from P175 DNA, and the bottom panel shows the trace from P175 cDNA (RNA). The arrows mark the location of the SNP. (**C**) Exon-specific RT-PCR for expression of *FUT9*. PCRs were carried out, with primers designed to detect exons 2 and 3 from the spliced *FUT9* cDNA, either with (+RT) or without (-RT) reverse transcriptase, on RNA from P175 and the mono-chromosomal hybrids L(Hyg)-1 and L(Neo)-38. PCR products derived from a cloned *FUT9* cDNA indicated correctly spliced products from P175 and L(Neo)-38. PCR products were not detected from P175 DNA, because the intron spanned by these two primers is >90 kb. (**D**) Exon-specific RT-PCR for expression of *MANEA*. PCRs, using primers designed to amplify exons 2 and 3 from the *MANEA* cDNA, either with (+RT) or without (-RT) reverse transcriptase, on RNA from P175 cells and the mono-chromosomal hybrids L(Hyg)-1 and L(Neo)-38. Sequencing reactions confirmed the exon junctions for properly spliced *MANEA* cDNA (not shown). (**E**) RNA-DNA FISH for expression of *MANEA* and *FUT9*. P175 cells were subjected to RNA FISH (green) using a

cDNA probe for *MANEA* (left panel), a cDNA probe for *FUT9* (middle panel), and an intronic probe for *FUT9* (right panel). Slides were subsequently fixed and processed for DNA FISH (red) using BAC 959I6, located distal to *FUT9* (Fig. 2.4E). Arrows mark the sites of expression, and the nuclear DNA was detected with DAPI. RNA FISH signals associated with DNA FISH signals were scored. (**F**) Quantification of the single (blue) and double (red) sites of RNA FISH signals for *MANEA* and *FUT9*. Error bars indicate the error of the standard mean. (**G**) Random mono-allelic expression of *ASAR6*. Non-clonal and clonal lymphoblastoid (GM1310) cell lines were assayed for expression of *ASAR6*. DNA sequence traces from PCRs designed to detect SNP rs6912187. The top panels show the traces from non-clonal DNA and cDNA (RNA). The bottom panels show the traces from DNA and cDNA (RNA) isolated from four independent clones (A12, A13, A24, and A26) isolated from GM13130. The arrows mark the location of the SNP. (Figure on next page).

**Figure 2.5. Mono-allelic expression of *FUT9* and random mono-allelic expression of *ASAR6*.**

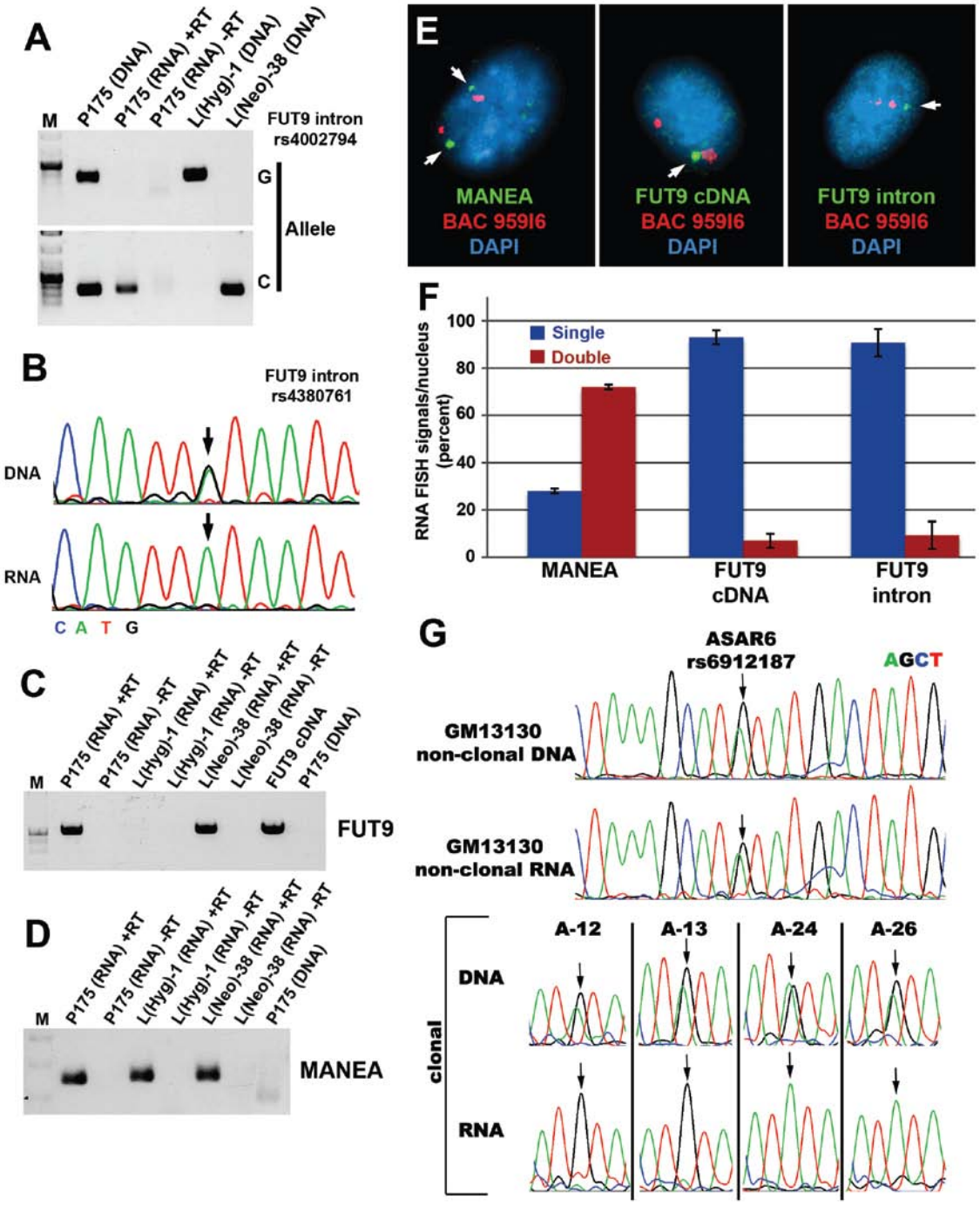


**Figure 2.6. Asynchronous replication of random mono-allelically expressed genes on chromosome 6.** (**A**) Schematic diagram of chromosome 6 showing the location of the genes and loci assayed for asynchronous replication. The ~1 Mb region of chromosome 6 between *MANEA* and *FHL5* is expanded on the right. The coordination in asynchronous replication of chromosome 6 mono-allelically expressed genes with *ASAR6* was found to be either in *cis* or in *trans* as indicated (**A**). Examples of the SD pattern in single nuclei for *ASAR6* (**B**) and *FUT9* (**C**). The BrdU staining is not shown. (**D**-**O**) Coordination in the asynchronous replication timing along chromosome 6. Low-passage primary human skin fibroblasts were analyzed using a two-color FISH assay combining BAC probes for *ASAR6* (green) with BAC or Fosmid probes for *FUT9, FHL5, ME1*, and *HTR1E* (red; see Appendix 1, Table A1.4 for probes). DNA was stained with DAPI. Arrows mark the location of the FISH signals. Dashed white lines show the outline of the nuclei. Single dots (S) represent un-replicated loci and double dots (D) represent replicated loci. The percentage of cells that simultaneously displayed the SD pattern for both probes was determined in 100 cells for the *cis* versus *trans* configuration. (Figure on next page).

**Figure 2.6. Asynchronous replication of random mono-allelically expressed genes on chromosome 6.**

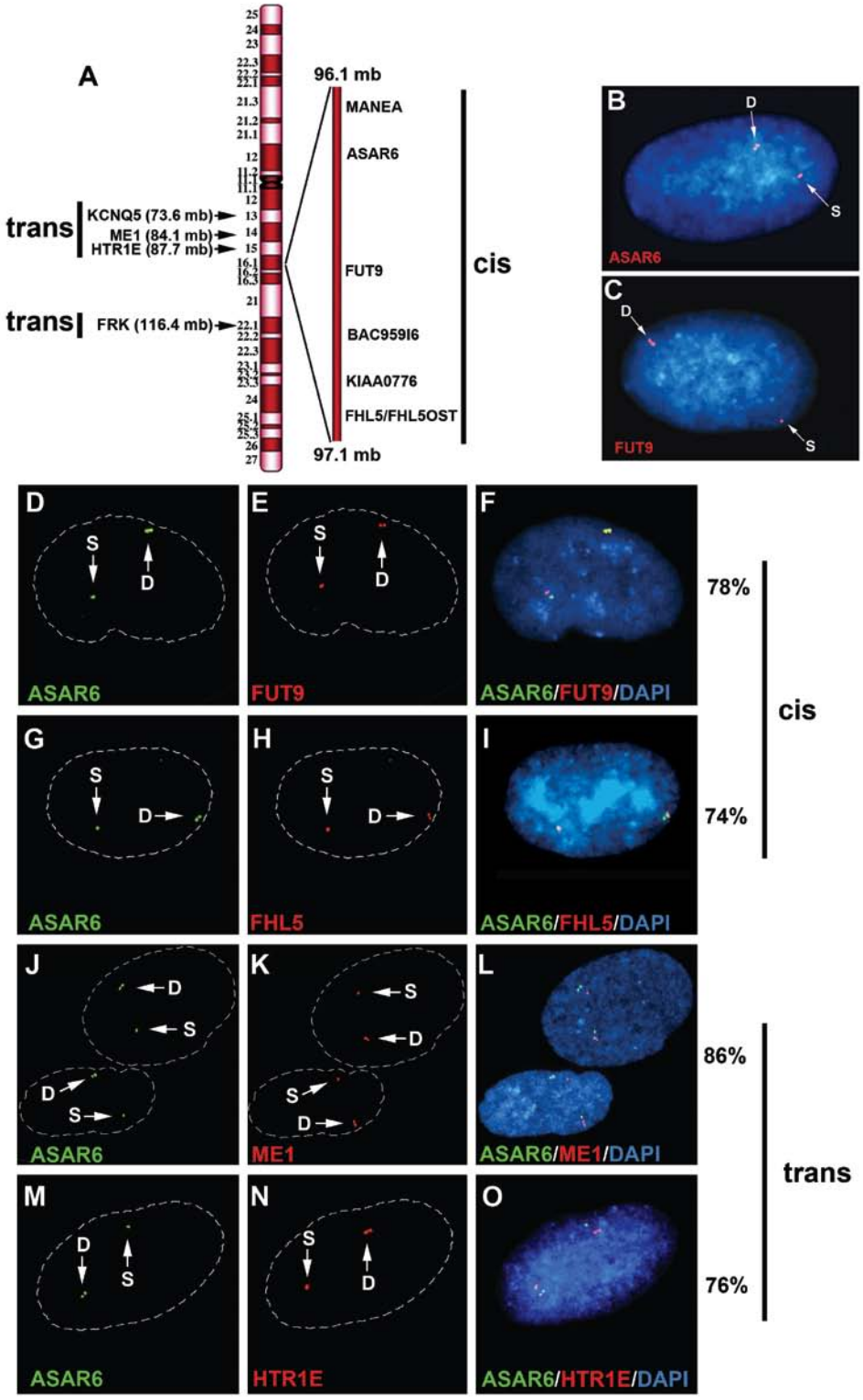
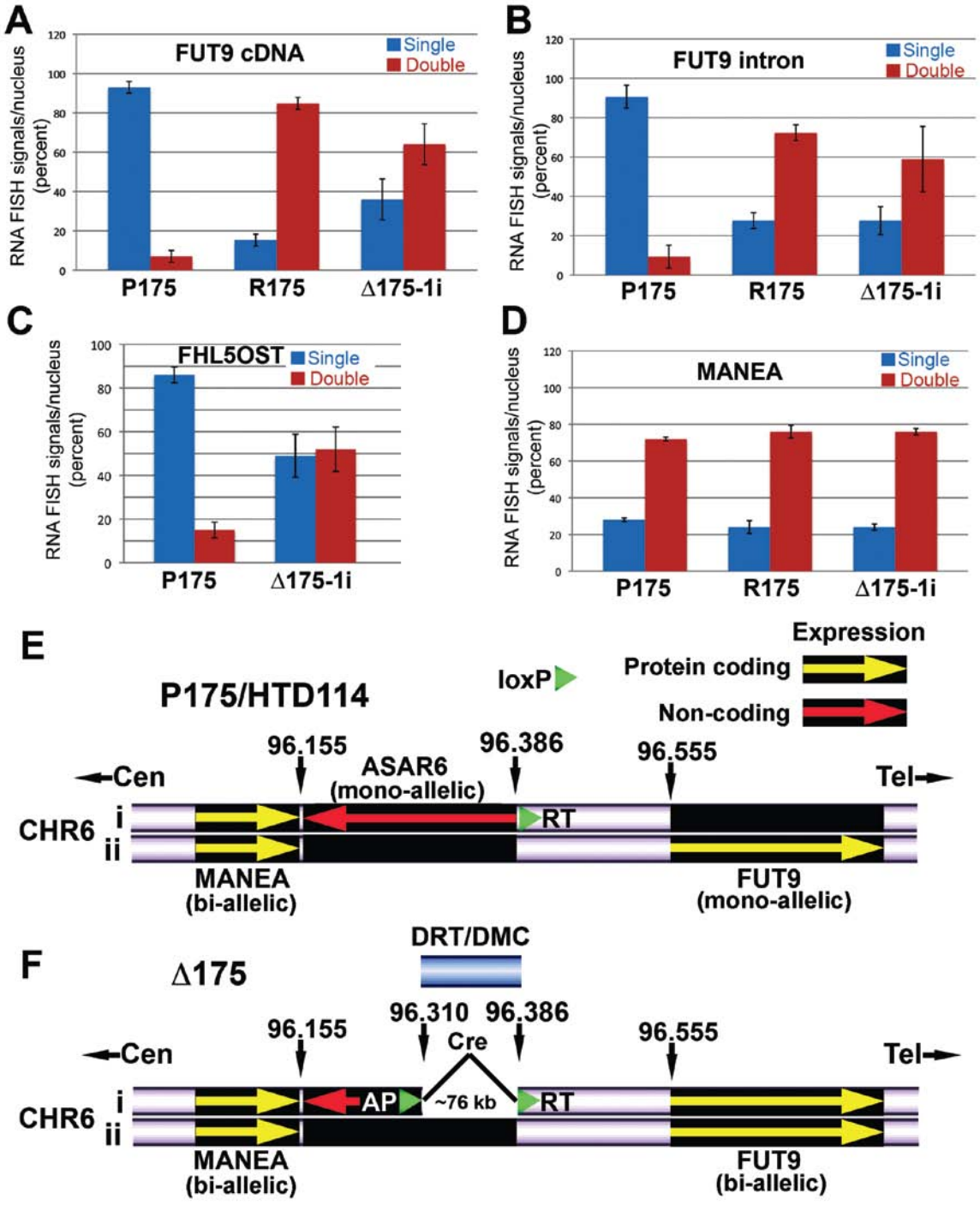


**Table 2.1. FISH analysis of autosomal loci.** The percentage of the single-double (%SD) pattern was determined using FISH. The five FISH probes located at 6q16.1 are within the ~1 Mb region of the *lox*P-3'RT integration site on chromosome 6. The four other probes from chromosome 6 are mono-allelically expressed genes located a distance away from this region. The remaining FISH probes were used as controls for synchronously replicating loci.

| Locus | Position | %SD |
|---|---|---|
| MANEA | 6q16.1 | 36 |
| ASAR6 | 6q16.1 | 43 |
| FUT9 | 6q16.1 | 37 |
| BAC959I6 | 6q16.1 | 44 |
| FHL5 | 6q16.1 | 34 |
| HTR1E | 6q15 | 33 |
| KCNQ5 | 6q13 | 33 |
| ME1 | 6q14.1 | 39 |
| FRK | 6q22.1 | 34 |
| PTK6 | 20q13.3 | 19 |
| LARP | 5q33.2 | 24 |
| C9orf43 | 9q32 | 19 |

**Figure 2.7. Loss of mono-allelic expression at 6q16.1.** (**A**-**D**) Bi-allelic expression of previously mono-allelically expressed genes in R175 and Δ175-1i. Slides were processed for RNA-DNA FISH and examined for sites of RNA hybridization using *FUT9* cDNA (**A**), *FUT9* intron (**B**), *FHL5OST* Fosmid (G248P86054G4) (**C**), or *MANEA* cDNA (**D**) probes. The DNA probe was BAC 959I6 (Fig. 2.4E). The percentage of cells (P175, R175 and Δ175-1i) from three independent hybridizations, that displayed single (blue bars) or double (red bars) sites of RNA hybridization are indicated (also see Appendix 1, Figs A1.28-35). Error bars indicate the error of the standard mean. (**E**) Schematic diagram of the region of the long arm of chromosome 6 near the *lox*P-3'RT integration site in P175. Both chromosome 6s are indicated (**i** and **ii**) showing expression of protein coding genes (yellow arrows) and non-coding regions (red arrow). (**F**) Schematic diagram of the same region of chromosome 6 shown in (**E**). The smallest Cre/*lox*P-mediated deletion that we have generated that induces DRT is ~76 kb. (Figure on next page).

**Figure 2.7. Loss of mono-allelic expression at 6q16.1.**



# CHAPTER THREE

A rAAV-mediated targeted deletion on chromosome 6

## Introduction

In the study presented in Chapter 2 and Appendix 1 of this dissertation, we found that a deletion as small as ~76 kb on chromosome 6, at 6q16.1, resulted in a chromosome-wide delay in replication timing and loss of mono-allelic expression, in *cis*. This ~76 kb locus was identified by making translocations and deletions on chromosome 6, anchored at a *lox*P site previously integrated randomly into this locus [10] (see Appendix 1 for detailed methods of this analysis). Recombination between the two *lox*P sites following transfection of a Cre recombinase-expressing plasmid results in deletion of the intervening sequence and reconstitution of a selectable marker (*Aprt*). Despite generating a large number (>50) of random deletions at this locus, I was unable to isolate any that were smaller than ~76 kb. The goal of this experiment was to identify the minimal deleted region that could induce the DRT phenotype by making smaller, targeted deletions at this locus on chromosome 6. I used a recombinant adeno-associated virus (rAAV) approach to make an ~46 kb targeted deletion of chromosome 6.

For both medical and research applications, much effort has gone into finding better ways to deliver gene-therapy vectors into the genome. Initially, random approaches were used; however, these approaches came with a variety of problems, ranging from unpredictable and potentially mutagenic integration of the vector to lack of control over how many integrations occurred and how much expression was achieved [122]. Researches switched to targeted methods of

delivery because these methods offer predictability in the site and number of integrations and in the level of expression of the inserted gene. Unfortunately, these methods require a great deal more effort in constructing the delivery vectors, and they have a much lower efficiency of genome integration than random approaches.

The problem of efficient integration of vectors arises because targeted events require the use of homologous recombination (HR) between the delivery vector and the target sequence in the genome. This is difficult to achieve in human cells due to the size of the human genome (over three billion bp, making it difficult for the targeting-vector to find the homologous region of DNA), and to the low efficiency of HR in human cells. HR occurs infrequently in mammalian somatic cells, with an observed frequency of only 1 in $10^6$-$10^7$ [122-123]. Conversely, random integration is a more efficient process because mammalian cells prefer to use non-homologous end-joining (NHEJ) as a method of DNA repair, which requires no homology. Therefore, vectors randomly integrate via NHEJ in mammalian cells at a frequency closer to 1 in $10^3$ (see [124-125] for reviews of NHEJ and HR, respectively).

A 1998 report indicated that recombinant AAV can integrate in mammalian cells via HR at a much higher frequency than previous targeting strategies dependent on HR [126]. AAV is a single-stranded DNA virus from the parvovirus family that primarily infects humans (see [122] for a review of gene targeting with AAV). The AAV2 subtype has no known pathology associated with it, and has thus been used often for research purposes. The AAV genome is 4.7 kb and

contains inverted terminal repeats (ITRs) that form hairpin structures. Recombinant versions of AAV have been developed that have the *rep* and *cap* open reading frames (ORFs) removed, which encode proteins needed for replication and packaging of the virus, respectively. Removal of these ORFs creates space in which to insert a gene or sequence of interest to be targeted to a specific DNA sequence in the genome, and also prevents replication of the virus. Targeting is achieved by constructing 5' and 3' homologous sequence arms on each end of the sequence to be inserted. The *rep* and *cap* viral proteins can then be provided in *trans* in order to allow production of the rAAV.

Numerous experiments have been done to gather information on the optimal means of gene targeting and gene repair using rAAV [122]. Due to the size limitation imposed on construction of rAAV vectors [commercially available kits (Stratagene) allow ~3 kb of space to insert homologous sequence of interest], experiments have been done to look at both the minimal and optimal size of homologous targeting necessary to achieve correct and efficient targeting [127], as well as the optimal levels of virus to use in infections [127-128, 122-123]. While homologous arms of 150 bp on the 5' end and 500 bp on the 3' end were sufficient for correct targeting, increasing the length of homologous sequence arms significantly increases targeting efficiency. Increasing the amount of virus used to infect cells (the MOI) led to an increase in random integration of the virus (up to 10% of cells had non-targeted integrations), but also increased targeting efficiency (correct targeting events occurred in up to 1% of cells) [122]. Therefore, the higher the viral titer, the more targeting events occur, but the

background of non-targeted integrations also increases. For this experiment, non-targeted integrations are not problematic, so we used a high MOI in order to increase our chances of recovering targeted integrations.

In these experiments, I used a modified version of our 5'AP-*lox*P-Blas construct [see Appendix 1 for details about this construct] to make a targeted deletion at a specific locus on chromosome 6, ~46 kb from the original *lox*P-3'RT integration site (See Fig. 3.1A for a schematic of the rAAV-targeting strategy). This location was chosen because it was approximately half-way between the smallest randomly generated deletion we had at the time and the *lox*P-3'RT integration site on chromosome 6. The 5'AP-*lox*P-Blas construct was designed so that infected cells will give Blasticidin-resistance, but following Cre transfection, the *Blas* gene is deleted and the *Aprt* gene is re-constituted (Fig 3.1A). This strategy allows for the selection of Aprt+ clones that have undergone recombination between *lox*P sites, followed by screening for loss of Blas-resistance to identify clones that made a deletion rather than another type of rearrangement. My goal was to generate this targeted deletion and then test the deletion-containing clone for the DRT phenotype.

## Results

### Generation of a rAAV-targeting vector for chromosome 6 deletions

An rAAV vector was generated that contained the 5'AP-*lox*P-Blas construct as well as left and right arms with sequence homology to a position on chromosome 6 that is ~46 kb away from the *lox*P-3'RT integration site (Fig. 3.1A). The homologous arms were designed to a region of unique sequence with no repeat elements (96,339,539-96,340,672 on chromosome 6 [NCBI build 36/hg18]) and were generated by PCR. Virus produced from this construct was capable of integrating into P175 DNA and generating Blas-resistant cells. The rAAV construct had a functional 5'AP cassette, as evidenced by the fact that Aprt-positive clones were generated following transient transfection with a Cre-expressing plasmid. These data indicate that this rAAV-targeting vector (AAV-5'AP-*loxP*-Blas) is functional and that the AAV Helper-Free System (Stratagene) can incorporate the vector plus homologous sequence arms.

### rAAV-targeted deletions on chromosome 6

P175 cells were infected with 1, 10, or 100 µl of AAV-5'AP-*lox*P-96.33 virus and selected in Blas media. Each of the eleven pools of Blas-resistant cells was transiently transfected with a Cre-expression plasmid and underwent selection for Aprt+ clones. Two independent clones originating from the same pool (infected with 100 µl of virus) were isolated that were Blas-sensitive and showed loss of the 3'RT plasmid-genome junction (Fig. 3.1B shows an example analysis

for one of these two clones), indicating that they contained chromosome 6 deletions. Additionally, these clones retained heterozygosity of a SNP ~67 kb from the integration site, meaning they were small deletions (smaller than ~67 kb) (Fig. 3.1B). To confirm that these clones contained properly targeted deletions, they were subjected to Southern blot hybridization. Each clone was confirmed to be a properly targeted ~46 kb deletion (Fig. 3.1C shows a Southern blot for one of these properly targeted deletion clones). Finally, deletions were also confirmed by the absence of a signal from a Fosmid probe, G248P86031E7 (G248E7), located within the deleted region, by DNA FISH. A non-deleted chromosome 6 retains the Fosmid signal, while a deleted chromosome 6 loses the signal. Figure 3.2A shows this analysis for one of the deletion clones.

## Replication timing of rAAV-targeted deletions on chromosome 6.

We next assayed the replication timing of the chromosome 6s in these two deletion clones. Cultures were incubated with BrdU for 5 h and mitotic cells were harvested, processed for BrdU incorporation and subjected to FISH analysis using chromosome 6-specific probes. We combined the BrdU incorporation assay with FISH using a chromosome 6 centromeric probe plus Fosmid G248E7 from the deleted region. The FISH signal from the chromosome 6 centromeric probe allowed us to identify both chromosome 6s, and the presence or absence of the Fosmid G248E7 probe allowed us to distinguish between the non-deleted or deleted chromosomes, respectively. Comparing the BrdU incorporation pattern of the deleted chromosome 6s to the non-deleted chromosome 6s from both

targeted-deletion clones indicated that the chromosomes containing the ~46 kb deletion do not appear to display DRT. Figure 3.2 shows representative replication timing data from one of the two targeted-deletion clones.

## Discussion

The experiments described in this chapter showed that it is possible to adapt our 5'AP-*lox*P-Blas lentiviral deletion strategy for use with a rAAV system in order to make targeted rearrangements on chromosomes. The rAAV vector is size-limited, but by shortening the AP-cassette, we were able to clone in sufficient homology for targeting by the rAAV. The proper-targeting of the rAAV achieved in these experiments provides proof-of-principle for the use of rAAV to target our *lox*P constructs to specific sites in the genome. This will allow us to make deletions or rearrangements at precise locations. This strategy may be useful for making targeted deletions on other chromosomes that show DRT/DMC. Random deletions have worked well as a method for collecting a large number of deletions, but it can be difficult to screen for and isolate deletions in a specific size range. rAAV-mediated targeting could also be used for placing *lox*P constructs in specific locations on chromosomes in order to test if a particular region shows DRT/DMC if disrupted. We can also use this strategy to make the same deletions or rearrangements in other cell types to test for any cell type-specific differences in induction of DRT/DMC.

We used a rAAV strategy to make a targeted ~46 kb deletion on chromosome 6. The deletion was confirmed by genome-junction PCR, LOH analysis, Southern

blot hybridization (Fig. 3.1B and C), and DNA FISH (3.2A). The ~46 kb deletion was tested for a delay in replication timing, and it was determined that this deletion was not sufficient to induce the DRT phenotype (Fig. 3.2B). We have previously shown that an ~76 kb deletion at this locus is sufficient to induce the DRT/DMC phenotype (Chapter 2; Appendix 1). The data from these two deletions indicate that there may be an important genetic element within the 30 kb that differentiates those deletions, or that there is a size threshold for the deletions in order to induce a delay in replication timing that is severe enough for accurate detection.

## Methods

### Cell culture

P175 cells [10] are a derivative of HTD114, which are a human *APRT*-deficient cell line derived from the HT1080 fibrosarcoma cell line [119]. All of the experiments described in this appendix were carried out with P175ΔNH, and are referred to as simply P175 throughout this section [see Appendix 1 for an explanation of the difference]. P175ΔNH3 cells, AAV-293 cells, mouse L cells and L-cell hybrid cells were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (Hyclone), except under low serum conditions during viral production, when serum was 2-3%. P175 derivatives infected with rAAV were grown with 10 µg/ml Blasticidin S HCl (Invitrogen). The P175 cell line, R-line, Δ-line, and ΔAAV-line derivatives were grown in AA media: DMEM supplemented with 10% dialyzed fetal bovine serum (Hyclone), 10 mg/ml azaserine (Sigma) and 10 mg/ml adenine (Sigma) to facilitate selection for Aprt+ cells. Cells were grown in a humidified incubator at 37°C in a 5% $CO_2$ atmosphere.

### Vector construction

The 5'AP-*lox*P-Blas construct was initially cloned into the pBluescript KSII+ vector (Stratagene). The 5'-AP cassette was modified from that used in the lentiviral construct (see Appendix 1 for details); this version has no introns, a C-T mutation in the splice donor site, and a bovine Growth Hormone polyadenylation signal (bHGpA). This enables a much smaller construct, thereby leaving space to

include homologous targeting sequences and still fit the entire construct into the rAAV cloning vector. Homologous sequence arms corresponding to 96,338,434-96,339,539 (1,106-bp) and 96,339,540-96,340,672 (1,133-bp) on chromosome 6 [NCBI Build 36/hg18] were generated by PCR with KpnI or SacII overhangs, respectively, on the PCR primers (KpnI arm: F-TTTCGGTACCTGCCTCTTATGTTCCTGTCATTC, R-TTTCGGTACCTAAAAGTCATCTGGAGACCTCTGG; SacII arm: F-TTTCCCGCGGTCCTGTTATACGGGTCAGTTCC, R-TTTCCCGCGGCACAGAAAGAGGGTCATGTGAG). PCR products were gel purified using standard methods. These homologous arms were then cloned into the KpnI or SacII restriction sites in the 5'AP-*lox*P-Blas KSII+ vector. The entire 5'AP-*lox*P-Blas vector including the homologous arms was cut out of the KSII+ backbone with BssHII, and cloned into the pAAV-MCS vector (Stratagene), which was cut with KpnI and blunted. Colonies were plated under ampicillin + Blas selection to prevent recombination events occurring between the rAAV ITRs. Correct orientation of the rAAV-5'AP-*lox*P-Blas-96.33 construct was confirmed by restriction digest analysis and DNA sequencing with the M13r primer. Sequencing of plasmid products was carried out at the Vollum Institute DNA Sequencing Core facility.

## Production and titering of rAAV

AAV-5'AP-*lox*P-Blas virus was produced in AAV-293 cells (AAV Helper-Free System, Stratagene) which express the adenovirus E1 gene necessary for vial production. Viral production was carried out using a modified version of the

provided manual (Stratagene). The AAV-5'AP-*lox*P-Blas plasmid was co-transfected along with the pHelper and pAAV-RC plasmids, each containing *trans* factors needed for viral production, into AAV-293 cells. Ten 15 cm dishes of AAV-293 cells, ~70-80% confluent, were transfected with 16 µg of each plasmid in 2.5 ml of .3M CaCl2 and 2.5 ml of 2x HBS. After 6 h, the media was replaced with low serum media (2-3% FBS). Virus was harvested after 3 days.

For viral harvest, 0.5 ml of EDTA was added to each dish, incubated for 15 min, the cells were tapped loose, and then cells and media from all of the dishes were collected in a 500-ml centrifuge bottle. The bottle was spun for 15 min at 3000 r.p.m. to pellet cells. Supernatant was taken off and the pellet was resuspended in 3.5 ml of Benzonase buffer (50 mM Tris, pH 8.5, 2 mM $MgCl_2$) and transferred to a 15-ml Falcon tube. The cells were then exposed to three freeze-thaw cycles using a dry-ice ethanol bath to freeze and a 37°C water bath to thaw, 10 min at each step. After the second freeze-thaw, 200 U/ml of Benzonase enzyme (Sigma) was added before proceeding to the third freeze-thaw cycle. Following the final thaw, the lysate was incubated at 37°C for 1 h. The lysate was then spun at 3,000 r.p.m. for 15 min to pellet the cell debris. The supernatant was transferred to a new tube and 1/39$^{th}$ volume of 1 M $CaCl_2$ was added. Viral lysate was kept on ice for 1 h and then spun at max speed for 30 min at 4°C. Supernatant was then transferred to new tubes and stored at -80°C.

To determine the titer of the virus, we infected P175 cells with serial dilutions of the virus. The cells were infected, re-fed with complete media after 24 h, and then plated in Blas media in 10 cm dishes. Blas-resistant colonies were counted

after 8 days and the titer was determined to be ~2x$10^6$/ml for Blas-resistant colonies.

## Viral infection

P175 cells were infected with 1, 10, or 100 µl of rAAV lysate, as above. Pools of Blas-resistant cells were pooled and transfected with a Cre-expression plasmid. 24 h after Cre transfection, cells were re-fed with complete DMEM media. 24 h later, cells were plated on 15 cm dishes in AA media to select for Aprt+ clones.

## Identifying targeted deletions

Aprt+ clones were screened for loss of Blas-resistance. Clones that lost Blas-resistance were grown and processed for DNA. Clones were then screened by PCR using primers designed to amplify the plasmid-genome junctions of both the 5' and 3' ends of the *lox*P-3'RT integration cassette. Clones that lost the 3'RT plasmid-genome junction and Blas-resistance were thus confirmed to contain deletions in the correct orientation on chromosome 6. Primers for the 3'RT junction were as follows: forward 5'-GCCAGGAATTAAGTCAAGAAGG-3', and reverse 5'-GTATTACCGCCTTTGAGTGAGC-3' primers. Primers for the 5'-RT junction were: forward 5'- TGTGCTACTGGAGATCAATAAACC-3' and reverse 5'-AGGAGTGGCCTTTTTCTTTTTC-3'. Clones were then screened to look for small deletions. Using a heterozygous SNP located ~67 kb from the integration site, clones that had lost the 3'RT junction, but retained heterozygosity at that SNP using the following allele-specific primers: “A allele” forward – 5'-

AAAGGCTAGTTTAGTAATTCAGACA-3', “G allele” forward – 5'-AAAGGCTAGTTTAGTAATTCAGACG-3' and the common reverse primer – 5'-CTGGTCCTCATCCAGAGCTTAC-3' were known to contain deletions smaller than ~67 kb. Clones that showed LOH were deletions larger than ~67 kb and therefore not targeted deletions. Clones confirmed to contain deletions smaller than ~67 kb were subjected to Southern blot hybridization to determine whether they were targeted deletions. ΔAAV clone DNA, along with P175, R175, a mouse L-cell hybrid containing the *lox*P-3'RT chromosome 6 (see Appendix 1), and the mouse L-cell DNA, were digested with SacI and SpeI restriction enzymes (NEB). A Southern probe located outside of the targeted region (96,336,161-96,336,758 base pairs on chromosome 6) was generated with the following PCR primers: forward – 5'- TTGTTGGTTCACGTTTACATGC-3' and reverse – 5'-TACCCATAAACCCTGCCTACAG-3'. Deletion clones were also confirmed by the absence of a Fosmid signal by DNA FISH (see below).

## PCR analysis and generation of homologous arms

Genomic DNA was isolated from HTD114 cells and its sub-clones . PCR was performed in a 25-50 µl volume using 50-100 ng of genomic DNA, 1x Standard Taq Buffer (New England Biolabs, Inc.), 200 µM of each deoxynucleotide triphosphate, 0.2 µM of each primer, and 3 U of Taq DNA Polymerase (New England Biolabs, Inc.) [except for the generation of homologous arms, which used 2 U of the PfuTurbo DNA Polymerase (Stratagene)] under the following reaction conditions: 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 55-60°C for 45 s, 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.

## DNA FISH

Trypsinized cells were centrifuged at 1000 r.p.m. for 10 min in a swinging bucket rotor. The cell pellet was re-suspended in 75 mM potassium chloride for 15-30 min at 37°C, re-centrifuged at 1000 r.p.m. for 10 min 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 [120]. Slides with mitotic spreads were baked at 85°C for 20 min and then treated with 0.1 mg/ml RNAase for 1 h at 37°C. After RNAase treatment, the slides were washed in 2x SSC (1x SSC is 150 mM NaCl and 15 mM sodium citrate) with three changes for 3 min each and dehydrated in 70%, 90%, and 100% ethanol for 3 min each. The slides were denatured in 70% formamide in 2x SSC 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 and Fosmid probes***: Mitotic chromosome spreads were prepared as described above. Slides were treated with RNase at 100 µg/ml for 1h at 37$^{O}$C and washed in 2x SSC and dehydrated in 70%, 90% and 100% ethanol. Chromosomal DNA was denatured at 75°C for 3 min in 70% formamide/2x SSC, followed by dehydration in ice-cold 70%, 90% and 100% ethanol. Fosmid DNA was nick-translated using standard protocols to incorporate biotin-11-dUTP or digoxigenin-dUTP (Invitrogen). Fosmid DNA was 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/µl. Centromeric probe cocktails (Vysis) plus Fosmid DNA [Fosmid G248P86031E7 (G248E7)], were denatured at 75°C for 10 min and prehybridized at 37°C for 30 min. Probes were applied to slides and incubated overnight at 37°C. Post-hybridization washes consisted of three 3-min rinses in 50% formamide/2x SSC, three 3-min rinses in 2x SSC, and finally three 3-min rinses in PN buffer (0.1 M $Na_2HPO_4$ + 0.1 M $NaH_2PO_4$, pH 8.0, +2.5% Nonidet NP-40), all at 45°C. Signal detection was carried out as described [121]. 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 (2.5 µg/ml) or DAPI (15 µg/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 µg/ml of BrdU (Sigma) for 5 h. Mitotic cells were harvested in the absence of colcemid, treated with 75 mM KCl for 15-30 min 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 2x SSC at 70°C for 3 min 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.

## Figures

**Figure 3.1. Chromosome 6 rAAV-targeted deletions.** (**A**) Diagram of chromosome 6 showing the orientation and integration site (in megabases) of the original *lox*P-3'RT cassette, structure of the rAAV-5'AP-*lox*P-Blas-96.33, including the locations of the homologous arms used for targeting, and the locations of the two nearest protein-coding genes, *MANEA* and *FUT9*. (**B**) PCR screen for small proximal deletions on chromosome 6. PCR was done for both the proximal and distal genome-plasmid junctions at the 3'RT cassette integration site (**i**). DNA from P175 shows both junctions intact. Δ175-2c is an ~21.9 Mb proximal deletion and shows loss of the proximal junction. Δ175F-11a is an ~1.4 Mb distal deletion and shows loss of the distal junction. ΔAAV33-1b, -1d, and -7b are Aprt+ clones that were isolated following infection with the rAAV-5'AP-*lox*P-Blas-96.33 virus, selection for Blas-resistance, were pooled, and then transiently transfected with a Cre-expression plasmid. Clones ΔAAV33-1b and 7b retained the proximal junction, while ΔAAV33-1d lost the proximal junction, indicating a deletion. These clones were simultaneously screened for LOH at an SNP ~67 kb proximal to the 3'RT integration site (**ii**). P175 shows both alleles present, the mouse L cell hybrids, L(Hyg)-1 and L(Neo)-38, show only the A or the G allele, respectively. All three of the ΔAAV33 clones show both alleles present, confirming that ΔAAV33-1b and -7b were not deletions (since they also retained the proximal junction), and that ΔAAV33-1d is a deletion smaller than ~67 kb (it has lost the proximal junction but retains both alleles at ~67 kb). (**C**) Southern

blot confirming the AAV-targeted deletion. The target locus is shown (**i**) with the location of the SacI and SpeI sites, the rAAV-5'AP-*lox*P integration site, and the location of the probe. Digestion of genomic DNA with SpeI and SacI results in a 6.2 kb fragment following a properly targeted deletion. The Southern blot (**ii**) shows the band sizes for P175, ΔAAV33-1d (a properly targeted deletion), R175 (containing the 6;10 translocation), the hybrid cell line L(Hyg)-1 and the L cell parental line. (Figure on next page).

**Figure 3.1. Chromosome 6 rAAV-targeted deletions.**

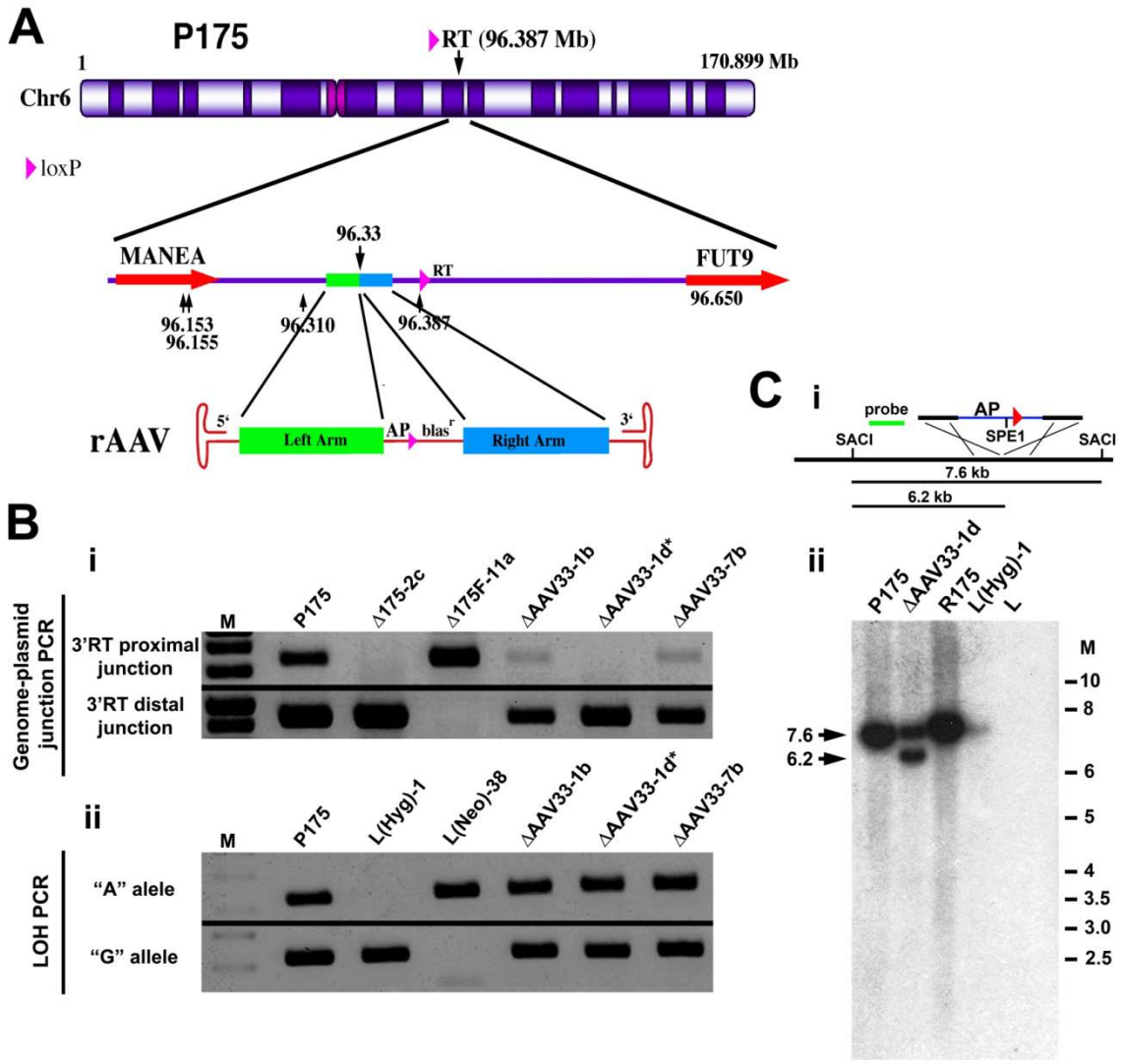
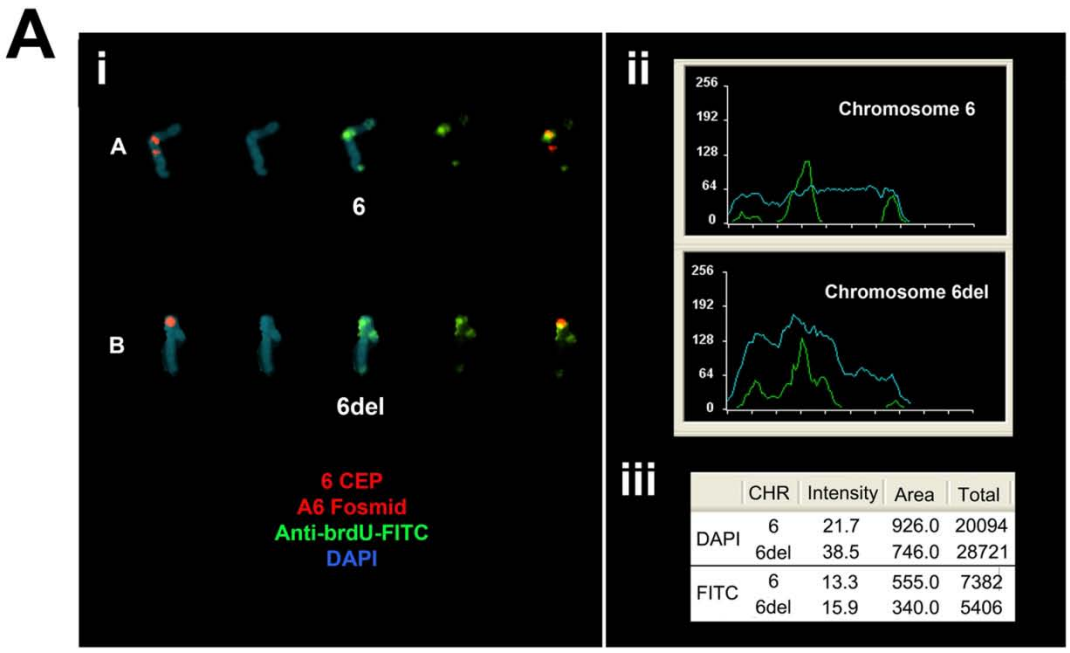
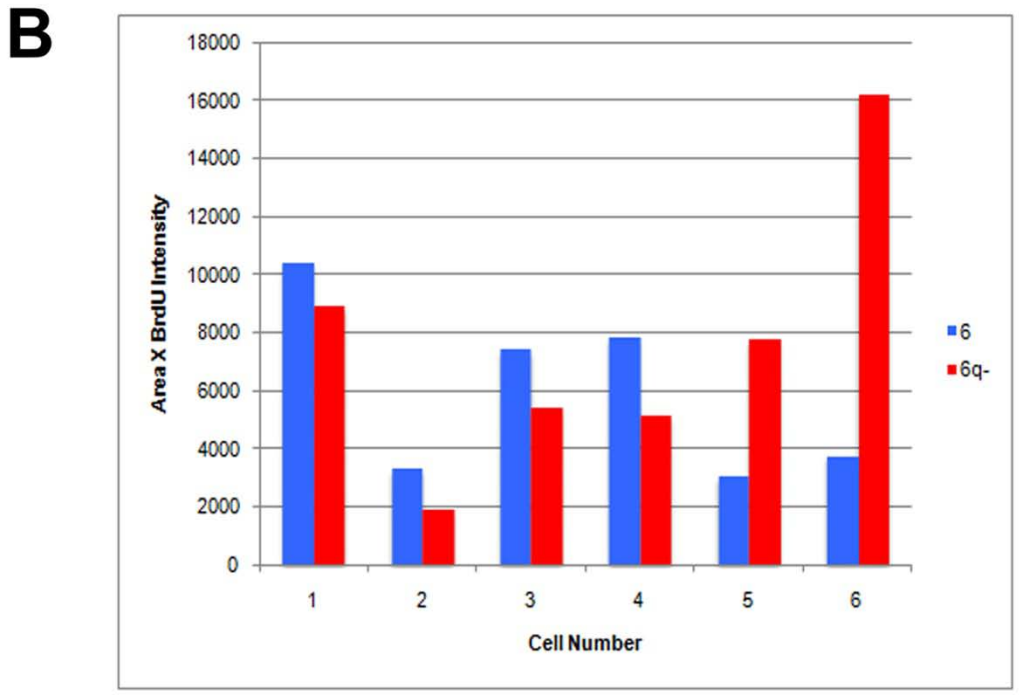


**Figure 3.2. Delayed replication timing assay on the ~46 kb targeted deletion (96,339,540).** (**A**) Cells were incubated with BrdU for 5 h, harvested and processed for FISH using a chromosome 6 centromeric probe (red) plus Fosmid G248E7 (red) and DAPI (blue). The two chromosome 6s were cut out and aligned (**i**). Pixel intensity profiles of the BrdU incorporation (green), and DAPI (blue) staining along the two chromosome 6s from panel i (**ii**). The pixel intensity (average intensity x area) for each chromosome showing the total amount of BrdU incorporation or DAPI staining is calculated (**iii**). (**B**) Quantification of the BrdU incorporation in seven different pairs of chromosome 6s. The red and blue bars represent the two chromosomes identified by the chromosome 6 paint; the blue bars represent the chromosome 6 containing Fosmid G248E7 (the non-deleted chromosome 6), while the red bars represent the chromosome that lacks the Fosmid [the ~46 kb targeted-deletion (6del, 6q-)]. (Figure on next page).

**Figure 3.2. Delayed replication timing assay on the ~46 kb targeted deletion (96,339,540).**



# CHAPTER FOUR

Conclusions

## Summary

The data presented in this dissertation detail a genetic analysis of the DRT/DMC phenotype in a cell line that displays DRT/DMC on one derivative chromosome following a translocation between chromosomes 6 and 10 [10]. The experiments described here explain how it was determined that a locus on chromosome 6, at 6q16.1, is important for the DRT/DMC phenotype (Figs 2.1 and 2.2). We show that disruptions of this locus by translocation or deletion result in a chromosome-wide delay in replication timing. Using a Cre/*lox*P chromosome engineering strategy, we made more than 50 deletions in chromosome 6 and discovered that a deletion as small as ~76 kb at a discrete locus could induce the DRT/DMC phenotype (Fig 2.3).

In characterizing this locus, I discovered several mono-allelically expressed transcripts. The deletions that result in DRT/DMC disrupt a large, intergenic non-coding RNA, *ASAR6* (asynchronous replication and autosomal RNA on 6), that is mono-allelically expressed (Fig 2.4). The transcript is only expressed from the chromosome that contains our *lox*P-3'RT integration cassette. I also determined that this mono-allelic expression was due to random mono-allelic inactivation and not to imprinting (Fig 2.5G). A nearby protein-coding gene, *FUT9*, is also mono-allelically expressed; but, the expressed allele is restricted to the opposite chromosome 6 (Fig 2.5). Another transcript is mono-allelically expressed from this opposite chromosome 6, in the anti-sense orientation from the protein-coding gene *FHL5*, and has thus been named *FHL5* opposite strand transcript, or *FHL5OST* (Fig A1.26). In addition to inducing DRT/DMC, disruptions of *ASAR6*

result in the re-activation of the previously silent alleles of *FUT9* and *FHL5OST*. Taken together, these data indicate that *ASAR6* controls chromosome-wide replication timing and mono-allelic expression of nearby genes, in *cis* (Fig. 2.7).

Mono-allelically expressed genes, such as imprinted genes, allelically excluded genes, and genes on the female X chromosome, replicate asynchronously [84]. We tested *ASAR6* and the other mono-allelically expressed genes at this locus on chromosome 6 for asynchronous replication and found that they do replicate asynchronously (Table 2.1).

It has been documented that the asynchronous replication of random mono-allelically expressed genes is coordinated in *cis* along the length of autosomes [99-100]. This means that all of the early replicating alleles are on the same chromosome and the late replicating alleles are on the opposite chromosome. We thus looked for coordination of our asynchronously replicating genes. All of the probes used from the ~1 Mb region around the *lox*P-3'RT integration site showed coordination of asynchronous replication timing in *cis* (Fig 2.6). However, this region was coordinated in *trans* with other asynchronously replicating loci located at distances proximal or distal to this region (Fig 2.6). The early replicating allele of *ASAR6* was on the same chromosome as the late replicating alleles of these other genes. This is a novel finding for coordination of asynchronous replication timing. It is unclear at this time whether this unique *trans* coordination is linked with the delayed replication timing phenotype or whether it may serve as a predictor of replication timing elements on other autosomes. The Thayer laboratory has, however, identified a locus on

chromosome 15 where disruptions induce the DRT/DMC phenotype, and it too displays asynchronous replication timing (N.D., L.S., and M.J.T, unpublished data). It is currently not known whether the asynchronous replication timing shows *trans* coordination with other asynchronous regions, but perhaps the asynchronous replication timing itself will prove to be a common characteristic of loci involved in chromosome-wide replication timing.

In Chapter 3 of this dissertation, I showed that it is feasible to target our 5'AP-*lox*-Blas construct into the genome using rAAV; and to subsequently use that construct in generating Cre-mediated deletions (Fig. 3.1). The data indicated that an ~46 kb deletion anchored at the *lox*P-3'RT site on chromosome 6 was not sufficient to induce the DRT/DMC phenotype (Fig 3.2). However, as a proof-of-principle experiment, it may prove very valuable. Using this technology we will be able to deliver *lox*P sites to specific regions of the genome in order to make deletions or rearrangements and test for DRT/DMC. For example, if we define characteristics that seem predictive of a locus that controls the replication timing of a chromosome, we can test such a locus by using rAAV-targeting to make rearrangements at that locus. We will also be able to recapitulate the rearrangements that show DRT/DMC in our HTD114 derivative cells in other cell types.

## Future Studies

### Genetic analysis

There are several experiments currently being done by the Thayer laboratory to continue the genetic analysis of the 6q16.1 locus. One experiment is a rescue experiment designed to replace the ~76 kb fragment that is lost in our smallest deletion with the same sequence from a BAC (see Fig. 4.1). This is a technically challenging experiment, but if it is successful, will tell us whether this ~76 kb fragment is sufficient to prevent DRT/DMC, and thus whether loss of the ~76 kb fragment is responsible for DRT/DMC.

The second experiment is to determine whether the disruptions of the *ASAR6* locus causes DRT/DMC due to loss of a genetic element, or whether it is due to placing an 'incompatible domain' out of its normal context and thereby inducing DRT/DMC. From our alternate partner analysis, we observed that DRT/DMC always segregated with the telomeric derivative of chromosome 6. This data could result from the telomeric portion being separated from the *ASAR6* locus (loss of the genetic element), or due to the some DNA element at the telomeric locus that is now placed into a different context.

To address that question, the Thayer laboratory is engineering constructs that will allow us to make the original 6;10 translocation, but in the context of the chromosome 6 distal deletion clones (see Fig. 4.2). These clones do not show DRT/DMC (no loss of the *ASAR6* locus); if they show DRT/DMC following the

translocation, the telomeric derivative would again be separated from the *ASAR6* locus, implicating the loss of that genetic element in the phenotype. However, if making that translocation does not induce DRT/DMC, it might indicate that deleting something from the distal 3'RT region on chromosome 6 takes out an incompatible DNA element that would otherwise result in DRT/DMC when removed from its normal context on chromosome 6. Alternatively, this same question is being addressed in a slightly modified way by making proximal deletions on chromosome 6; but again in the distal deletion clones (see Fig. 4.3).

We have also been performing the same types of analyses that we have done on chromosome 6 on other chromosomes that displayed DRT/DMC. Deletions on chromosome 15 in the P268 cell line show DRT/DMC (N.D., L.S., and M.J.T, unpublished data). Further deletions are being generated and analyzed to determine the minimal deleted region necessary to induce the phenotype. This locus on chromosome 15 shares similarities with the locus on chromosome 6. First, although most expression in this region appears to be bi-allelic, this locus does display asynchronous replication timing, indicating that some of the transcripts at this locus may be mono-allelically expressed in other cell types. Additionally, this region, like the locus on chromosome 6, has a high LINE-1 density, which is a predictor of mono-allelic expression [87], and has also been implicated in the process of X-chromosome inactivation [78, 77, 80, 76]. We hope that further characterization of the locus on chromosome 15 and on other DRT/DMC chromosomes will help elucidate characteristics that define these replication timing regulatory loci.

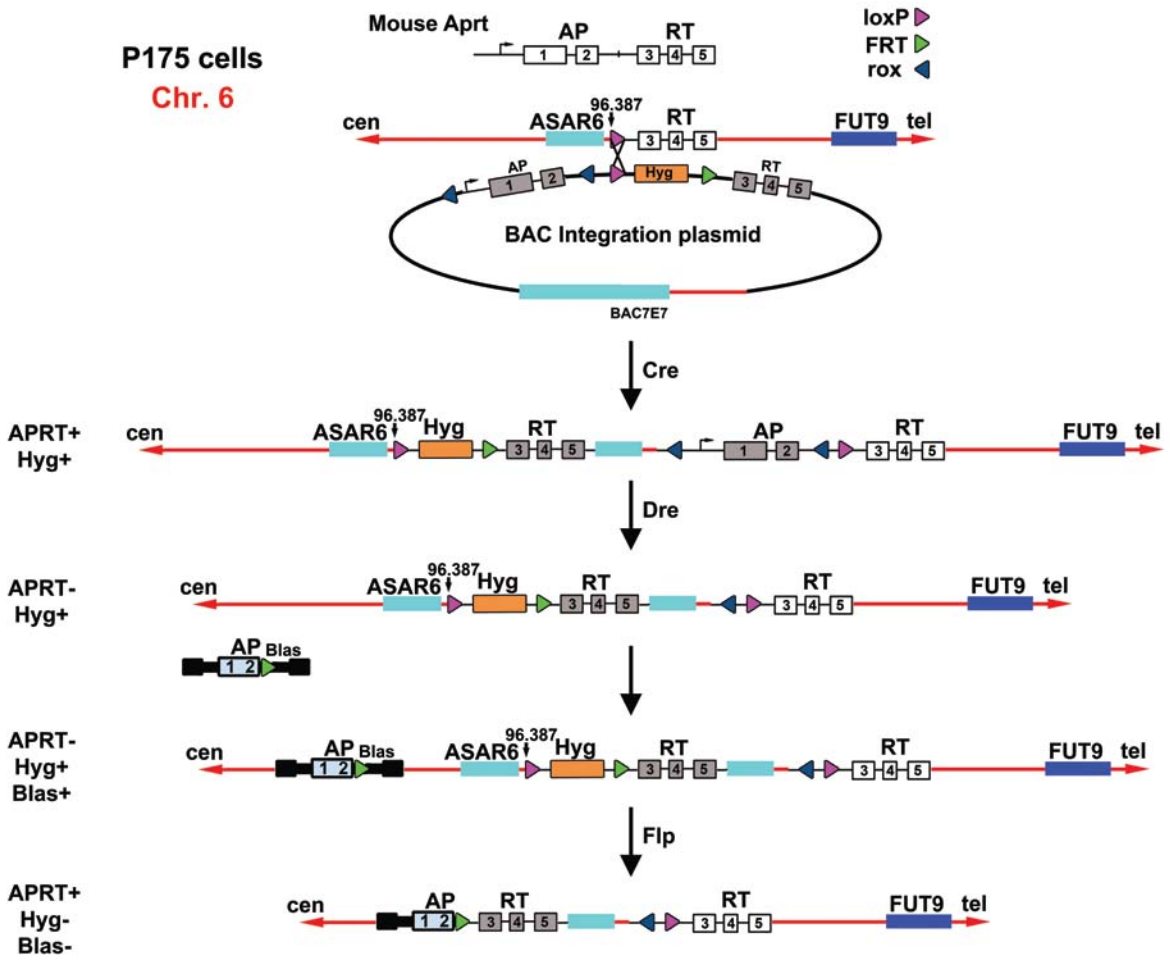



**Figure 4.1. Chromosome 6 DRT/DMC rescue.** This is a diagram of the strategy for the rescue of the DRT/DMC phenotype on chromosome 6. First, an integration plasmid containing the BAC RP11-767E7 (BAC7E7), a Hygromycin resistance gene, a 3′RT cassette, and the 5′AP cassette will be integrated at the original *lox*P-3'RT integration site on chromosome 6 in P175 cells via Cre recombinase and *lox*P sites. The resulting cells will be both APRT-positive and Hyg-positive. 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) [129-130]. Following

transfection with Dre, the 5′AP cassette gets excised, disrupting the APRT gene. APRT-negative cells are infected with a lentivirus containing the 5′AP cassette and an FRT site. Blas-positive cells are transfected with Flp recombinase. Proximal deletions on chromosome 6 result in reconstitution of the APRT selectable marker as well as loss of both Blas- and Hyg-resistance. These deletions excise various lengths of DNA on chromosome 6, proximal to the original *lox*P-3'RT integration site, and will include part 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 re-made, 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 not drawn to scale).

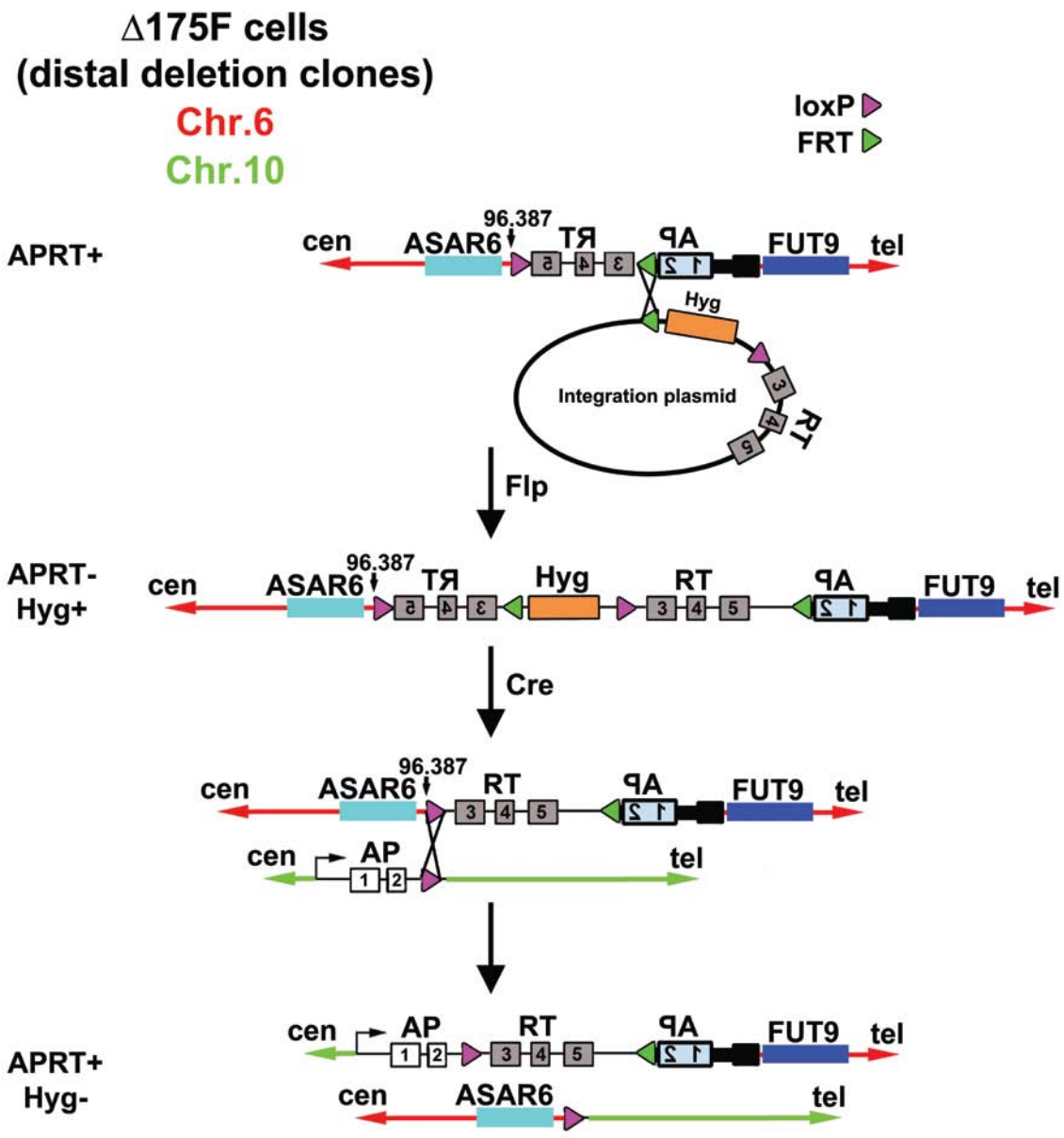



**Figure 4.2. Delta-translocation assay.** In order to test whether DRT/DMC results from loss of a genetic element at the ASAR6 locus, or due to an incompatible domain located telomeric from the *lox*P-3'RT integration cassette, the Thayer laboratory is attempting to re-make the 6;10 translocation (which displays DRT/DMC) in the context of chromosome 6 distal deletions (which do not show DRT/DMC). If the phenotype is due to an incompatible domain, that

domain will have been removed by the distal deletion prior to re-making the 6;10 translocation, thus preventing DRT/DMC from occurring. If DRT/DMC is caused by loss of the ASAR6 locus, these 6;10 translocations will still have DRT/DMC after the removal of the distal portion of chromosome 6. To accomplish this, the APRT gene is disrupted by an integration plasmid via Flp recombinase and FRT sites. The integration plasmid contains a Hygromycin resistance gene, and a *lox*P-3'RT cassette in the same orientation as the original integration cassette. Clones that properly integrate this plasmid become APRT-negative and Hyg-positive. Following transfection with Cre, these clones will flox out the Hyg gene, and clones that make the 6;10 translocation will reconstitute the APRT gene. (Figure not drawn to scale).

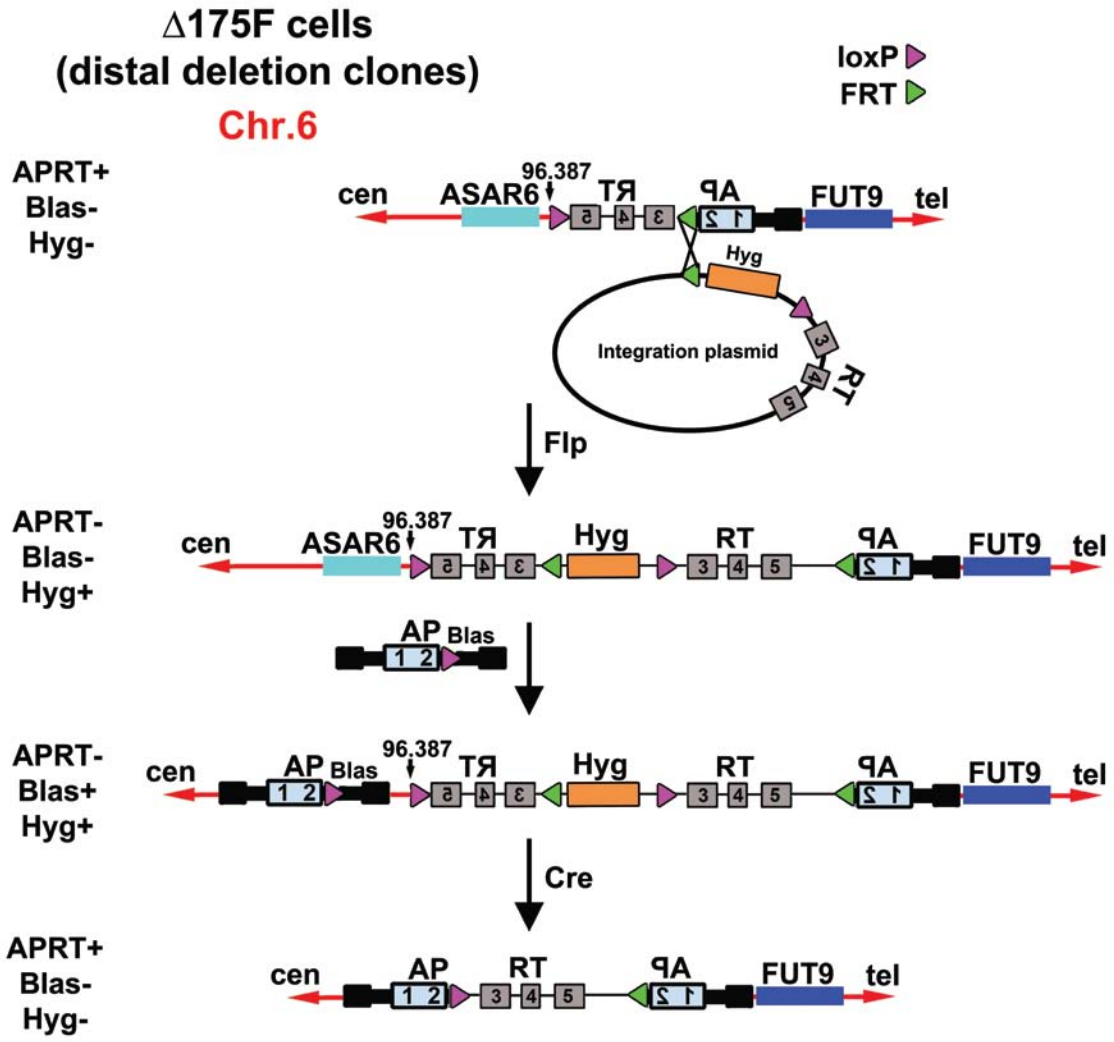



**Figure 4.3. Delta-delta assay.** In order to test whether DRT/DMC results from loss of a genetic element at the ASAR6 locus, or due to an incompatible domain located telomeric from the *lox*P-3'RT integration cassette, the Thayer laboratory is attempting to make proximal deletions on chromosome 6 (which display DRT/DMC) in the context of chromosome 6 distal deletions (which do not show DRT/DMC). If the phenotype is due to an incompatible domain, that domain will have been removed by the distal deletion prior to making the proximal deletions

on chromosome 6, thus preventing DRT/DMC from occurring. If DRT/DMC is caused by loss of the ASAR6 locus, these proximal deletions will still have DRT/DMC after the removal of the distal portion of chromosome 6. To accomplish this, the APRT gene is disrupted by an integration plasmid via Flp recombinase and FRT sites. The integration plasmid contains a Hygromycin resistance gene, and a *lox*P-3'RT cassette in the same orientation as the original integration cassette. Clones that properly integrate this plasmid become APRT-negative and Hyg-positive. Following infection with the 5′AP lentivirus, clones with viral integrations will also become Blas-positive. Following transfection with Cre, proximal deletions on chromosome 6 will reconstitute the APRT gene while losing Blas and Hyg resistance. (Figure not drawn to scale).

## *ASAR6* characterization

I have been working on characterizing the *ASAR6* RNA. *ASAR6* is transcribed by RNA polymerase II and is not-polyadenylated (data not shown). I also have preliminary data from RNA immunoprecipitation experiments (RIP) that *ASAR6* may interact with EZH2, a subunit of the Polycomb group protein, PRC2. We are looking into whether *ASAR6* may recruit PRC2 to the *FUT9* and *FHL5OST* loci, in *cis*, where PRC2 could then place repressive chromatin marks, such as H3K27me3. We are analyzing the chromatin at the *lox*P-3'RT locus by chromatin immunoprecipitation (ChIP) to determine whether there are allele-specific chromatin differences, and whether those differences are lost upon deletion of the *ASAR6* locus and re-activation of the preciously silent alleles of *FUT9* and *FHL5OST*. We also plan to look at whether the *ASAR6* RNA is necessary for the DRT/DMC phenotype or the silencing of *FUT9* and *FHL5OST* by knocking down the RNA and assaying for a delay in replication timing, the re-activation of *FUT9* and *FHL5OST*, or both.

## Mechanistic models

With some of the future experiments mentioned above, we hope to gain insight into the mechanism responsible for the DRT/DMC phenotype following disruptions of the *ASAR6* locus. Although there are no other examples of autosomal loci that control chromosome-wide replication timing, we do have models to use when thinking about possible mechanisms. First, there is the X-inactivation center (Xic), which is involved in the mono-allelic expression and

replication timing of the female X chromosomes [82-83, 62]. The Xic encodes many non-coding RNAs (ncRNAs), which play various roles, in *cis*, to silence genes on the Xi [67]. One of the functions of the ncRNA *Xist* is to bind to Polycomb group proteins and recruit them to the Xic; and, following the spread of *Xist*, to spread the PcG proteins and their repressive chromatin marks along the Xi [67].

On autosomes, there are examples of long ncRNAs having a functional role in the silencing of protein-coding genes [92-93, 67, 91, 131]. The *Kcnq1ot1* ncRNA has been shown to bind to the G9a methyltransferase and to PRC2 to mediate chromatin methylation and transcriptional silencing [92-93]. A recent report showed that a long ncRNA, antisense non-coding RNA in the INK4 locus, *ANRIL*, binds to a component of PRC2 and recruits it for the chromatin-mediated repression of the $p15^{INK4B}$ locus [131]. These studies provide clues about the potential role of *ASAR6* in silencing transcription of *FUT9* and *FHL5OST* in *cis*. It is unclear whether this mechanism, if true for *ASAR6*, will also help explain its role in controlling chromosome-wide replication timing.

Another line of investigation into the mechanistic role of *ASAR6* is to look at nuclear positioning and *trans* communication involving that locus. Pairing between the X chromosomes has been shown to be important for X-inactivation [70-72], and deletions of *Xist* have been shown to have *trans* effects on the replication timing of the X chromosomes [83]. Pairing of the X-chromosomes occurs prior to targeting of the Xi to the perinucleolar compartment during S-phase [74], and the Kcnq1ot1 RNA has also been implicated in targeting its

imprinted cluster to the perinucleolar region [92]. Taken together, these findings raise questions about whether the *ASAR6* locus has a role in either pairing of the chromosome 6 homologs or in the nuclear positioning of the chromosome that is silent for *FUT9* and *FHL5OST*, such as targeting those silent regions to the repressive perinucleolar space. We do have preliminary data showing that after making the ~230 kb deletion on chromosome 6 in the Δ175-1i clone, the distance between the two chromosome 6s during interphase is reduced (L.S. and M.J.T, unpublished data), suggesting that there are *trans* effects resulting from disrupting this locus. This deletion could alter the timing of homologous pairing or of the nuclear positioning of the deleted chromosome 6. We would like to more thoroughly investigate the location of the chromosome 6s before and after disruptions with respect to each other and to other nuclear landmarks such as the perinucleolar compartment.

We would also like to explore the mechanism by which DRT/DMC chromosomes undergo frequent secondary rearrangements. A recent report by Stephens et al. posits that certain tumors that show a vast number of rearrangements of a single chromosome do so via a mechanism they term chromothripsis [19]. The term chromothripsis comes from the Greek words *chromos*, for chromosome, and *thripsis*, meaning shattering into pieces. The model is that the vast numbers of chromosomal rearrangements occur following a single catastrophic event that shatters the chromosome, the pieces of which are then stitched back together in a patchwork manner, generating many intra-chromosomal rearrangements and a few inter-chromosomal rearrangements

[19]. Data from our lab, however, indicates that a high number of chromosomal rearrangements can result from a simpler origin. Rather than a catastrophic shattering of the chromosome, a single genetic event, such as a translocation or deletion of a chromosome, which results in DRT/DMC, can initiate genomic instability and the accumulation of a large number of chromosome rearrangements in a small number of generations. We hypothesize that the chromosome-wide delay in replication timing, followed by the delay in mitotic chromosome condensation, results in numerous fragile sites along the chromosome that are susceptible to DNA damage (see Fig 4.4). These sites of DNA damage along the chromosome results in multiple rearrangements of the chromosome, many of which would be intra-chromosomal, and some of which would be inter-chromosomal.

As part of the chromosome engineering project, cells containing a chromosome with DRT/DMC were followed for 20 cell cycles [10]. After 20 cell cycles, these chromosomes show a 30-80-fold increase in the number of new gross chromosomal rearrangements [10]. Although we know these chromosomes have a high number of new gross chromosomal rearrangements, we would like to use next-generation sequencing to map the breakpoints at a high resolution for all of the chromosome rearrangements in these clones in order to detect the smaller intra-chromosomal rearrangements that may have been missed by the G-banding and FISH assays used in the original analysis. We hope to determine whether this sequence data aligns with what was seen by Stephens et al. that led to their model for chromothripsis [19], and thus establish DRT/DMC as the

possible driver of these massively rearranged karyotypes. Because chromosomes with DRT/DMC acquire a susceptibility to damage and rearrangements due to their delay in replication timing and subsequent delay in mitotic chromosome condensation [8], DRT/DMC is an intriguing model to explain the findings seen in that study.

## Conclusions

In this dissertation, I have shown that disrupting a locus on human chromosome 6 resulted in a chromosome-wide delay in replication timing and the bi-allelic expression of previously mono-allelically expressed genes. Chromosomal rearrangements that result in this delayed replication timing phenotype are unstable [8-10], indicating that this phenotype may play a role in the genomic instability commonly associated with cancer cells and in cells exposed to ionizing radiation.

Our data also show that deletions at this locus on chromosome 6 disrupt a mono-allelically expressed, intergenic, long non-coding RNA, *ASAR6*. Other mono-allelically expressed transcripts at this locus, *FUT9* and *FHL5OST*, become bi-allelically expressed following disruptions of the *ASAR6* locus. We have shown that the *ASAR6* locus replicates asynchronously and that the asynchronous replication is coordinated in *cis* within this ~1 Mb region. However, this asynchronous replication is coordinated in *trans* with other asynchronously replicating regions on chromosome 6 – a result that has not previously been seen for the coordination of asynchronously replicating loci.

I hypothesize that all human autosomes will have one or more loci that are capable of affecting the replication timing of the entire chromosome. It has previously been shown that the X chromosome has such a locus (the XIC) [82-83], the data in this dissertation describes a similar locus on chromosome 6, and we also have data that a locus with these properties exists on chromosome 15 as well (N.D., L.S., and M.J.T, unpublished data). The loci on chromosomes 6 and 15 were identified due to their involvement in the DRT/DMC phenotype; and, since several other chromosomes display DRT/DMC following translocations, it is likely that they too will contain loci that are important for chromosome-wide replication timing. Additionally, *cis*-regulatory sequences have previously been shown to control replication timing of local origins (reviewed in [28, 132]), indicating that, while novel, our findings are not without parallel. In fact, Pope et al. stated that it "is reasonable to expect that cis-acting regulatory elements for replication timing exist in mammals" [132].

It remains to be seen whether the loci involved in the DRT/DMC phenotype actively control the replication timing program of an entire chromosome or whether disrupting these loci interferes with a separate process (such as asynchronous replication timing), which has a downstream consequence of delaying the replication timing of the chromosome. These *cis*-acting loci, at the very least, are likely dependent on their genomic context. For example, the locus control region (LCR) of the human beta-globin gene has been implicated in the chromatin structure and replication timing of this locus [133]. However, while certain sequences from this region are capable of regulating replication timing in

a mouse transgene context, these sequences are not needed for replication timing regulation in the native genomic context [134-135].

In the case of DRT/DMC loci, I think at least one component of the regulatory role these loci play will be due to their genomic context; for example, being located within asynchronously replicating regions. These asynchronously replicating regions may contain genetic elements (such as ncRNAs) or specific epigenetic modifications that help establish the asynchronous replication pattern, and loss of these elements or modifications of the elements result in misregulation of replication timing for the entire chromosome. Alternatively, these regions of asynchronous replication timing may contain barriers between asynchronously and synchronously replicating regions that if disrupted, cause defects in the replication timing signals or the replication machinery.

Identification of the loci involved in DRT/DMC is important for several reasons. First, chromosomes with DRT/DMC provide a model by which cancer cells and cells exposed to IR can acquire genomic instability. Mapping the precise locations of these loci will help elucidate one mechanism by which genomic instability occurs. Our model for the process by which DRT/DMC leads to instability is that the delay in replication timing leads to a subsequent delay in mitotic condensation, and this failure to properly condense prior to mitosis predisposes the chromosomes to DNA damage followed by chromosomal rearrangements (see Fig. 4.4). Although we think this is a likely picture of what happens as a result of DRT/DMC, we still want to understand the genetic and

molecular origins of the phenotype; identification of the loci involved in DRT/DMC is an important first step in that understanding.

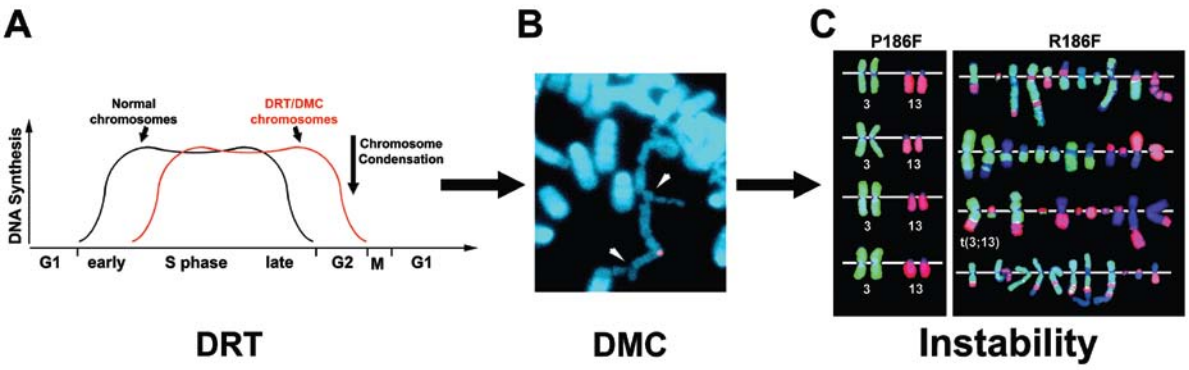



**Figure 4.4. Delayed replication timing, delay in mitotic condensation and genomic instability.** Chromosomes with DRT/DMC replicate 2-3 h later than normal chromosomes (**A**). DRT results in a subsequent delay in mitotic chromosome condensation (DMC) (**B**). DMC chromosomes are susceptible to damage [(**B**), white arrows]. DRT/DMC chromosomes participate in frequent rearrangements (**C**). Panel C shows the non-rearranged chromosomes 3 and 13 in four clones of the P186 cell line. Following transfection with Cre, a t(3;13) translocation is generated, with both derivatives showing DRT/DMC. After 20 cell cycles, the R186 clones show numerous rearrangements involving both chromosomes 3 and 13.

In addition to its potential role in the progression of cancer, identifying the loci involved in DRT/DMC may provide insight into other important biological questions. For instance, since these loci can affect chromosome-wide replication timing, identifying them may help us to understand the establishment and

maintenance of chromosome replication timing programs. DRT/DMC loci may also play a role in the asynchronous replication timing associated with random mono-allelically expressed genes, or in the regulation of the mono-allelic expression itself. Therefore, identifying these loci may allow us to learn more about the process by which asynchronous replication timing and random mono-allelic expression are established and regulated, and the link between the two processes.

Therefore, we hope to genetically and molecularly define what the regions involved in DRT/DMC will look like on every chromosome and thus have the ability to predict and test for such loci. The data presented here is the first attempt to genetically and molecularly characterize one of these regions on an autosome. We have shown that an autosomal locus on chromosome 6 is capable of controlling the replication timing of an entire chromosome in *cis*, and that disruption of such a locus can result in genomic instability, a common phenotype associated with cancer and cells exposed to ionizing radiation.

# APPENDIX I

## Supplemental Data: An Autosomal Locus that Controls Chromosome-wide Replication Timing and Mono-Allelic Expression

Eric P. Stoffregen, Nathan Donley, Daniel Stauffer, Leslie Smith, and Mathew J. Thayer*

Department of Biochemistry and Molecular Biology

Oregon Health & Science University

3181 S. W. Sam Jackson Park Road

Portland, Oregon 97239, USA

***Corresponding author.**

TEL: (503) 494-2447

FAX: (503) 494-7368

Email:thayerm@ohsu.edu

**Human Molecular Genetics** (2011) **20** (12), 2366-2378.

## Introduction

The following material was provided along with the material in Chapter 2 in our publication in Human Molecular Genetics [136]. The text and figures provide more detail about the chromosome engineering strategy used to make the original translocations that display DRT/DMC [9-10], the lentiviral strategy to make alternative partner translocations, and the strategies to make both the proximal and distal deletions on chromosome 6. It also includes more details on the characterization and analysis of the alternative partner translocations and the chromosome 6 deletion clones (including screening for, and confirmation of, deletions, as well as LOH analysis of two of the deletions). The details about the replication timing assay, data showing the replication timing analysis of several of the normal chromosomes in P175, as well as the replication timing analysis of chromosome 6 in some of the deletion clones is included. This appendix includes more information on both the assay for asynchronous replication timing and for the coordination of asynchronous replication timing. This material will also go more in-depth into the mono-allelic expression analysis at the *lox*P-integration locus on chromosome 6 and the loss of mono-allelic expression following deletions at the locus. The List of Figures and Tables at the front of this dissertation provides more information about what is included in this appendix and will help navigate the material in this section.

## Chromosome engineering

Our chromosome engineering strategy is based on the Cre/*lox*P site-specific recombinase system (reviewed in [56]) to generate chromosomal rearrangements [9-10]. Because the Cre/*lox*P system is relatively inefficient at mediating inter-chromosomal events ($<1 \times 10^{-3}$), we are using reconstitution of a selectable marker to isolate the cells that undergo Cre-mediated recombination. A schematic representation of this strategy is shown in Figure A1.1. This strategy involves reconstitution of a cloned murine *Aprt* gene in a human cell line, HTD114, which is APRT-deficient [119]. The strategy involves the generation of parental cell clones (P-clones); each containing two independently inserted *lox*P cassettes. One cassette contains the 5' portion of the *Aprt* gene linked to a G418 resistance gene (5'AP-*lox*P), and the other cassette contains the 3' portion of the *Aprt* gene linked to a Hygromycin B resistance gene (*lox*P-3'RT). Each cassette contains *lox*P sites in the second intron of the *Aprt* gene. Thus, following Cre-mediated recombination, the *Aprt* gene is reconstituted at the *lox*P sites within the second intron, and reciprocal exchanges are generated as recombinant clones or pools (R-clones/pools). In addition, work from Dr. M. Jasin's laboratory indicates that chromosome translocations can be generated from DNA double strand breaks induced by the rare cutting restriction enzyme I-Sce1 [137]. Therefore, as an alternative approach to generate translocations using these same cassettes, we introduced I-Sce1 recognition sequences into the second intron of the *Aprt* gene in the 5' and 3' cassettes (Fig. A1.1).

This system allows for the generation of chromosome rearrangements at similar breakpoints using two distinct mechanisms: 1) site-specific recombination mediated by Cre, and 2) non-homologous end joining of DNA double strand breaks generated by I-Sce1. One of the advantages of this system is that the same rearrangements can be generated in multiple, independent events. In addition, because the 5' and 3' *Aprt* cassettes contain *Neo* and *Hyg* genes flanked by *loxP* sites (i.e. floxed) (Fig. A1.1), Cre-mediated intra-chromosomal deletions of the *Neo* and *Hyg* genes can be accomplished without the generation of translocations – owing to the fact that Cre is more efficient when *lox*P sites are in close proximity [106]. These small internal deletions serve as controls for Cre-mediated events at specific chromosomal locations, and to assay any non-specific affects of Cre forced expression [10]. Furthermore, deletion of the *Neo* and *Hyg* genes allows us to reuse our original 5'AP-*lox*P and *lox*P-3'RT cassettes, as the resulting cells are sensitive to G418 or Hygromycin B, respectively.

## Balanced Translocations display DRT/DMC

To identify chromosome translocations that display the DRT/DMC phenotype, we previously isolated 97 P-clones that contained independent, random integrations of the two *Aprt* cassettes. Subsequently, each P-clone was transiently transfected with a Cre expression vector and subsequently cultured in media that selects for Aprt+ cells. Pools of Aprt+ colonies were expanded and harvested for karyotypic analysis. Mitotic spreads from each R-pool, were

analyzed for the presence of chromosomes with DMC. One important consideration for detection of the DMC phenotype is to omit a colcemid pretreatment step prior to the mitotic harvest [8]. Therefore, all karyotypic assays for DMC were carried out in the absence of colcemid. We found that the majority of R-pools displayed a relatively low frequency of DMC, ranging from 0-4%. In contrast, 5 R-pools contained chromosomes with DMC in a greater fraction of mitotic spreads, ranging from 6% to 50%.

This analysis led to the identification of five different balanced translocations that display DRT/DMC ([10]; also see Fig. A1.2). In addition, detailed karyotypic analysis of the resulting balanced translocations indicated that, on 4 of 5 translocations, only one derivative chromosome displayed the DMC phenotype [10]. For example, Figure A1.3 shows a schematic diagram (panel A), partial G-banded karyotype (panel B), and the DMC phenotype (panel C) on the chromosome 10 derivative of a balanced translocation involving chromosomes 6 and 10. This t(6;10) can be generated by treating the P175 cell line with either Cre or I-Sce1 (Fig. A1.1), selecting for Aprt+ cells, and isolating pools or clones of R175 cells.

## DRT/DMC and Genomic Instability

Using this chromosome engineering system, we utilized a fluctuation analysis to show that cells containing chromosomes with DRT/DMC have a 30-80-fold increase in the rate at which new gross chromosomal rearrangements occur,

indicating that cells with DRT/DMC display translocation instability [10]. In addition, we have found that DRT/DMC chromosomes activate the DNA replication and spindle assembly checkpoints [54]. Furthermore, cells with DRT/DMC chromosomes have centrosome amplification, abnormal spindle assembly, experience non-disjunction events at an increased frequency, experience endoreduplication, and display aneuploid karyotypes, indicating that chromosomes with DRT/DMC are associated with chromosome instability. For these reasons, all of the assays carried out in the present study were performed on low passage (<5) cells to ensure that secondary rearrangements and chromosome losses and gains were at a minimum. In addition, because our chromosome engineering system allows for the generation of the same genetic alteration in multiple independent events, all of the replication timing assays for the translocations and deletions described here were carried out on multiple independent clones for each genetic alteration.

## ‘Alternative Partner’ analysis

We previously identified a Cre/*lox*P-dependent balanced translocation, t(6;10)(q15;q11.2) induced by expressing Cre in the P175 cell line, that displays DRT/DMC [10] (see Figs A1.2 and A1.3). To generate alternative partner translocations with both chromosomes 6 and 10, we took advantage of the fact that the *lox*P cassettes inserted in P175 contain ‘floxed’ *Neo* and *Hyg* genes (Fig. A1.1), and the fact that we had previously generated P175 cells that had deleted both selectable markers via Cre transient expression, but had not generated the

t(6;10) (P175ΔNH) [10]. Deletion of the *Neo* and *Hyg* genes allowed us to reuse our original 5'AP-*lox*P and *lox*P-3'RT cassettes, as P175ΔNH cells are sensitive to either G418 or Hygromycin B. All of the experiments in this manuscript were carried out with P175ΔNH, and are referred to as simply P175 throughout this report. We isolated 20 clones containing new integrations of the 5'AP-*lox*P cassette [named P175(5')-1…20] and 20 clones containing the *lox*P-3'RT cassette [named P175(3')-1…20]. In addition, we have also used lentiviral vectors, containing new 5'AP and 3'RT *lox*P cassettes, to integrate new *lox*P sites into the chromosomes of P175 to generate additional P175(5') and P175(3') clones (see deletion analysis below). We next transiently expressed Cre in each new P175(5') and P175(3') clone and generated Aprt+ clones from each. Note that each new P175(5') and P175(3') clone can potentially generate two different translocations, the original t(6;10) or a new translocation involving either chromosome 6 or 10 and a new chromosome, depending on the cassette and 'new' integration site (Fig. 2.1A).

## Chromosome 6 Alternative Partners

We next screened each new Aprt+ clone using Southern blot hybridizations that allowed us to distinguish between the original t(6;10) and any new translocation. Two examples are shown in Figure A1.4. For this analysis, we digested genomic DNA, isolated from P175, R175, P175(5')-5, P175(5')-6, and clones R175(5')-5-1…9 and R175(5')-6-1…4, with the Bcl-1 restriction enzyme. Note that Bcl-1 restriction sites are not present in either of the original *lox*P

cassettes, and therefore the size of the restriction fragment generated by Bcl-1 digestion is dependent upon the genomic site of integration. In addition, upon Cre treatment, the size of the Bcl-1 restriction fragments are expected to change, with the new size of the fragments dependent on the location of the Bcl-1 restriction site present at the genomic integration site of the other partner cassette. For example, Bcl-1 digestion of DNA from P175 results in a single relatively large (~20 kb) restriction fragment when probed with the AP half of the *Aprt* gene (Fig. A1.4B and C). This ~20 kb band represents the original chromosome 10q11.2 integration site of the 5'AP-*lox*P cassette. Furthermore, two new restriction fragments are generated in R175, ~19 kb and ~14 kb, suggesting that P175 has two AP cassettes integrated at the same location in chromosome 10. Additional Southern blot analysis of P175 using other restriction enzymes and probes indicated that P175 does indeed contain two 5'AP-*lox*P cassettes integrated at the same location (not shown). Regardless, the ~19 kb and ~14 kb restriction fragments are diagnostic for the t(6;10) present in the original R175 genomic DNA. Figure A1.4 shows the Southern blot hybridizations for two new P175(5') clones, P175(5')-5, and P175(5')-6 and their Aprt-positive R175(5')-5 or -6 clones. P175(5')-5 DNA contains a new Bcl-1 restriction fragment (~4.5 kb) that hybridizes to the AP probe (Fig. A1.4B).

This new AP restriction fragment represents the new 5'AP-*lox*P cassette integrated into another site in the genome. In addition, all of the R175(5')-5 clones have experienced a Cre-mediated event that is distinct from the original R175 rearrangement based on the following criteria: 1) the new AP restriction

fragment present in P175(5')-5 is no longer detectable in the R175(5')-5 clones, indicating that the new 5'AP-*lox*P cassette participated in a Cre-dependent rearrangement; 2) the original ~20 kb AP restriction fragment present in P175 is still present in the R175(5')-5 clones, indicating that the original 5'AP-*lox*P cassette did not experience a Cre-mediated event in these clones; 3) the ~19 kb and ~14 kb Bcl-1 fragments are not present in any of the R175(5')-5 clones, indicating that the t(6;10) was not generated in these clones; and 4) a new Bcl-1 restriction fragment (~9 kb) is present in the R175(5')-5 clones, indicating that a new Cre-dependent AP fragment was created. All of these changes in restriction fragment size are consistent with a new rearrangement involving chromosome 6 in all of the R175(5')-5 clones. Karyotypic analysis of the R175(5')-5 clones indicated the presence of a new balanced translocation involving chromosomes 6 and 17, t(6;17)(q15;q25) (Fig. A1.5).

In contrast, a similar analysis of the Aprt+ R-clones isolated from a different P175(5') clone, P175(5')-6, identified R-clones with the original t(6;10) restriction pattern only (Fig. A1.4C). Karyotypic analysis of one of these R175(5')-6 clones indicated the presence of the original t(6;10) (not shown), and were not analyzed further.

Screening additional P175(5') clones, and their respective Aprt+ R175(5') clones using this Southern hybridization scheme, identified three additional alternative partner balanced translocations with chromosome 6 (data not shown): t(6;7)(q15;q36) (Fig. A1.6), t(6;9)(q15;p21) (Fig. 2.1; Figs A1.7 and A1.8), and t(6;8)(q15;q24.1) (Fig. A1.9).

One of the advantages of our system is that the same balanced translocations can be generated in multiple independent events, which allows for the analysis of the same translocation in independent clones. For example, we isolated 9 independent clones, R175(5')-5-1…9, containing the new t(6;17) translocation (Fig. A1.4B). Therefore, to rule out any clonal variability in replication timing we characterized a minimum of three independent clones for each of the new alternative partner translocations both for the presence of the alternative partner translocation and for replication timing changes (see below).

## Chromosome 10 Alternative Partners

Using a similar Southern hybridization screening strategy, as described above for the chromosome 6 alternative partners, we analyzed new P175(3') and Aprt+ R175(3') clones for new rearrangements involving chromosome 10. We processed the Bcl-1 digested genomic DNA by Southern blot hybridizations using the 3'RT half of the *Aprt* gene as probe (data not sown). This allowed us to distinguish between the original t(6;10) and new translocations involving chromosome 10. From this analysis we identified three new balanced translocations involving chromosome 10, a t(1;10)(p23;q11.2), a t(4;10)(q25;q11.2), and a t(5;10)(q35.1;q11.2) (Figs A1.10-12, respectively).

## Replication Timing Assay

The DRT phenotype is characterized by a >2 h delay in the initiation as well as the completion of DNA synthesis along the entire length of the chromosome, while the other chromosomes within the same cell show normal patterns of DNA synthesis. To characterize the replication timing of the chromosomes involved in this study prior to any Cre-mediated rearrangements, we carried out an extensive analysis of chromosome replication timing in parental P175 cells. Our most sensitive assay for detection of DRT was to use pulse-labeling with BrdU followed by detection of the incorporated BrdU by fluorescent immunostaining combined with FISH using chromosome-specific probes on metaphase spreads. For this analysis, we cultured P175 cells with BrdU for 4.5, 5, 6, 7, 8 and 9 h and harvested mitotic cells in the absence of a colcemid pretreatment step. This 'terminal labeling' protocol detects the latest replicating regions of chromosomes at the shorter time periods (4.5-6 h) of BrdU incorporation and middle S-phase plus the late replicating regions during the longer BrdU pulses (7-9 h; see Chapter 2, Fig. 2.1B). Figure A1.13 shows examples of this replication timing assay for chromosome 6 in P175 cells. The earliest detectable BrdU incorporation, representing the last DNA to replicate, occurred during the 6-h time point, indicating that G2 in P175 cells was at least 5.5 h. Note that the two chromosome 6s present in each P175 cell show very similar amounts of BrdU incorporation at each time point and that the 'banded' pattern of incorporation is similar between the two chromosome 6s within any given cell. In addition, the banded pattern of BrdU incorporation at any given time point is quite similar

between cells, and this banded pattern of incorporation is consistent with previous replication timing maps for human chromosome 6 [104-105]. Furthermore, quantification of the BrdU incorporation was determined for the two chromosome 6s by measuring the total amount of fluorescent signal present. For this analysis, we determined the amount of BrdU immunostaining by measuring the pixel intensity along each chromosome, and multiplying the average pixel intensity by the area occupied by those pixels (Fig. A1.14A-D). For comparison, we also quantified the total amount of DAPI staining on each chromosome pair using the same metrics (Fig. A1.14D). Applying this analysis to 7 different mitotic spreads indicated that the two chromosome 6s within individual P175 cells incorporated very similar amounts BrdU, indicating that the chromosome 6s in P175 replicate synchronously (Fig. A1.14E).

We also carried out a similar replication timing assay for chromosomes 1, 4, 5, 7, 8, 9, 10, and 17. Figure A1.15 shows this replication timing assay for chromosome 10 in P175 cells. Note that the two chromosome 10s present in P175 show very similar amounts of BrdU incorporation and that the banded pattern of incorporation is similar between the two chromosomes. The amount of the BrdU incorporation was determined by quantifying the total amount of fluorescent signal present on each chromosome (Fig. A1.15C and D).

This analysis indicated that the two chromosome 10s in individual P175 cells incorporated very similar amounts of BrdU, indicating that the chromosome 10s replicate synchronously in P175 cells (Fig. A1.15E). A similar replication timing analysis indicated that the replication timing of chromosomes 1, 4, 5, 7, 8, 9, and

17 was synchronous between homologous pairs (Fig. A1.16), and that the banded pattern of BrdU incorporation (not shown) was consistent with the known replication timing maps for each of these chromosome pairs [104-105].

We next assayed the chromosome replication timing of the chromosome 6 alternative partner translocations. Cultures were incubated with BrdU for 4.5, 5 and 6 h and mitotic cells were harvested and processed for BrdU incorporation and for FISH using a chromosome 6 paint plus centromeric probes specific to the alternative partner chromosomes, as described above. In total, DRT was detected on three of the four chromosome 6 alternative partner translocations, a t(6;17)(q15;q25) (Fig. A1.5), a t(6;7)(q15;q36) (Fig. A1.6), and a t(6;9)(q15;p21) (Fig. 2.1; Fig. A1.8). Analysis of a fourth alternative partner translocation involving chromosome 6, t(6;8)(q15;q24.1), indicated that it does not display DRT (Fig. A1.9).

We next assayed the chromosome replication timing of the chromosome 10 alternative partner translocations. Cultures were incubated with BrdU for 4.5, 5, and 6 h and mitotic cells were harvested, and processed for BrdU incorporation and subjected to FISH using a chromosome 10 paint and chromosome-specific centromeric probes as described above. DRT was not detected on any of the three chromosome 10 alternative partner translocations (Figs A1.10-12).

## Engineering Deletions in Chromosome 6

### Chromosome 6 Proximal Deletions

Our results obtained from the alternative partner analysis suggested that the DRT phenotype segregated with the chromosome 6 *lox*P integration site and not with the chromosome 10 integration site in P175 cells. Therefore, we next generated internal deletions in chromosome 6, nested at the original *lox*P-3'RT integration site in P175 cells. First, we determine the exact location of the chromosome 6 *lox*P cassette integration site using inverse PCR [138]. For this analysis, we used Southern blot hybridizations to characterize the plasmid-genome junctions of 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-3'RT cassette (data not shown). Direct sequencing of the PCR product indicated that the *lox*P-3'RT cassette was inserted at position 96,386,321 base pairs (NCBI Build 36 version 2) of chromosome 6. This integration site was confirmed by PCR using primers designed to amplify the plasmid-genome junctions of both the 5' and 3' ends of the inserted cassette (Fig. A1.17D; Table A1.1). A schematic view of the *lox*P-3'RT integration site, with the location of the cassette-genome junction primers, is indicated (Fig. A1.17C). In addition, this analysis revealed that a 70 bp deletion had occurred in chromosome 6 upon integration of the original *lox*P-3'RT cassette (Table A1.1).

The rationale behind the methodology used to engineer chromosome 6 deletions was based on previous studies showing that intra-chromosomal Cre events are markedly more efficient than inter-chromosomal Cre events (reviewed in [106]). The design of this set of experiments was to introduce new *lox*P cassettes, 5'AP-*lox*P, into P175 cells, obtain pools of cells containing these new integration sites and subsequently transiently express Cre in each pool of clones and isolate Aprt+ colonies. Accordingly, reconstitution of the *Aprt* gene with Cre is expected to occur at a higher frequency in cells containing lentiviral integrations near the original *lox*P-3'RT cassette. For these experiments we used a lentiviral vector (pL6-AP*lox*-Blas) that contains the 5' half of the *Aprt* gene (AP), a *lox*P site, and the selectable marker for Blasticidin (Blas) resistance. A schematic diagram of this lentiviral construct is shown in Figure A1.17A (also see Fig. 2.2A). In addition, this virus was designed so that the *Blas* gene would be deleted if the virus inserted into a *lox*P-3'RT-containing chromosome and if it is oriented in a way that would generate a deletion following Cre-mediated recombination with the original *lox*P-cassette integration site. In order to generate a large number of random integrations, we infected P175 cells with this AP-*lox*P-Blas virus at a MOI of <0.1, thus ensuring single copy integrations; from this, 45 independent pools, each containing ~5000-10,000 Blasticidin-resistant colonies, were isolated. Note that with this number of pools and integration sites a new *lox*P cassette is predicted to be introduced throughout the P175 genome approximately every 20 kb. Subsequently, each pool was transiently transfected with a Cre expression vector, selected for reconstitution of *Aprt*, and independent

Aprt+ clones isolated (>1,000 total clones). Due to random integration of the Lentivirus, only 1 in 4 integrations in chromosome 6 would contain the Lentivirus in the correct position and orientation to generate deletions proximal to the original *lox*P-3'RT integration site. Therefore, each Aprt+ clone was tested for growth in media containing Blasticidin, and all Blasiticidin-sensitive clones were expanded and analyzed by Southern blot hybridization for retention of the *Blas* gene and for novel lentiviral integrations. Figure A1.17B shows an example of this analysis for clones isolated from pools #1, #2, #3, #4, #5, and #6. Note that most of the Blasticidin-sensitive clones have lost the *Blas* gene as judged by probing the blot with a *Blas* probe. Furthermore, probing a second blot with the 5' portion of the *Aprt* gene (AP probe) indicated that some of the clones from each pool retained restriction fragments of similar size, i.e. clones Δ175-1a, -1b, -1c, -1d, -1e, -1f, and -1i, suggesting that they came from the same lentiviral integration site in Pool #1, and therefore represent the same deletion. In addition, other clones from each pool retain different sized AP restriction fragments, i.e. clones Δ175-1g and -1h, indicating that these clones came from a different lentiviral integration in Pool #1 and therefore represent different deletions. To confirm that the Blas-negative clones contain deletions in chromosome 6 proximal to the original *lox*P-3'RT integration site, we carried out PCR analysis with primers directed to the cassette-genome junctions. The primers were designed to detect either proximal or distal junction fragments at the original *lox*P-3'RT integration site (Fig. A1.17C).

Figure A1.17D shows this analysis for a subset of clones from Pool #1, #2, #3, #4, and #6, and indicates that, as expected, the distal junction is present in all of these clones and that the proximal junction is not. In total, we have isolated P175 sub-clones that contain >40 distinct lentiviral integrations that have lost the *Blas* gene and this proximal cassette-genome junction.

## Chromosome 6 Distal Deletions

One caveat to the relatively straightforward generation of deletions in chromosome 6, nested at the *lox*P-3'RT integration site, is that deletions could be easily generated in only one direction, i.e. proximally, from this integration site. Therefore, to generate deletions distal to the *lox*P-3'RT integration site, we used a second site-specific recombinase system, Flp/FRT, to engineer deletions distal to the *lox*P-3'RT cassette in chromosome 6 of P175 cells. Like Cre, the Flp recombinase can be used in combination with its recognition sequence, FRT, to generate site-specific rearrangements in mammalian chromosomes [56]. For these experiments, we first integrated a plasmid containing a single *lox*P site located in the second intron of an AP-*lox*P cassette (Fig. A1.18). Thus, in a Cre-dependent manner this plasmid is directed to the original *lox*P-3'RT integration site in chromosome 6 of P175, and properly targeted integrations are recovered by selection for reconstitution of the *Aprt* gene. The integrated plasmid also contained a second 3'RT cassette (positioned in the opposite orientation from the original cassette) linked to an FRT site in its second intron. The integration

plasmid also contained a second FRT site located 5' of the AP cassette in the configuration shown in Figure A1.18.

This strategy allowed us to remove the AP portion of the integrated plasmid by transient expression of Flp, which allowed us to re-isolated Aprt-negative cells and to reuse the Aprt selectable marker. Subsequently, these Aprt-negative cells were infected with a lentivirus (L6-AP-FRT-Blas) containing an AP-FRT cassette, as shown (Fig. A1.18), and pools of Blasticidin-resistant colonies were isolated and transiently transfected with a Flp expression vector followed by the re-isolation of Aprt+ colonies.

Using the same strategy that we used to identify and characterize the proximal deletions in chromosome 6, described above (Fig. A1.17), we isolated four Flp/FRT-mediated deletions in chromosome 6 distal to the original *lox*P-3'RT integration site, ranging in size from ~26 Mb to as small as ~18 kb (Table A1.1; Fig. A1.19).

We next used LAM-PCR [107] 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). We have successfully cloned and sequenced LAM-PCR products from 21 of the >50 lentiviral integrations (Table A1.1). In addition, we have confirmed all 21 integration sites using PCR with primers (Table A1.3) directed at the lentiviral

5'LTR-genome junction sites (not shown). Furthermore, we have confirmed that these 21 different lentiviral integration sites generated deletions in chromosome 6 using LOH analysis (see below) and FISH with probes representing BAC RP11-374I15 (Fig. 2.2; Fig. A1.23), BAC CTD-2231M1 (Figs A1.19, A1.25A and B) or with Fosmid G248P86031E7 (Fig. 2.2A and 2.3F).

## Affymetrix Genotyping, Copy Number, and Loss of Heterozygosity

In order to characterize these chromosome 6 deletions in molecular detail, we have carried out a genotyping and copy number analysis of the parental cell lines HTD114 and P175 and two different clones containing proximal deletions in chromosome 6, Δ175-2d (~16.6 Mb deletion) and Δ175-1i (~231 kb deletion) using the Affymetrix 500K SNP array. The array hybridizations were subjected to a performance evaluation. The metrics that were used for this evaluation were those suggested by Affymetrix, Inc. The quality of the hybridizations were assessed using two sets of measurements: 1) the genotype call rates across the array as determined by the Affymetrix Dynamic Modeling (DM) algorithm and 2) pair-wise comparison of the genotypes from 50 SNPs that are tiled onto both the NspI and the StyI arrays. Passing call rates suggested by Affymetrix are >93%. The call rates for our initial analyses were 95.84% for the Nsp1 array and 95.33% for the Sty1 array. In addition, the OHSU Affymetrix core uses 48 identical calls out of the 50 common SNPs (96%) as a passing threshold. Our analysis yielded 50 out of 50 identical calls on all samples.

In order to characterize the extent of the deletions in chromosome 6, present in the Δ175-2d and Δ175-1i cell lines, we carried out LOH and copy number analyses, using the micro-array SNP data and Partek Genomics Suite software. Figures A1.20 and A1.21 show the results of these analyses and indicate that both cell lines contain LOH and decreased copy number at the chromosomal positions predicted by the two integrated *lox*P cassettes in each clone. For example, the original *lox*P-3'RT integration site in P175 is at position 96,386,321 and in both deletion cell lines LOH begins at the first heterozygous SNP proximal to this integration site.

In addition, the LOH continues in each cell line until the first heterozygous SNP proximal to the lentiviral integration site for each clone (Figs A1.20 and A1.21; Table A1.1). Similarly, we find decreased copy number at the same genomic positions as observed in the LOH analysis for each cell line (Figs A1.20 and A1.21). Finally, using PCR and sequencing, we have confirmed >100 of these heterozygous SNPs, located on chromosomes 6, 10, 13, and 15 (Figs 2.4B and C, 2.5B; Table A1.2; and not shown [for chromosomes 13 and 15]). In total, we have identified >119,000 heterozygous SNPs located throughout the genome of the HTD114 cell line, from which all of the chromosome rearrangements described in this study were generated. In addition, we have identified heterozygous SNPs that reside within the chromosome 6 deleted regions (Figs A1.20 and A1.21). Subsequent PCR and sequencing of the heterozygous SNPs located within the deleted regions confirmed that all of the deletion lines generated here have indeed lost heterozygosity, either proximal or distal to the

*lox*P-3'RT integration site, and retained heterozygosity in the opposite orientation (not shown). In addition, we have determined which SNP alleles are associated with the *lox*P-3'RT integration site in P175 by PCR and sequencing from DNA isolated from mono-chromosomal hybrids containing the individual chromosome 6s (Table A1.2). These data indicate that we have generated a set of nested deletions anchored at the original *lox*P-3'RT integration site in chromosome 6 of P175, and that the proximal deletions range in size from ~30 Mb down to ~76 kb (Table A1.1; Fig. 2.2A), and that the distal deletions range in size from ~26 Mb down to ~18 kb (Table A1.1; Fig. A1.19).

## Replication timing of chromosome 6 deletions

We next assayed the replication timing of the chromosome 6s in a subset of these deletion clones. Cultures were incubated with BrdU for 4.5 h and mitotic cells were harvested, processed for BrdU incorporation and subjected to FISH using chromosome 6-specific probes. For all of the deletion clones assayed, we first used FISH with a chromosome 6-specific paint in conjunction with the BrdU incorporation. This analysis allowed us to visualize all of the chromosome 6 material present in every cell. This was important as chromosomes with DRT/DMC are unstable and frequently participate in secondary rearrangements [10]. Thus, the paints allowed us to assay cells in which additional rearrangements in chromosome 6 were not present (for examples of this assay see Figs A1.3, A1.13, A1.14, A1.15, A1.22 and A1.24). However, this assay did not allow us to unambiguously identify the deleted chromosomes. Therefore, we

also combined the BrdU incorporation assay with FISH using a chromosome 6 centromeric probe plus a BAC or Fosmid from the deleted regions (Fig. 2.2A; Fig. A1.19) on mitotic preparations containing intact chromosomes, as judged by the chromosome paints, that were assayed above. The FISH signal from the chromosome 6 centromeric probe allowed us to identify both chromosome 6s, and the presence or absence of the BAC/Fosmid allowed us to distinguish between the non-deleted or deleted chromosomes, respectively. Comparing the BrdU incorporation pattern of the deleted chromosome 6s to the non-deleted chromosome 6s indicated that the deleted chromosomes present in ∆175-2c (~21.9 Mb deletion; Fig. A1.22), ∆175-2g (~2.1 Mb deletion; not shown), ∆175-3e (~235 kb deletion, not shown), ∆175-1i (~232 kb deletion; Fig. A1.23), ∆175-11d (~96 kb deletion; Fig. A1.24) were delayed in replication timing by at least 2 h. In addition, analysis of cells containing the smallest chromosome 6 proximal deletion (~76 kb deletion), present in the ∆175-23a cell line, indicated that the deleted chromosome was delayed in replication timing by >2 h (Fig. 2.3).

We next assayed the replication timing of the chromosome 6s containing deletions distal to the original *lox*P-3'RT integration site. Figure A1.25 shows the BrdU incorporation analysis, combined with FISH using a chromosome 6 centromeric probe plus BAC CTD-2231M1 (Fig. A1.19) for three independent deletions, and indicates that chromosomes containing deletions of ~25.9 Mb (A-C), ~1.4 Mb (D), and ~17 kb (E) did not display DRT.

## Characterization and Mono-allelic expression of FHLOST

We also attempted to determine if the *FHL5* gene was mono-allelic in our cell system. For this analysis we used RT-PCR followed by sequencing of heterozygous SNPs. We detected mono-allelic expression from within, and telomeric of, the *FHL5* gene (Fig. A1.26; Table A1.2). However, *FHL5* exon-spanning primers failed to generate RT-PCR products (Fig. A1.26B), and these transcripts were generated in the antisense orientation to *FHL5* (Fig. A1.26C), indicating that the *FHL5* protein-coding sequence was not expressed. We have named these *FHL5* opposite strand transcripts *FHL5OST*. This transcript is also known as ENSG00000246665 in Ensemble.

### RNA-DNA FISH for expression of *ASAR6*, *FUT9*, *MANEA* and *FHL5OST*

We have used allele-specific RT-PCR (Figs 2.4A and 2.5A), sequencing of RT-PCR products containing SNPs (Figs 2.4B and C and 2.5B; Table A1.2), RT-PCR on RNA isolated from mono-chromosomal hybrids (Fig. 2.4D; Table A1.2), and RNA-DNA FISH (Figs 2.4F and G, 2.5E and F, and 2.7C-F) to examine the allelic expression of *ASAR6*, *FUT9*, *MANEA* and *FHL5OST*. Figure A1.27 shows additional RNA-DNA FISH images, using three independent probes for *ASAR6*, on a panel of single nuclei from P175 cells. Figure A1.28 shows additional RNA-DNA FISH images, using the *FUT9* cDNA as probe, on a panel of single nuclei from P175 cells. Figure A1.29 shows additional RNA-DNA FISH images, using the *FUT9* cDNA as probe, on a panel of single nuclei from ∆175-1i cells. Figure

A1.30 shows additional RNA-DNA FISH images, using the *FUT9* intronic probe, on a panel of single nuclei from P175 cells. Figure A1.31 shows additional RNA-DNA FISH images, using the *FUT9* intronic probe, on a panel of single nuclei from ∆175-1i cells.

Figure A1.32 shows additional RNA-DNA FISH images, using the *MANEA* cDNA as probe, on a panel of single nuclei from P175 cells. Figure A1.33 shows additional RNA-DNA FISH images, using the *MANEA* cDNA as probe, on a panel of single nuclei from ∆175-1i cells. Figure A1.34 shows RNA-DNA FISH images, for *FHL5OST* using Fosmid G248P86054G4, on a panel of single nuclei from P175 cells. Figure A1.35 shows RNA-DNA FISH images, for *FHL5OST* using Fosmid G248P86054G4, on a panel of single nuclei from ∆175-1i cells.

## Asynchronous replication timing on chromosome 6

Chromosomal loci can be identified as replicating either synchronously or asynchronously using a FISH-based assay [109]. FISH analysis of interphase nuclei pulse-labeled with BrdU allows selective examination of cells in S-phase. This assay also utilizes a methanol/acetic acid fixation, which destroys the nuclear structure and allows for an accurate estimate of replication timing [110-111]. Using a probe to a particular chromosomal site, some cells display two single hybridization dots (an SS pattern), indicating that neither allele has replicated, other cells display two double dots (a DD pattern), indicating that both alleles have replicated, and a third class of cells contains one single dot and one

double dot (an SD pattern), indicating that only one of the two alleles has replicated. For this analysis, we used primary human skin fibroblasts and this FISH assay to quantify the number of hybridization signals per locus present in S-phase nuclei (Table 2.1). Table A1.4 shows the BAC and Fosmid probes and the corresponding chromosomal positions for the loci used in this analysis. Figure A1.35 shows examples of the SS, DD and SD patterns for *ASAR6*. In addition, using BAC probes for synchronously replicating loci [100], indicated that our assay scored the SD pattern within 2% of previously published data for both *LARP* (5q33.2) and *C9orf43* (9q32) (Table 2.1), indicating that our assay produced similar results as this previous study [100].

Previous studies have also shown that the random mono-allelic genes present on autosomes are coordinated in their asynchronous replication so that the alleles on one homolog replicate earlier than the alleles on the other [99-100]. Therefore, we next tested if the asynchronously replicating loci present on chromosome 6 are also coordinated in their asynchronous replication. The level of coordination was examined using a two-color FISH assay and scoring cells that simultaneously displayed the SD pattern for both loci, essentially as described [99-100]. The main text shows examples of coordination of asynchronous replication between *ASAR6* and *FUT9*, *FHL5*, *ME1*, and *HTR1E* (Fig. 2.6). Figure A1.36 shows additional examples of coordination in asynchronous replication in *cis* between *ASAR6* and the closely linked loci BAC959I6 (70/100 cells $P < 1 \times 10^{-4}$; Fig. A1.36A-C) and *MANEA* (78/100 cells $P < 1 \times 10\text{-}5$; Fig. A1.36D-F), and an additional example of coordination in trans

between *ASAR6* and *KCNQ5* (74/100 cells $P < 1 \times 10^{-5}$; Fig. A1.36G-I). In addition, Figure A1.36 also shows examples of coordination in *cis* between the previously reported [86] mono-allelic genes *ME1* and *KCNQ5* (80/100 cells $P < 1 \times 10^{-5}$; Fig. A1.36J-L), and in *trans* between *ME1* and *FUT9* (80/100 cells $P < 1 \times 10^{-5}$; Fig. A1.36M-O). The fact that the assay did not show coordination in all cells suggests that, while the assay is robust ($p < 1 \times 10^{-4}$ for a deviation from 50% in the above examples), it does not allow visualization of the coordination in every cell examined.

Our results indicate that *ASAR6* is coordinated in its asynchronous replication timing in *cis* with the relatively closely linked loci *MANEA*, *FUT9*, BAC959I6, and *FHL5*/*FHL5OST*, suggesting that *ASAR6* is located within a larger domain (>1 Mb) of *cis*-coordinated replication timing. In addition, our data also show that *ASAR6*, and its larger replication timing domain, is coordinated in asynchronous replication in *trans* with other chromosome 6 mono-allelic genes, located >8 Mb either centromeric or telomeric to *ASAR6*. Thus the earlier replicating *ASAR6* allele is coordinated with the later replicating alleles for *HTR1E*, *KCNQ5*, *ME1*, and *FRK* (see Chapter 2 text and Fig. 2.6).

## Expression of *MANEA*, *FHL5*, *FUT9*, and *ASAR6* in primary tissues

To determine the tissue distribution of *ASAR6* expression in relation to *MANEA*, *FHL5*, and *FUT9* we used RT-PCR analysis on a panel of RNA samples isolated from various human tissues. Figure A1.37 shows the RT-PCR analysis for expression of *MANEA*, *FHL5*, *FUT9* and *ASAR6* in the normal human tissue

samples indicated. *MANEA* and *FHL5* were expressed in all tissues assayed, while *FUT9* and *ASAR6* were expressed in a subset of tissues, including: brain, trachea, kidney, testes, and placenta. Interestingly, *FUT9* shows a wider expression profile than *ASAR6*, also showing expression in bladder, prostate, and adipose tissues. This expression pattern for *FUT9* is consistent with a previous report [139]. It will be interesting to determine if *FUT9* is mono-allelically expressed in the tissues that do not express *ASAR6*. Furthermore, it will be interesting to determine if disruption of the *ASAR6* region also results in altered replication timing and bi-allelic expression of these and other mono-allelic genes on chromosome 6 in cells that do not express the *ASAR6* transcripts.

## Figures and Tables

**Figure A1.1. Schematic diagram of our 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 *Neo* or *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 integrate randomly throughout the genome following linearization and electroporation. After Cre transient transfection, reciprocal translocations are 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 are excised as circles via highly efficient (determined to be >90%; data not shown) intra-chromosomal events. Next, Cre directs the remaining *lox*P sites to proceed through a low efficiency ($<1 \times 10^{-3}$) inter-chromosomal reciprocal exchange. This results 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 or pools) [9-10]. (Figure on next page).

**Figure A1.1. Schematic diagram of our chromosome engineering strategy.**

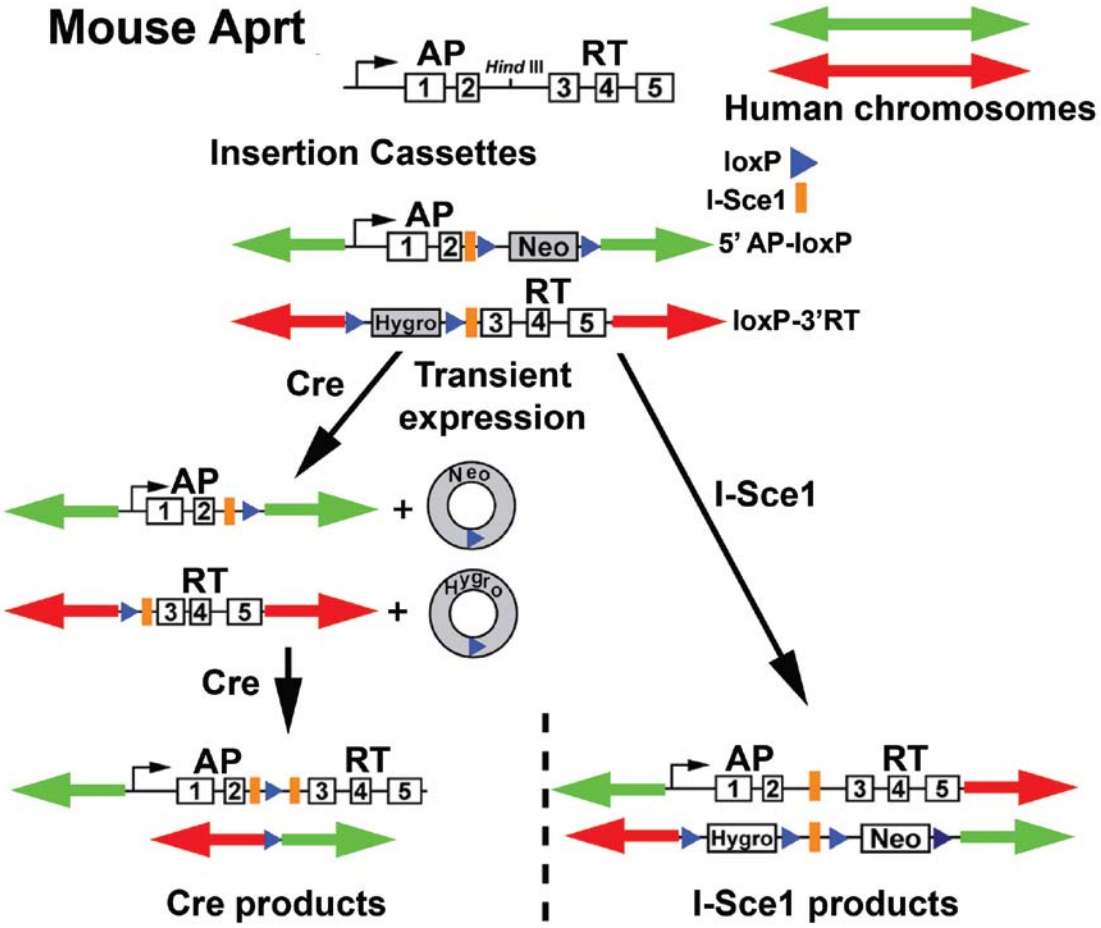


**Figure A1.2. DRT/DMC occurs on only one derivative chromosome of certain balanced translocations.** A schematic diagram of the (**A**) t(3;16)(p13;p13.3) in R27, (**B**) t(6;10)(q15;q11.2) in R175, (**C**) t(der5p;22)(p14;q11.2) in R276, and (**D**) t(15;16)(q24;q12.1) in R268. Analysis of a fifth balanced translocation, a t(3;13)(q29;q14) (not shown) present in R186, showed that both derivative chromosomes could display DMC, indicating that the phenotype is not restricted to a single derivative chromosome [10]. (Figure on next page).

**Figure A1.2. DRT/DMC occurs on only one derivative chromosome of certain balanced translocations.**

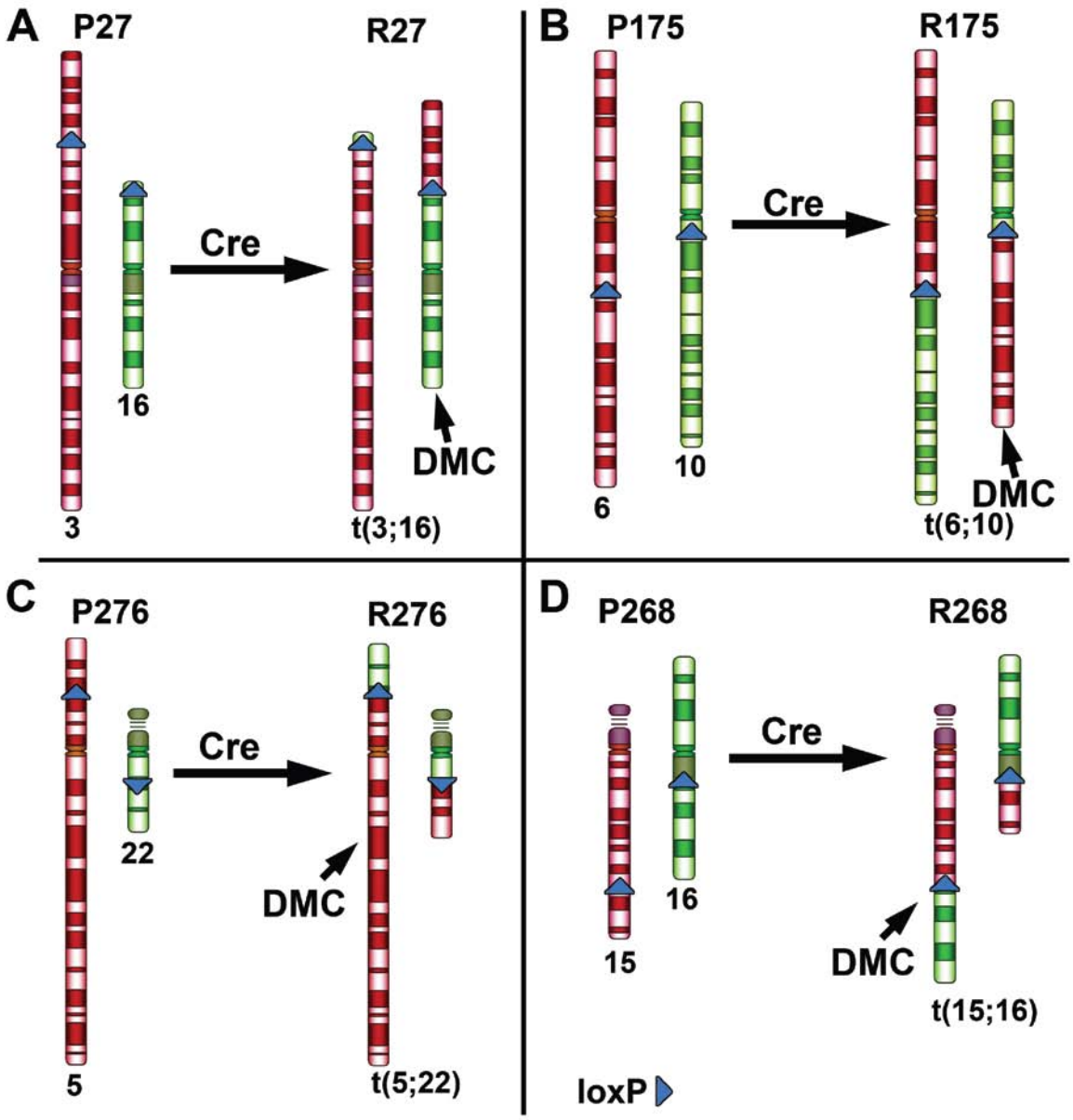


**Figure A1.3. DMC on the chromosome 10 derivative of a t(6;10) balanced translocation.** (**A**) Schematic diagram, (**B**) G-banded chromosomes, and (**C**) DMC on the chromosome 10 derivative. The approximate location of the reconstituted *APRT-lox*P cassettes are shown (**A**). Cells were harvested in the absence of colcemid, dropped on slides and processed by FISH using chromosome 6 and 10 centromeric probes (CHR 6 cen and CHR 10 cen). The non-rearranged chromosomes 6 and 10 and the t(6;10) derivatives were identified by reverse DAPI banding (not shown) and are indicated by arrows. P175, and this t(6;10) have been described previously [10]. (Figure on next page).

**Figure A1.3. DMC on the chromosome 10 derivative of a t(6;10) balanced translocation.**

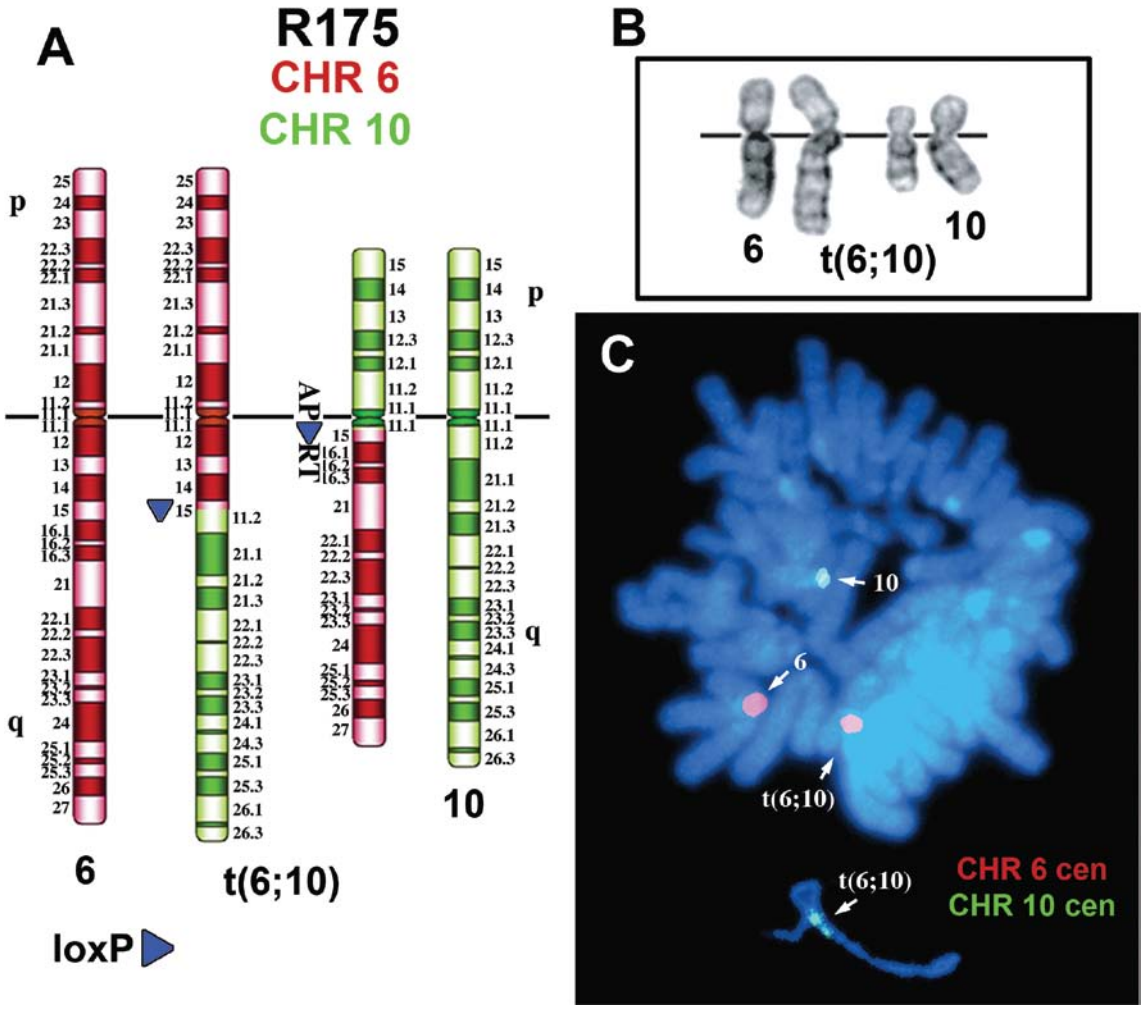


**Figure A1.4. Screen for alternative partner translocations.** (**A**) Genomic structure of the mouse *Aprt* gene. (**B** and **C**) Genomic DNAs from the indicated cell lines were digested with Bcl-1, separated on agarose gels, blotted to nylon membranes and probed with the AP probe, as shown in (**A**). As expected for a single site of integration, P175ΔNH contains a single Bcl-1 restriction fragment that hybridizes to the AP probe (**B** and **C**). In addition, all of the P175(5') clones contain an additional restriction fragment that hybridizes to the AP probe (AP*) but are not present in P175ΔNH (**B** and **C**). These new AP restriction fragments represent the 'new' *lox*P cassette integration sites. The following criteria were used to distinguish between new alternative partner translocations and the original t(6;10) in the R175(5') clones: 1) lack of the ~19kb and ~14kb restriction fragments that are present in the original R175; 2) the presence of a new AP-positive restriction fragment in the R175(5') clones (AP+); 3) loss of the AP-positive restriction fragment present in the P175(5') clones (AP*); and 4) retention of the original ~20 kb AP restriction fragment [AP(CHR10)]. All of the R175(5')-5 clones pass these criteria (**B**), and karyotypic analysis of three of the R175(5')-5 clones indicated that they contain a new alternative partner translocation between chromosomes 6 and 17 [t(6;17)(q15;q25); see below]. All of the R175(5')-6 clones fail these criteria (**B**), and karyotypic analysis of one clone indicated that it contains the original t(6;10), as expected. (Figure on next page).

**Figure A1.4. Screen for alternative partner translocations.**

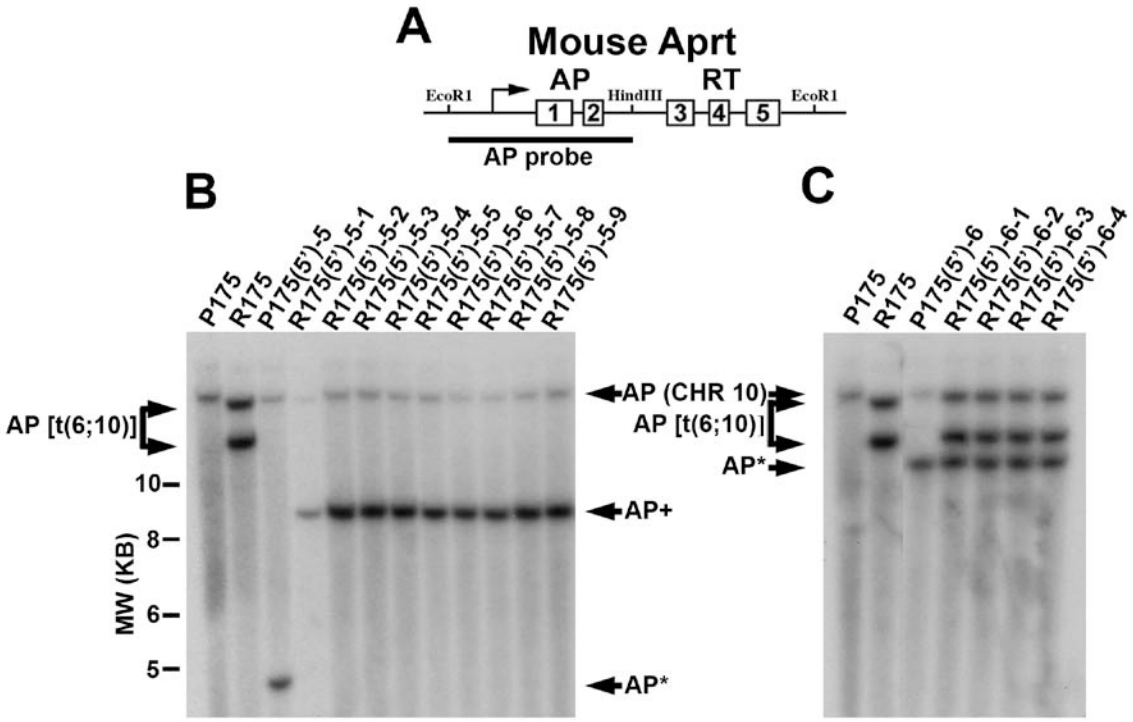


**Figure A1.5. The (6;17) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C**-**F**) on the t(6;17)(q15;q25). (**C**-**E**) R175(5')-17 cells were pulsed with BrdU for 4.5 h, harvested in the absence of colcemid, and processed for FISH using chromosome 6 paint (red) as probe and for BrdU incorporation (green). Inverted DAPI bands were also used to identify the chromosomes [(**D**); DAPI bands]. The non-rearranged chromosomes 6 and 17 and the t(6;17) derivatives are indicated by arrows in panel (**C**). (**C**-**E**) represent chromosomes from the same cell. (**F**) represents a partial karyotype and BrdU incorporation in a second metaphase cell showing the chromosomes 6, 17 and the chromosome 17 derivative of the t(6;17). Double arrows mark the breakpoints in chromosomes 6 and 17 (**E** and **F**). Note the increased BrdU incorporation in the t(6;17). (Figure on next page).

**Figure A1.5. The (6;17) alternative partner translocation.**

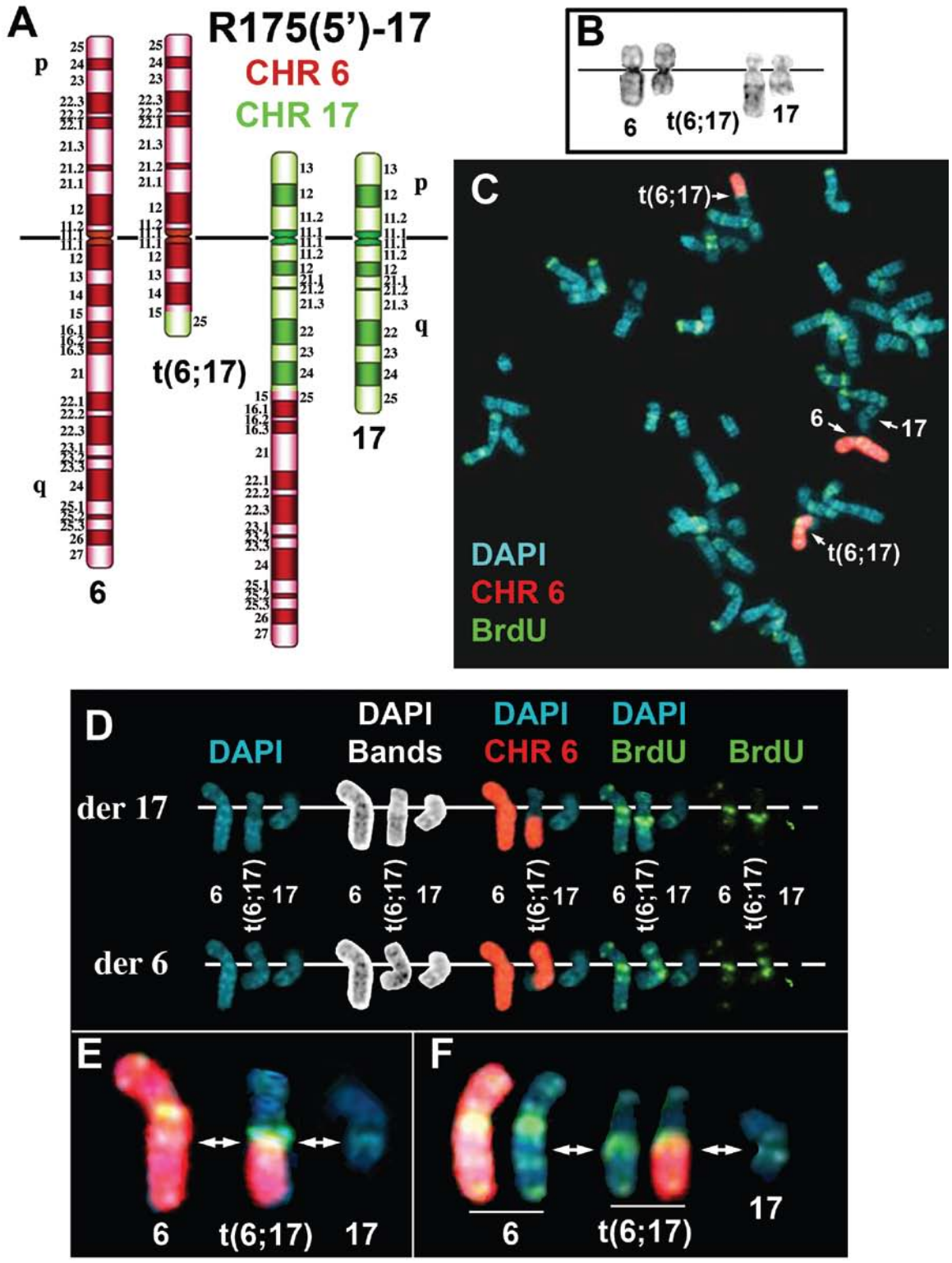
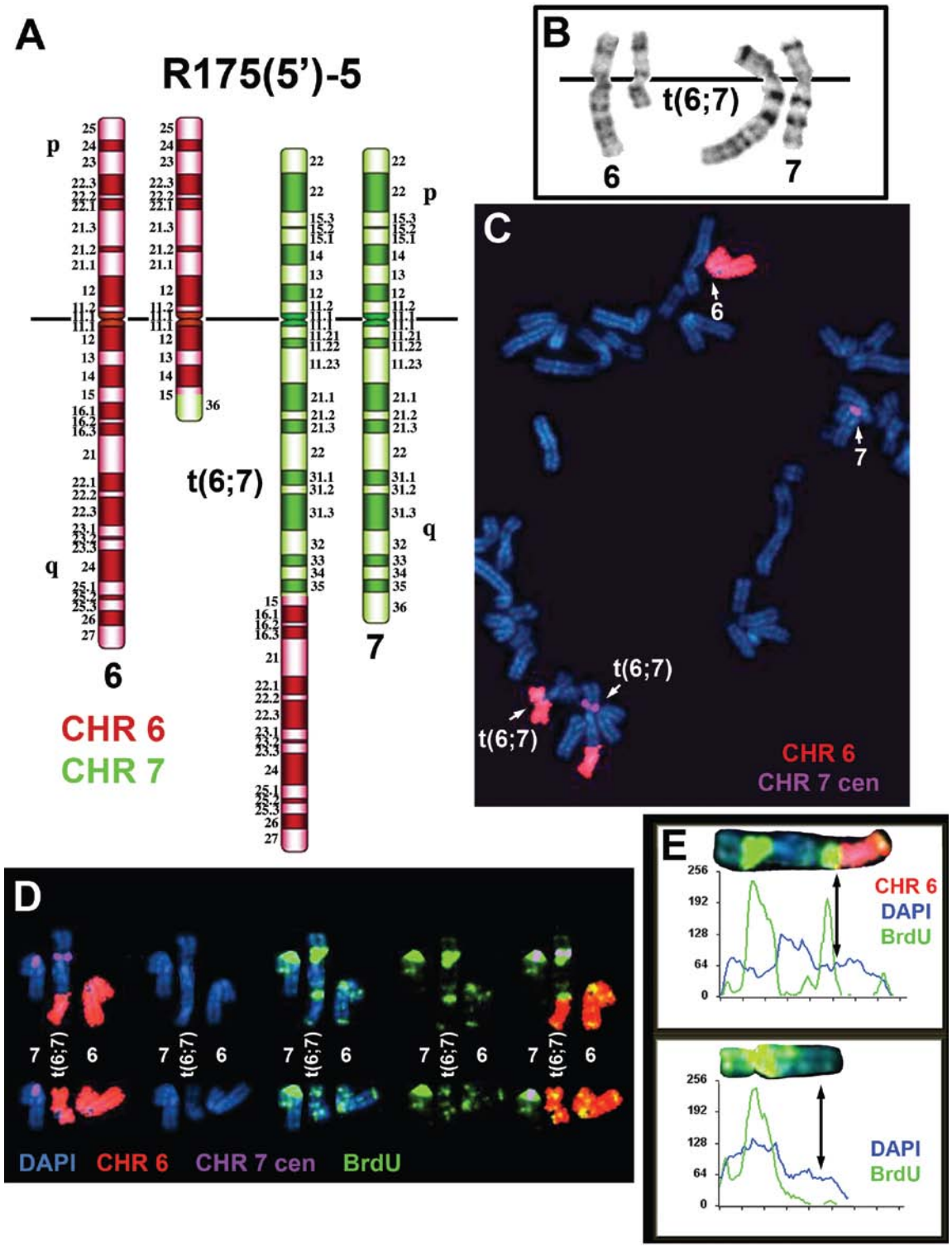


**Figure A1.6. The (6;7) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C**-**E**) on the t(6;7)(q15;q36). (**C**-**E**) R175(5')-5 cells were pulsed with BrdU, harvested in the absence of colcemid, and processed for FISH using chromosome 6 paint (red) and chromosome 7 centromeric (purple) probes and for BrdU incorporation [green, not shown in (**C**)]. Note that the chromosome 7 centromeric probe was pseudo-colored purple for clarity. (**C**-**E**) represent chromosomes from the same cell. The non-rearranged chromosomes 6 and 7 and the t(6;7) derivatives are indicated by arrows (**C**). (**E**) shows the pixel intensity profile for the DAPI staining and BrdU incorporation of the chromosome 7 derivative from the t(6;7) (top panel) and the non-rearranged chromosome 7 (bottom panel). The double arrows indicate the breakpoint. (Figure on next page).

**Figure A1.6. The (6;7) alternative partner translocation.**



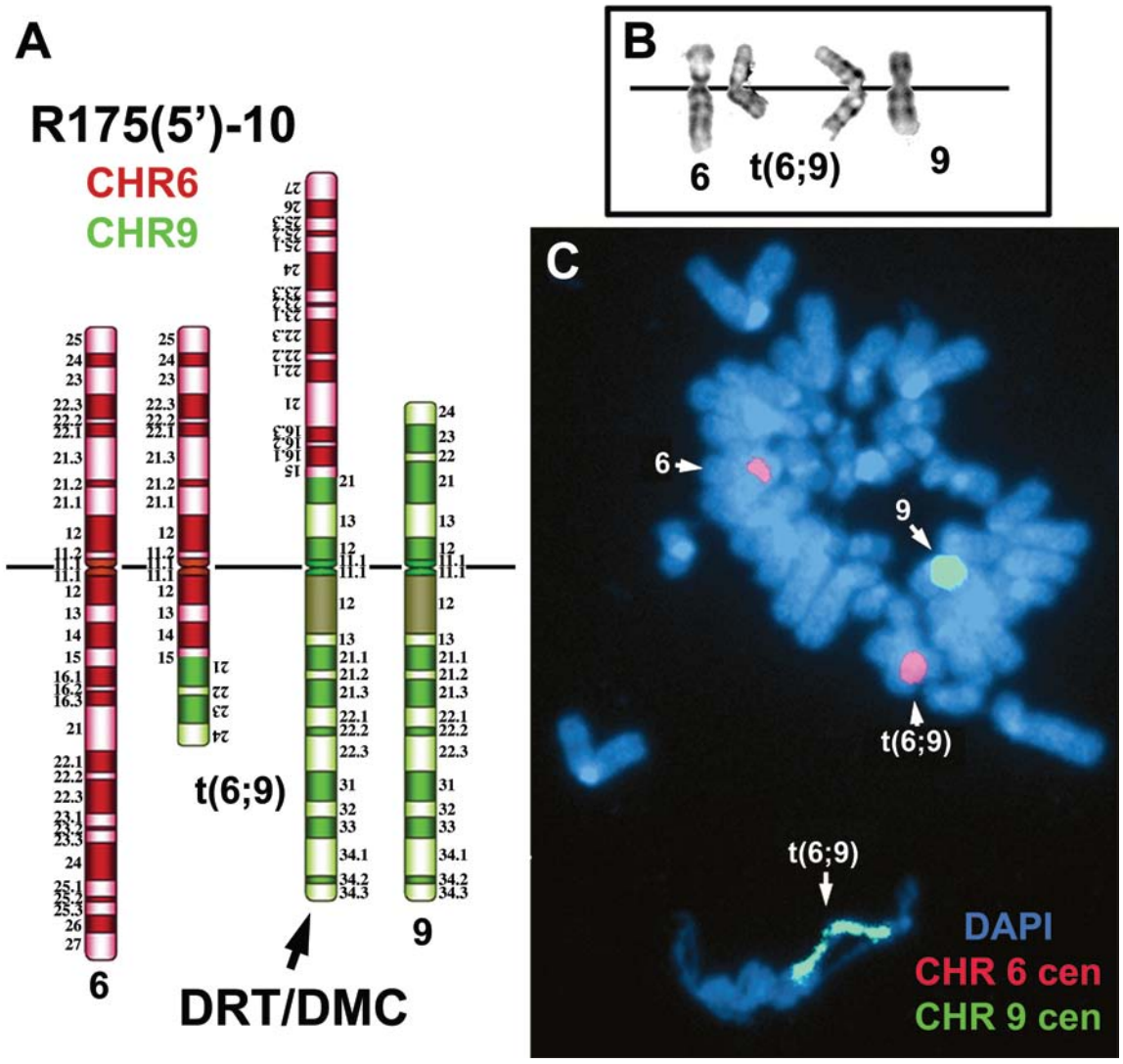



**Figure A1.7. The (6;9) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and DMC assay (**C**) on the t(6;9)(q15;p21). (**C**) R175(5')-10 cells were harvested in the absence of colcemid, and processed for FISH using chromosome 6 paint (red) and chromosome 9 centromeric (green) probes. The non-rearranged chromosomes 6 and 9 and the t(6;9) derivatives are indicated by arrows (**C**). See Figure 2.1 and Figure A1.8 for replication timing assays.

**Figure A1.8. Delayed replication on the t(6;9).** R175(5')-10 cells were incubated with BrdU for 4.5 h and processed for BrdU incorporation (green), and for FISH using probes for chromosome 6 (CHR 6 paint, red) and chromosome 9 centromere (CHR 9 cen, purple) (**A**-**F**). Note that the chromosome 9 centromeres were pseudo-colored from red to purple for clarity. The arrows mark the non-rearranged chromosomes 6 and 9, and both derivative chromosomes of the t(6;9). (**B**) The non-rearranged chromosomes 6 and 9 were aligned with both derivative chromosomes of the t(6;9). (**C**-**F**) Pixel intensity profiles of the BrdU incorporation, and DAPI staining along the long arm (q) of the non-rearranged chromosome 6 (**C**), the long arm of chromosome 6 of the chromosome 9 derivative of the t(6;9) (**D**), the long arm of the non-rearranged 9 (**E**), and the chromosome 9 long arm of the t(6;9) (**F**). The intensity profiles are aligned with the cut out long arms of chromosome 6 (red) or chromosome 9 (purple centromere) for each chromosome (**C** and **D**, or **E** and **F**, respectively). Note that the intensity profile for the BrdU incorporation on 6q of the t(6;9) (**D**) shows three major peaks (i.e. bands) on the chromosome 6 long arm portion, while the long arm of the non-rearranged chromosome 6 shows only a single peak of incorporation near the telomere (**C**). Similarly, the intensity profile for the BrdU incorporation in the chromosome 9 long arm of the chromosome 9 derivative (**F**) shows much more incorporation than the long arm of the non-rearranged chromosome 9 (**E**). The intensity profiles are aligned with the cut out of the long arms of chromosome 9 for each chromosome (**E** and **F**). (**G**) Quantification of the total amount (Area x Intensity) of BrdU incorporation (red) and DAPI staining

(blue) on the long arm of the non-rearranged chromosome 6 (CHR6 6q), the long arm of chromosome 6 on the t(6;9) [t(6;9) 6q], the long arm of chromosome 9 (CHR 9q), the long arm of chromosome 9 on the t(6;9) [t(6;9) 9q], the long arm of chromosome 6 plus the long arm of chromosome 9 on the non-rearranged chromosomes (CHR 6q + 9q), and both arms of the t(6;9) [t(6;9) 6q + 9q]. (Figure on next page).

**Figure A1.8. Delayed replication on the t(6;9).**

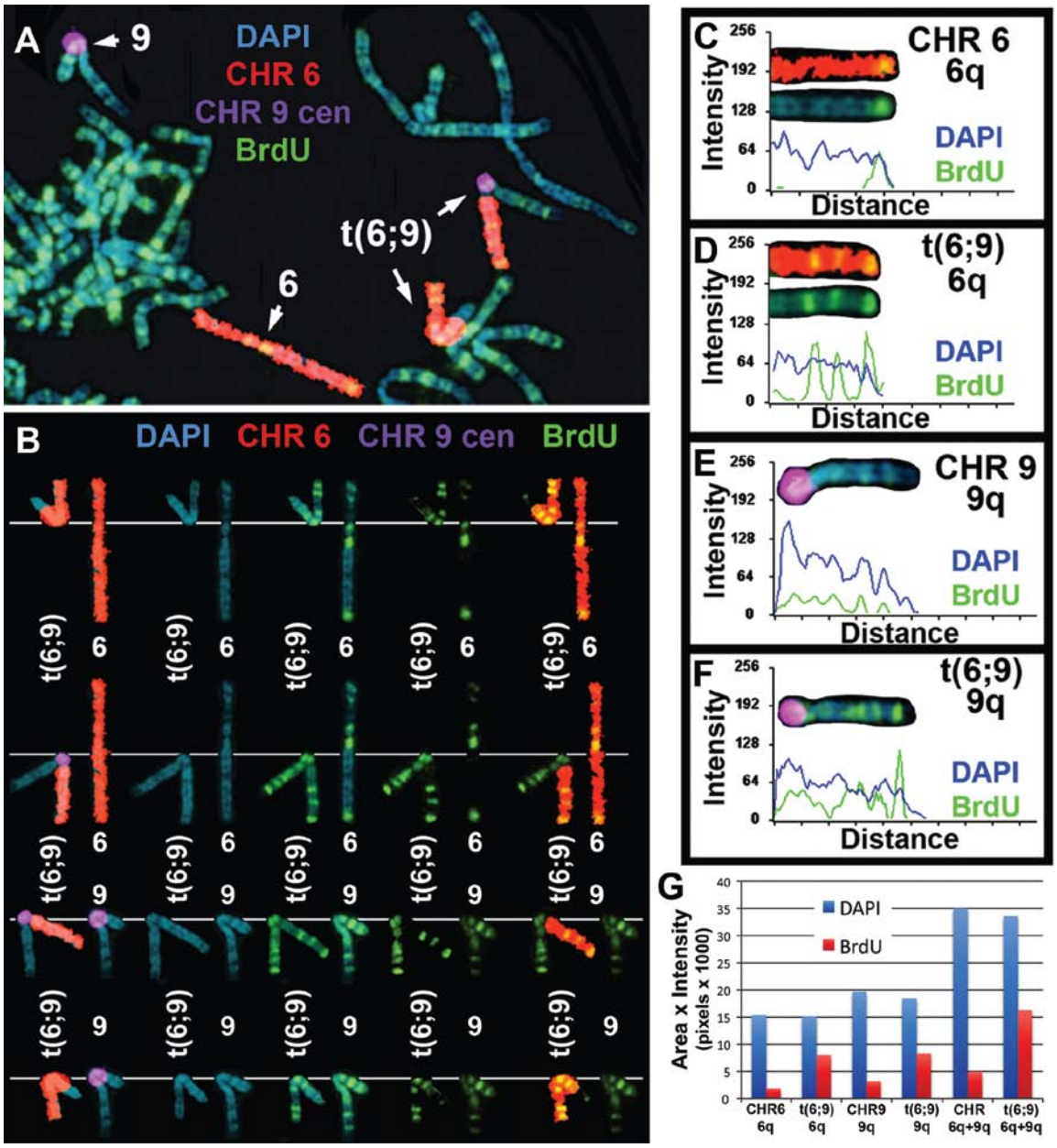


**Figure A1.9. The (6;8) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C** and **D**) on the t(6;8)(q15;q24.1). (**C** and **D**) R175(5')-9 cells were pulsed with BrdU, harvested in the absence of colcemid, and processed for FISH using a chromosome 6 paint (red) as probe and for BrdU incorporation (green). Inverted DAPI bands were also used to identify the chromosomes [(**D**), black and white]. The non-rearranged chromosomes 6 and 8 and the t(6;8) derivatives are indicated by arrows (**C**). The BrdU incorporation of the derivative chromosomes 6 (top line) and 8 (bottom line) of the t(6;8) are compared to the non-rearranged chromosomes 6 and 8 (**D**). (**C** and **D**) represent chromosomes from the same cell. (Figure on next page).

**Figure A1.9. The (6;8) alternative partner translocation.**

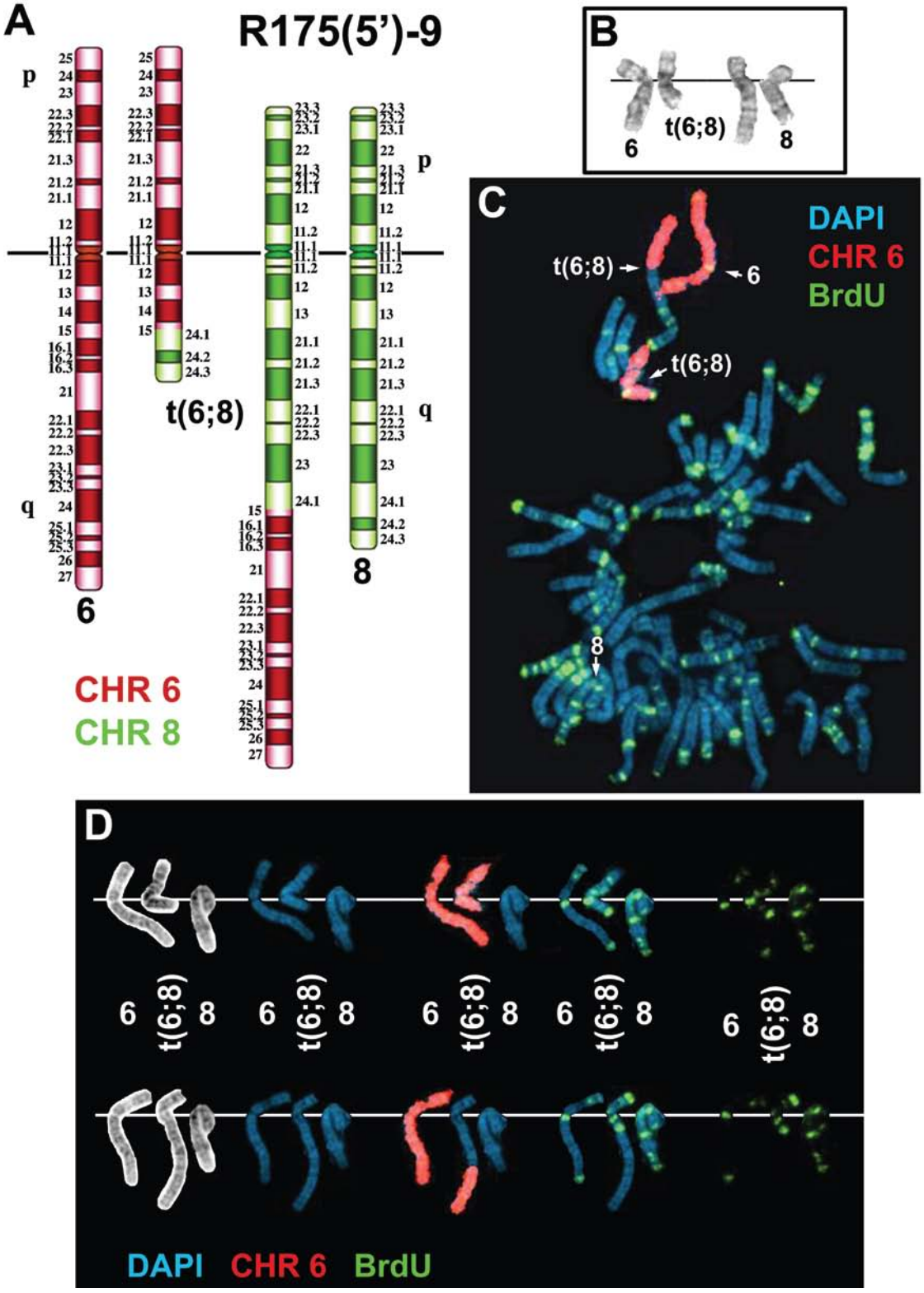


**Figure A1.10. The (1;10) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C** and **D**) on the t(1;10)(p22.3;q11.2). (**C** and **D**) R175(3')-12 cells were pulsed with BrdU, harvested in the absence of colcemid, and processed for FISH using chromosome 10 paint (red) and chromosome 1 centromeric (purple) probes and for BrdU incorporation (green). The non-rearranged chromosomes 1 and 10 and the t(1;10) derivatives are indicated by arrows (**C**). The BrdU incorporation into the derivative chromosomes (1 der and 10 der) of the t(1;10) is compared to the non-rearranged chromosomes 1 and 10 (**D**). (**C** and **D**) represent chromosomes from the same cell. (Figure on next page).

**Figure A1.10. The (1;10) alternative partner translocation.**

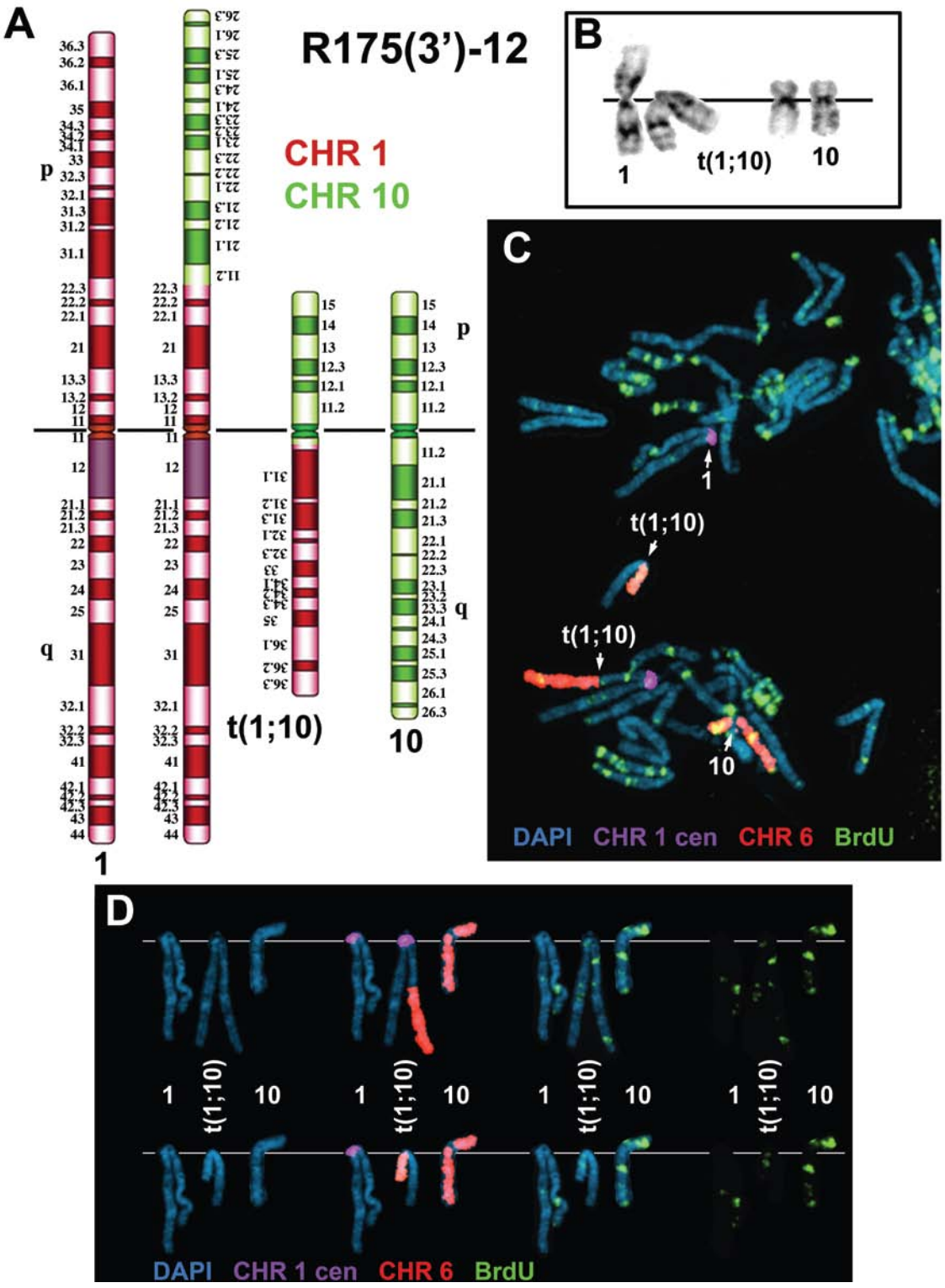


**Figure A1.11. The (4;10) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C** and **D**) on the t(4;10)(q25;q11.2). (**C**) R175(3')-4 cells were pulsed with BrdU, harvested in the absence of colcemid, and processed for FISH using chromosome 10 paint (red) and chromosome 4 centromeric (purple, pseudo-colored for clarity) probes and for BrdU incorporation (green). The non-rearranged chromosomes 4 and 10 and the t(4;10) derivatives are indicated by arrows (**C**). The BrdU incorporation into the derivative chromosomes 4 (top line) and 10 (bottom line) of the t(4;10) are compared to the non-rearranged chromosomes 4 and 10 (**D**). (**C** and **D**) represent chromosomes from the same cell. (Figure on next page).

**Figure A1.11. The (4;10) alternative partner translocation.**

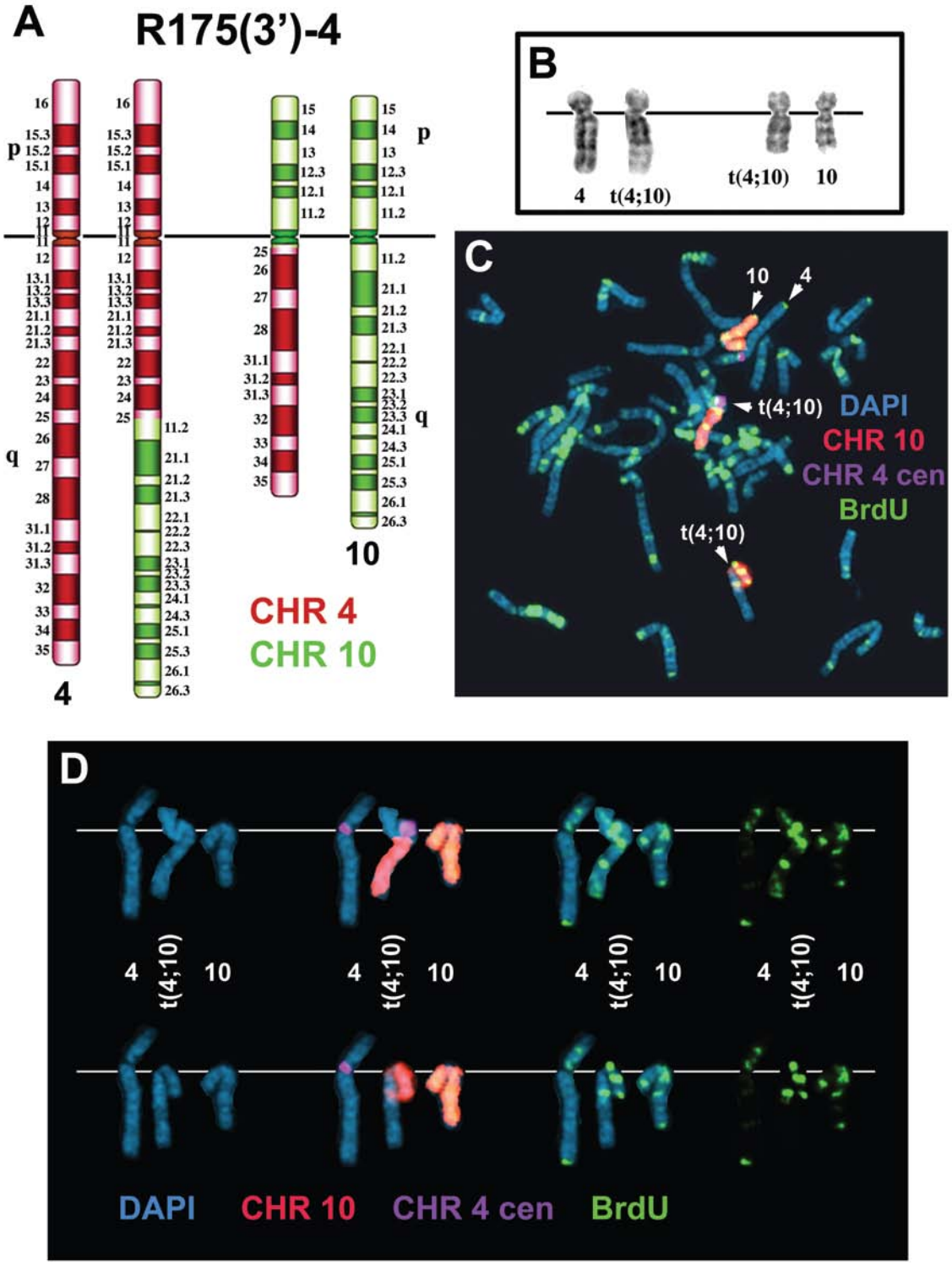


**Figure A1.12. The (5;10) alternative partner translocation.** Schematic diagram (**A**), G-banded chromosomes (**B**), and replication timing assay (**C** and **D**) on the t(5;10)(q35.1;q11.2). (**C** and **D**) R175(3')-6 cells were pulsed with BrdU, harvested in the absence of colcemid, and processed for FISH using a chromosome 10 paint (red) as probe and for BrdU incorporation (green). Inverted DAPI bands were also used to identify the chromosomes from this tetraploid cell (not shown). The non-rearranged chromosomes 5 and 10 and the t(5;10) derivatives are indicated by arrows (**C**). The BrdU incorporation into the derivative chromosomes 5 (top line) and 10 (bottom line) of the t(5;10) are compared to the non-rearranged chromosomes 5 and 10 (**D**). (**C** and **D**) represent chromosomes from the same tetraploid cell. (Figure on next page).

**Figure A1.12. The (5;10) alternative partner translocation.**

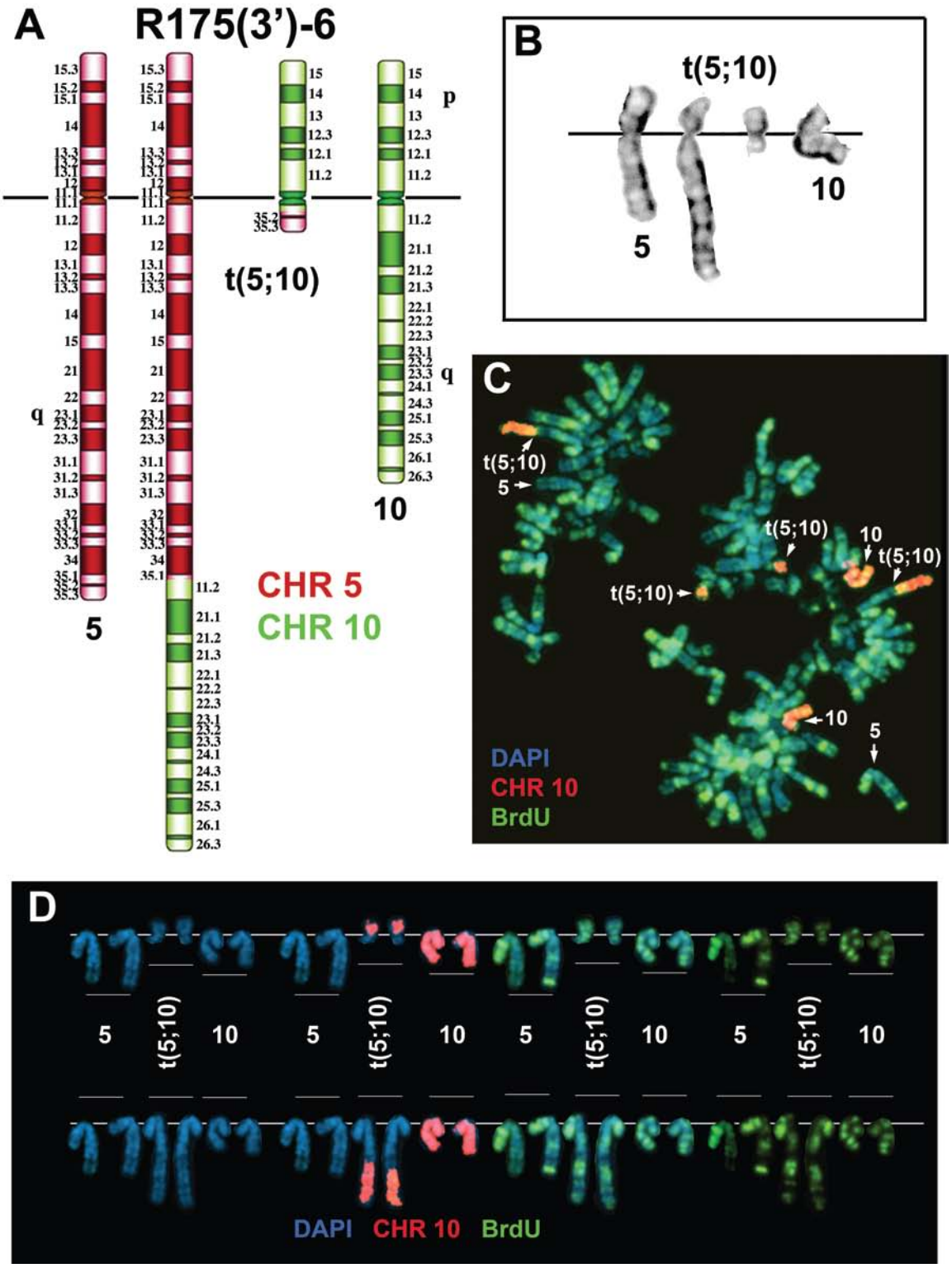


**Figure A1.13. Replication timing assay for chromosome 6 in P175.**

P175 cells were incubated with BrdU for increasing amounts of time followed by mitotic chromosome harvest (see Chapter 2 text and Fig. 2.1B). Fixed cells were dropped on microscope slides and processed for FISH using chromosome-specific paints and for BrdU incorporation. P175 cells were incubated with BrdU for 4.5, 5, 6, 7, 8, and 9 h, harvested for mitotic chromosomes in the absence of a colcemid pretreatment step, processed for FISH using a chromosome 6 paint (red), stained for BrdU incorporation (green), and counterstained with DAPI (blue). The two chromosome 6s from a single cell, 4.5 and 5 h, or four different cells, 6, 7, 8, and 9 h, are aligned with their centromeres positioned on a white line. Note that BrdU incorporation was not detected in any cells in any chromosomes at the 4.5 and 5 h time points, and only single cells as examples are shown. (Figure on next page).

**Figure A1.13. Replication timing assay for chromosome 6 in P175.**

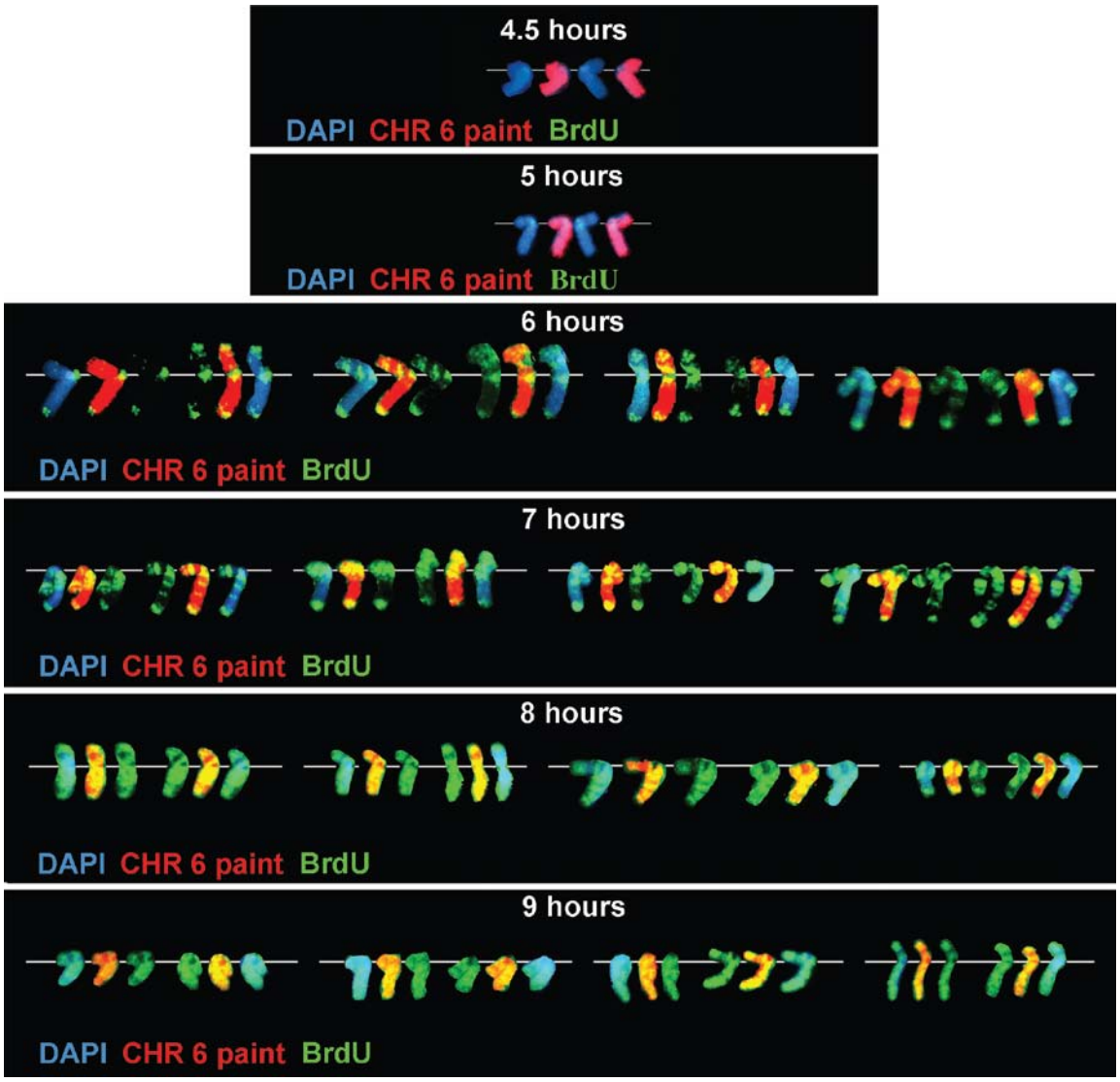
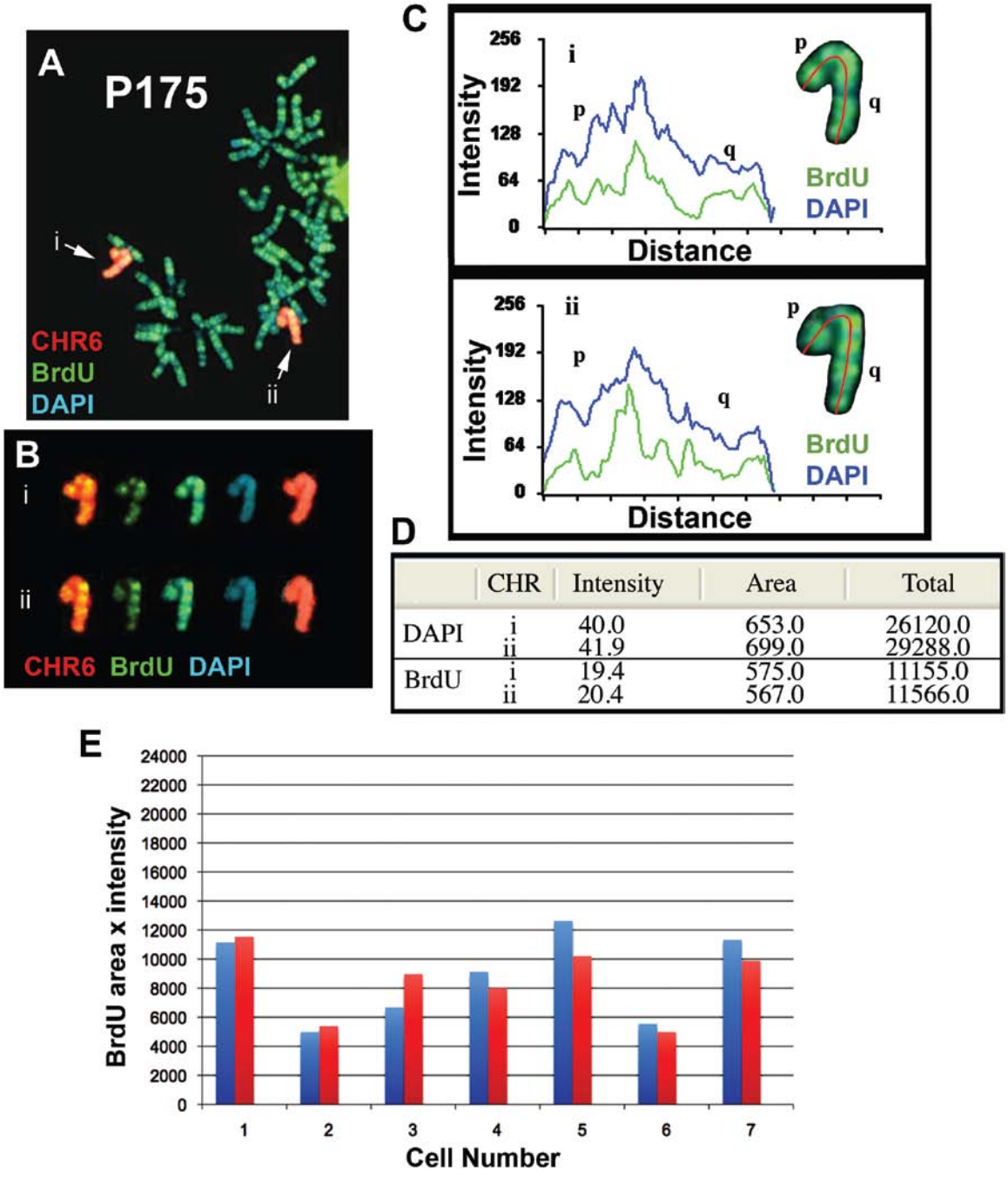


**Figure A1.14. Replication timing analysis of chromosome 6 in P175 cells.** (**A**) An example of a mitotic spread processed for BrdU incorporation (green) and FISH with a chromosome 6 paint as probe (red). (**B**) The two chromosome 6s were separated from the rest of the spread and the different fluors were overlaid in different combinations. (**C**) Intensity profile of the BrdU incorporation and DAPI staining for the two chromosome 6s shown in panels (**A**) and (**B**). The red line on the chromosomes represents the path, starting on the short arm, used to determine the position and location of the signal along the length of each chromosome, the short (p) and long (q) arms are indicated. (**D**) Quantification of the total amount of BrdU incorporation and DAPI staining for the two chromosome 6s shown in panels (**A**-**C**). The values represent pixels. (**E**) Total BrdU incorporation in the chromosome 6s in P175 cells. The red and blue bars represent the two chromosome 6s from 7 different mitotic cells (Cell Number). (Figure on next page).

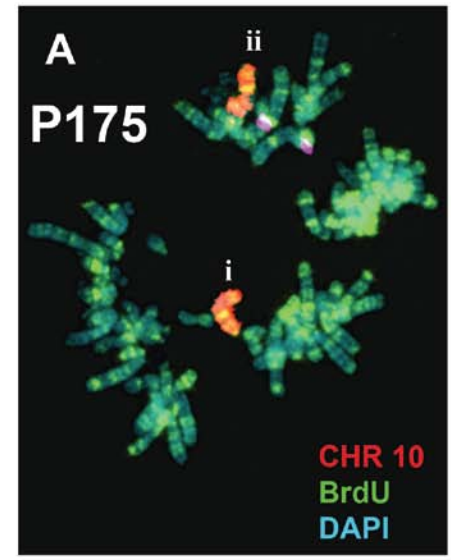
**Figure A1.14. Replication timing analysis of chromosome 6 in P175 cells.**

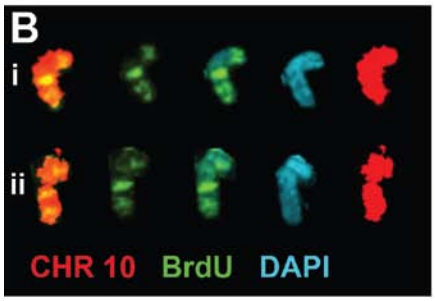


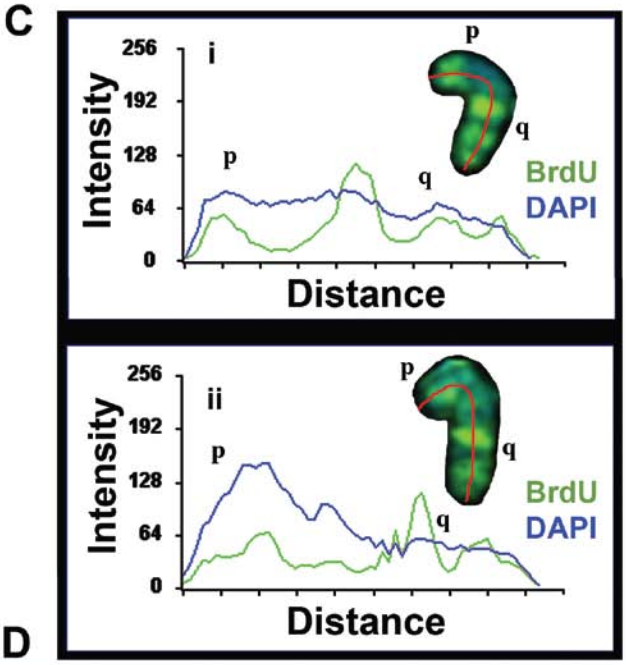

| | CHR | Intensity | Area | Total |
|---|---|---|---|---|
| DAPI | i | 40.0 | 653.0 | 26120.0 |
| | ii | 41.9 | 699.0 | 29288.0 |
| BrdU | i | 19.4 | 575.0 | 11155.0 |
| | ii | 20.4 | 567.0 | 11566.0 |

**Figure A1.15. Replication timing analysis of chromosome 10 in P175 cells.** (**A**) An example of a mitotic spread processed for BrdU incorporation (green), FISH with a chromosome 10 paint as probe (red), and FISH with a chromosome 4 centromeric probe (purple). (**B**) The two chromosome 10s, from panel (**A**), separated from the rest of the spread and aligned with the different fluors overlaid in different combinations. (**C**) Intensity profile for the BrdU incorporation and DAPI staining for the two chromosome 10s shown in panels (**A**) and (**B**). The red line on the chromosomes represents the path used to quantify the position and location of the signal along the length of each chromosome, the short (p) and long (q) arms of the chromosome 10s are indicated. (**D**) Quantification of the total amount of BrdU incorporation and DAPI staining (Total: pixel area x average intensity) for the two chromosome 10s in panels (**A**) and (**B**). (**E**) Total BrdU incorporation in chromosome 10 in P175 cells. The red and blue bars represent the two chromosome 10s from 7 different mitotic cells (Cell Number). The BrdU incorporation profile and quantification for chromosome 4 are not shown (see Fig. A1.16). (Figure on next page).

**Figure A1.15. Replication timing analysis of chromosome 10 in P175 cells.**




D

| | CHR | Intensity | Area | Total |
|---|---|---|---|---|
| DAPI | i | 26.0 | 569.0 | 14794.0 |
| | ii | 29.0 | 621.0 | 18009.0 |
| BrdU | i | 26.1 | 535.0 | 13963.5 |
| | ii | 29.0 | 621.0 | 18009.0 |

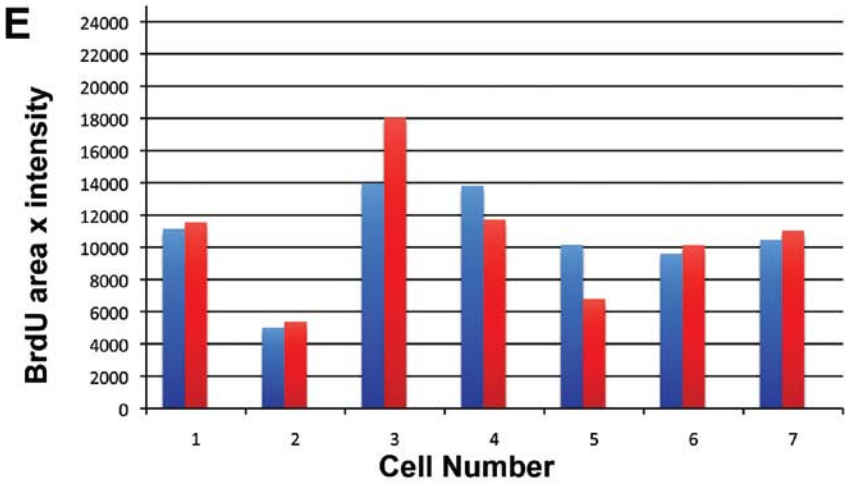


**Figure A1.16. Replication timing analyses of chromosomes 1, 4, 5, 7, 8, 9, and 17 in P175 cells.** Each graph shows the quantification of the total amount of BrdU incorporation (BrdU Area x Intensity) for each chromosome pair. P175 cells were incubated with BrdU for 6 h, processed for BrdU immunostaining, and subjected to FISH with chromosome-specific paints or centromeric probes. The red and blue bars represent the two homologs of the indicated chromosome pairs present in 7 different mitotic cells (Cell number). (Figure on next page).

**Figure A1.16. Replication timing analyses of chromosomes 1, 4, 5, 7, 8, 9, and 17 in P175 cells.**

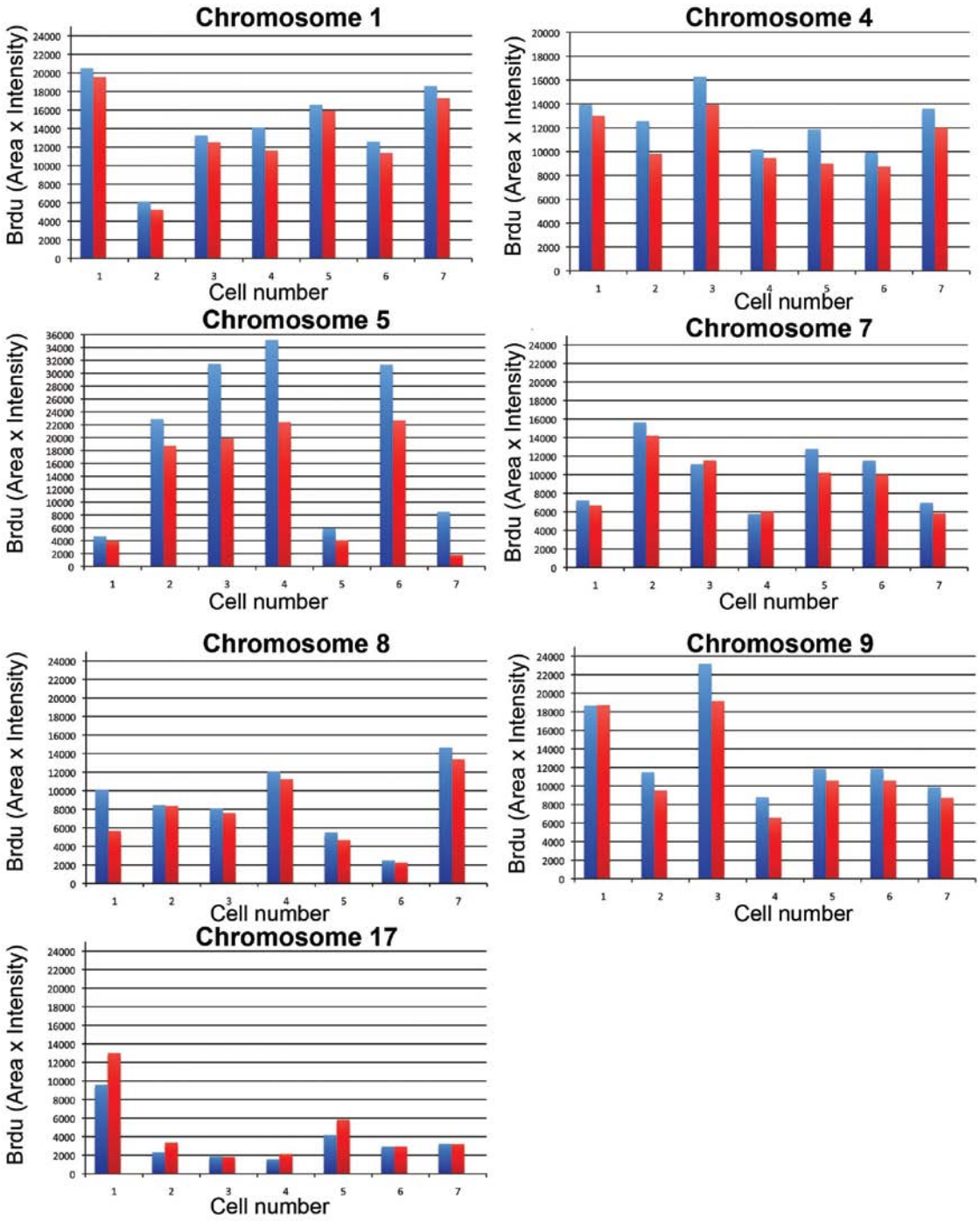


**Figure A1.17. Chromosome 6 proximal deletions.** (**A**) Schematic diagram of the lentiviral vector (AP-*lox*P-Blas) and the *lox*P-3'RT integration site in chromosome 6 of P175. Note that integration of the lentivirus in the correct region and orientation will result in a Cre-dependent loss of the *Blas* gene. An idealized location of LTR-genome junction PCR primers (red arrows) are indicated. The Bcl-1 and HindIII (HIII) sites are indicated. The original *lox*P-3'RT band is indicated (*). (**B**) Southern blot hybridizations showing similar and independent lentiviral integrations in clones from different pools of infected P175 cells, and the lack of the *Blas* gene in most of the Aprt+ clones. The top panel shows a blot (Bcl-1 digestion) hybridized with the AP probe (see panel **A**), and the bottom panel shows a blot (HindIII digestion) hybridized with the *Blas* probe. (**C**) Schematic view of the original *lox*P-3'RT integration site in P175. The integration site was determined by inverse PCR and is located at 96,386,392. The approximate location of the cassette-genome junction-PCR primers (red arrows) are indicated for both the proximal and distal junctions. (**D**) Junction PCR for the detection of deletions. Individual Aprt+ clones isolated from the indicated pools (#1, #2, #3, #4, and #6) were subjected to PCR reactions with the proximal and distal junction-PCR primers (see panel **C**). Note that all of the clones have lost the proximal junction but retain the distal junction. (Figure on next page).

**Figure A1.17. Chromosome 6 proximal deletions.**

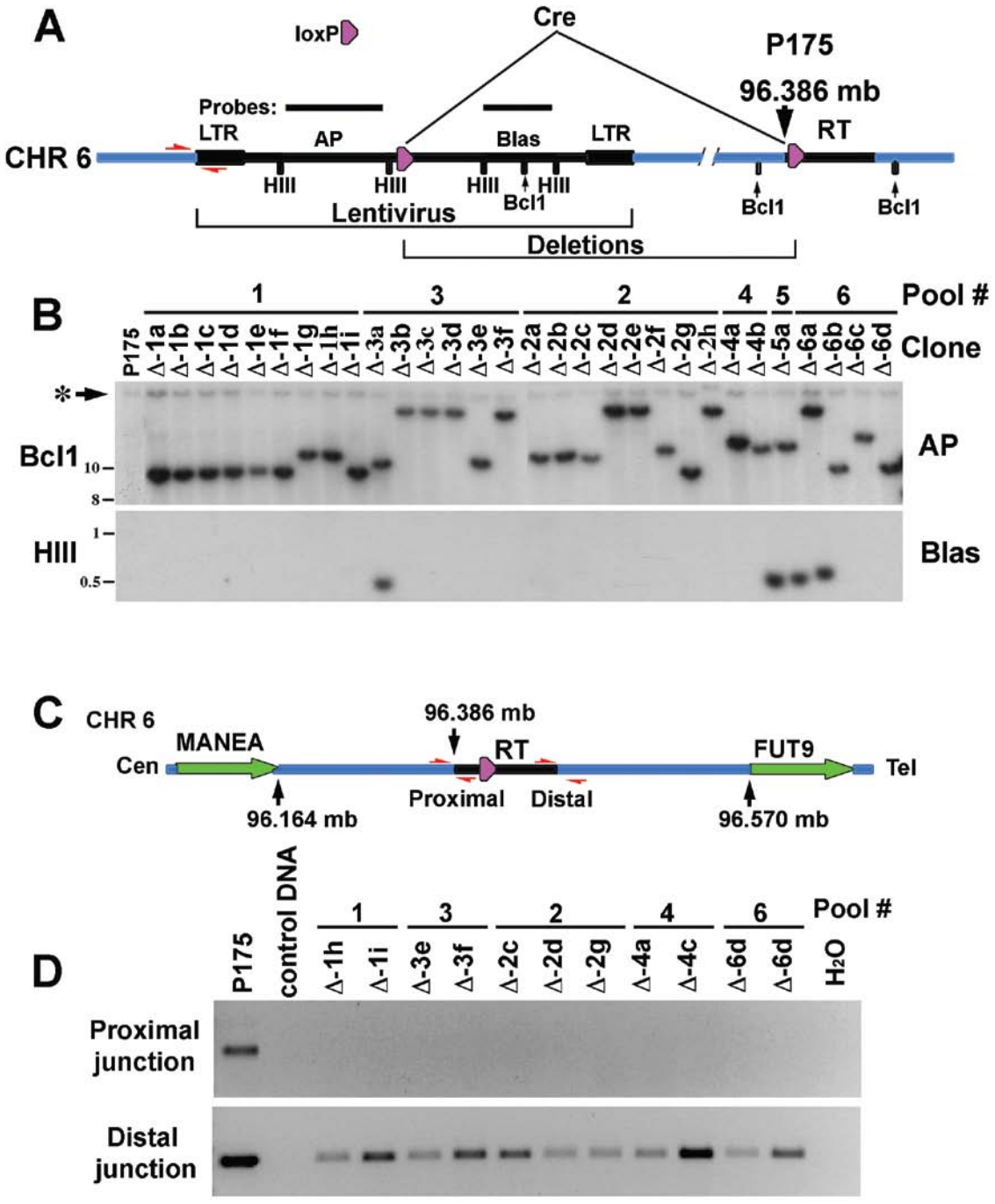
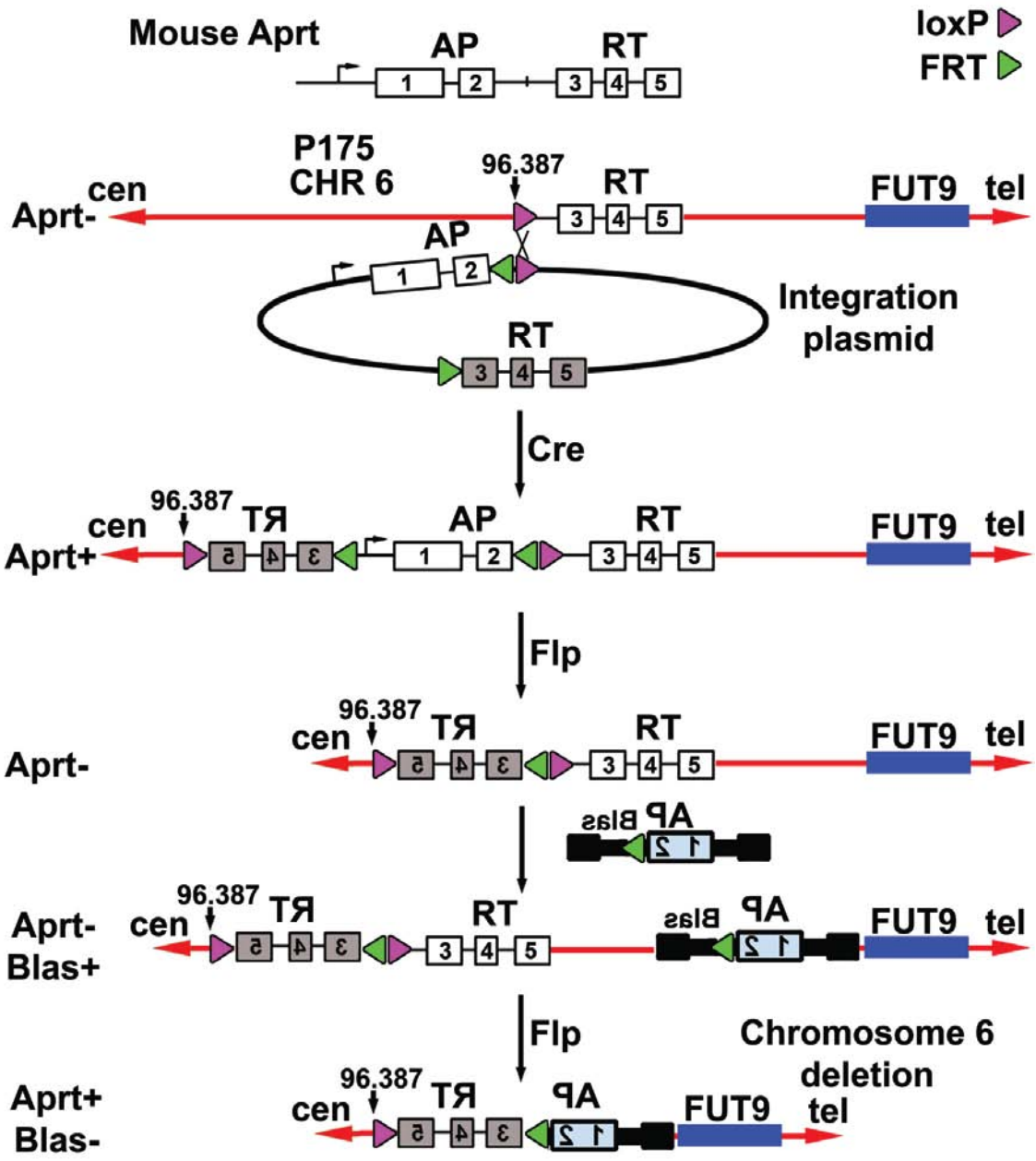


**Figure A1.18. Engineering scheme for chromosome 6 distal deletions.**

Schematic diagram of the mouse *Aprt* gene, the location and orientation of the original *lox*P-3'RT site, the 'integration' plasmid (containing a single *lox*P site and an FRT-RT cassette, the lentiviral vector (L6-AP-FRT-Blas, in reverse in black), and Flp-mediated deletion and reconstitution of *APRT*. Note that integration of the lentivirus in the correct region and orientation will result in a Flp-dependent loss of the *Blas* gene in the final step. (Figure on next page).

**Figure A1.18. Engineering scheme for chromosome 6 distal deletions.**



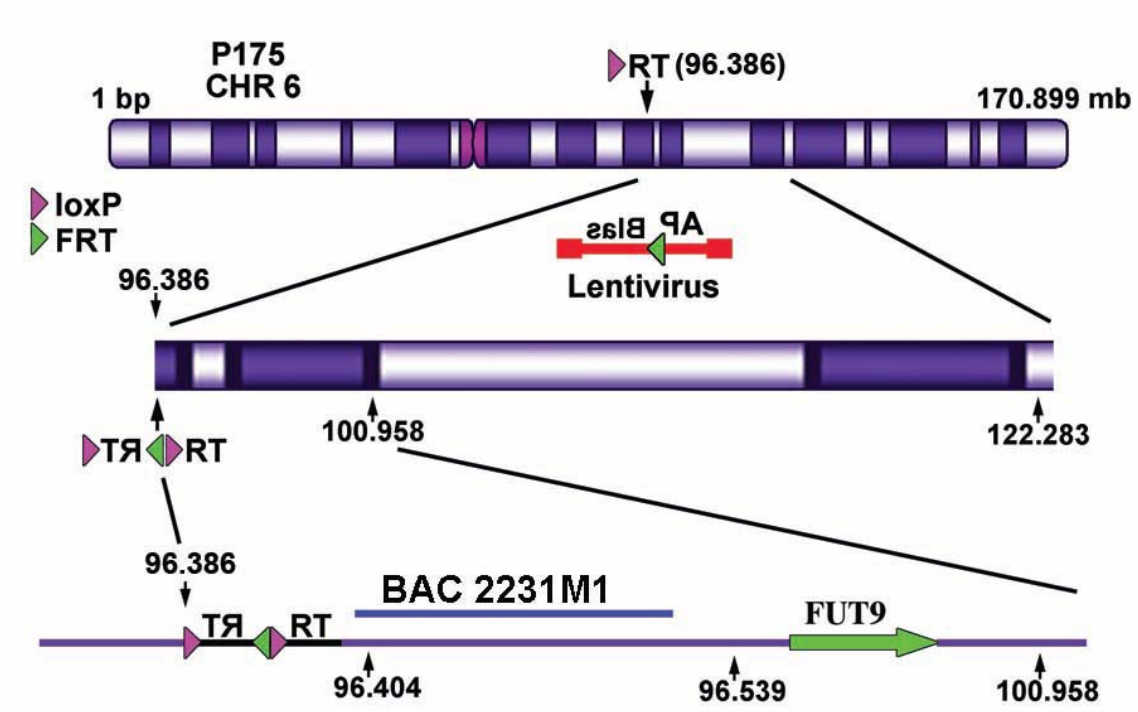



**Figure A1.19. Chromosome 6 distal deletions.** Locations of deletions made in chromosome 6 distal to the *lox*P-3'RT site. Schematic diagram of chromosome 6 in P175 cells showing the integration plasmid at the original *lox*P-3'RT integration site, the location of BAC CTD-2231M1, the lentiviral vector (L6-AP-FRT-Blas; in the reverse orientation), and the lentiviral integration sites along chromosome 6 (arrows, numbers are in megabases; also see Table A1.1). Note that integration of a lentivirus only in the correct region and orientation will result in Cre-dependent loss of the *Blas* gene.

**Figure A1.20. SNP analysis of an ~16.58 Mb chromosome 6 deletion.**

Affymetrix 500K SNP arrays were performed on two independent isolates of genomic DNA from the starting cell line HTD114 (H), and single genomic DNA isolates of P175 (P) and Δ175-2d (Δ). Δ175-2d is predicted to have an ~16.58 Mb deletion, based on the LAM-PCR integration site (Table A1.1). The left and middle panels show the heterozygous SNPs, defined as AB calls in both runs of the HTD114 DNA, comparing HTD114 to P175 and Δ175-2d. The genotype calls were imported into Excel and color-coded, with the AA genotype in green, the BB genotype in blue, and the AB genotype in red. The SNP ID#, chromosome (CHR), position and genotype calls are indicated. The middle panel shows an ~18 Mb region of chromosome 6, and the left panel shows an expanded view of this region. The genomic locations of the original *lox*P-3'RT and lentiviral integration sites are shown (* and +, respectively). The right panel shows the copy number and Heat Map analysis for a pair-wise comparison between P175 and Δ175-2d for chromosome 6. The positions of the short (p) and long (q) arms, as well as the centromere (c) are indicated. (Figure on next page).

**Figure A1.20. SNP analysis of an ~16.58 Mb chromosome 6 deletion.**

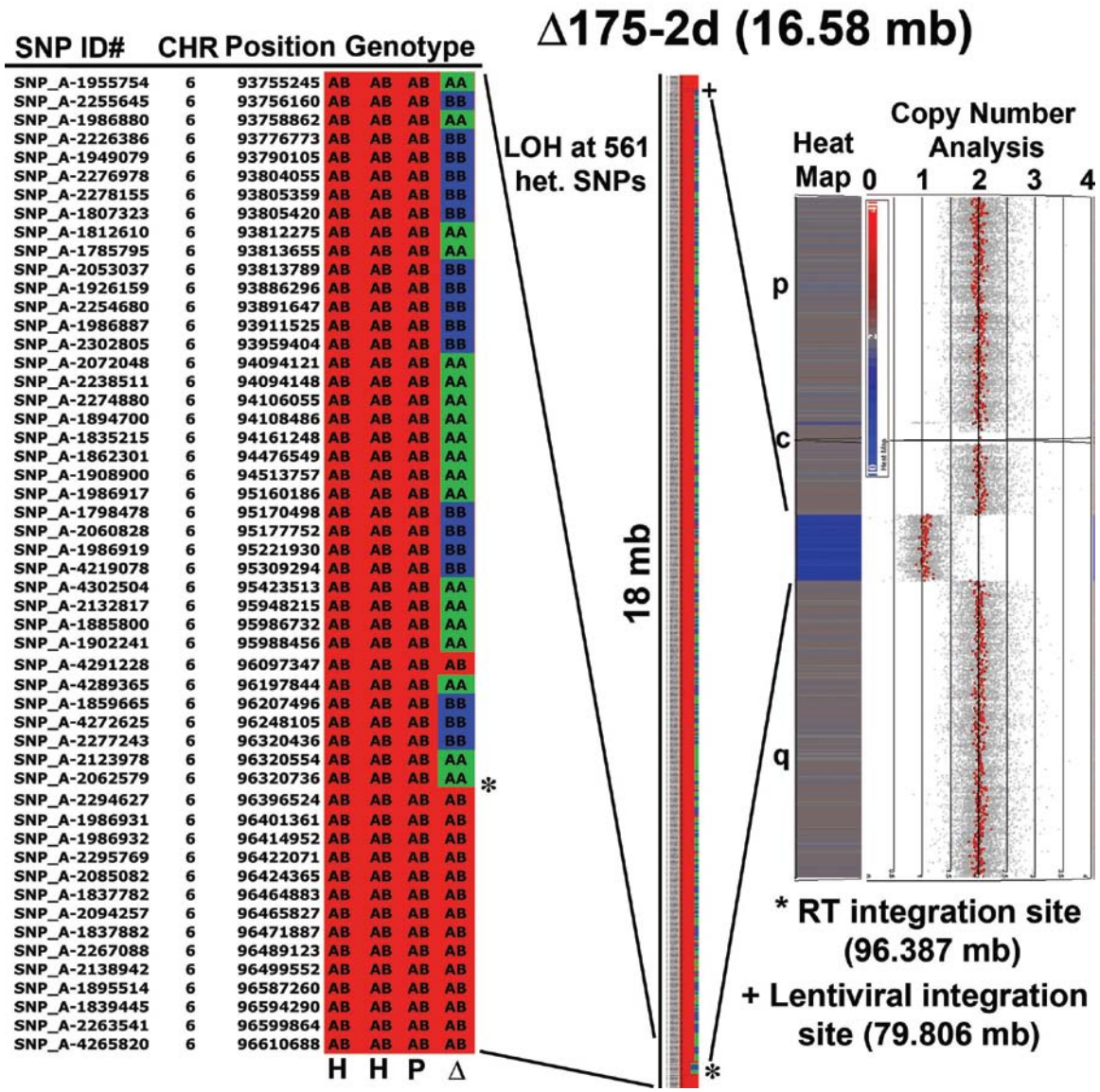
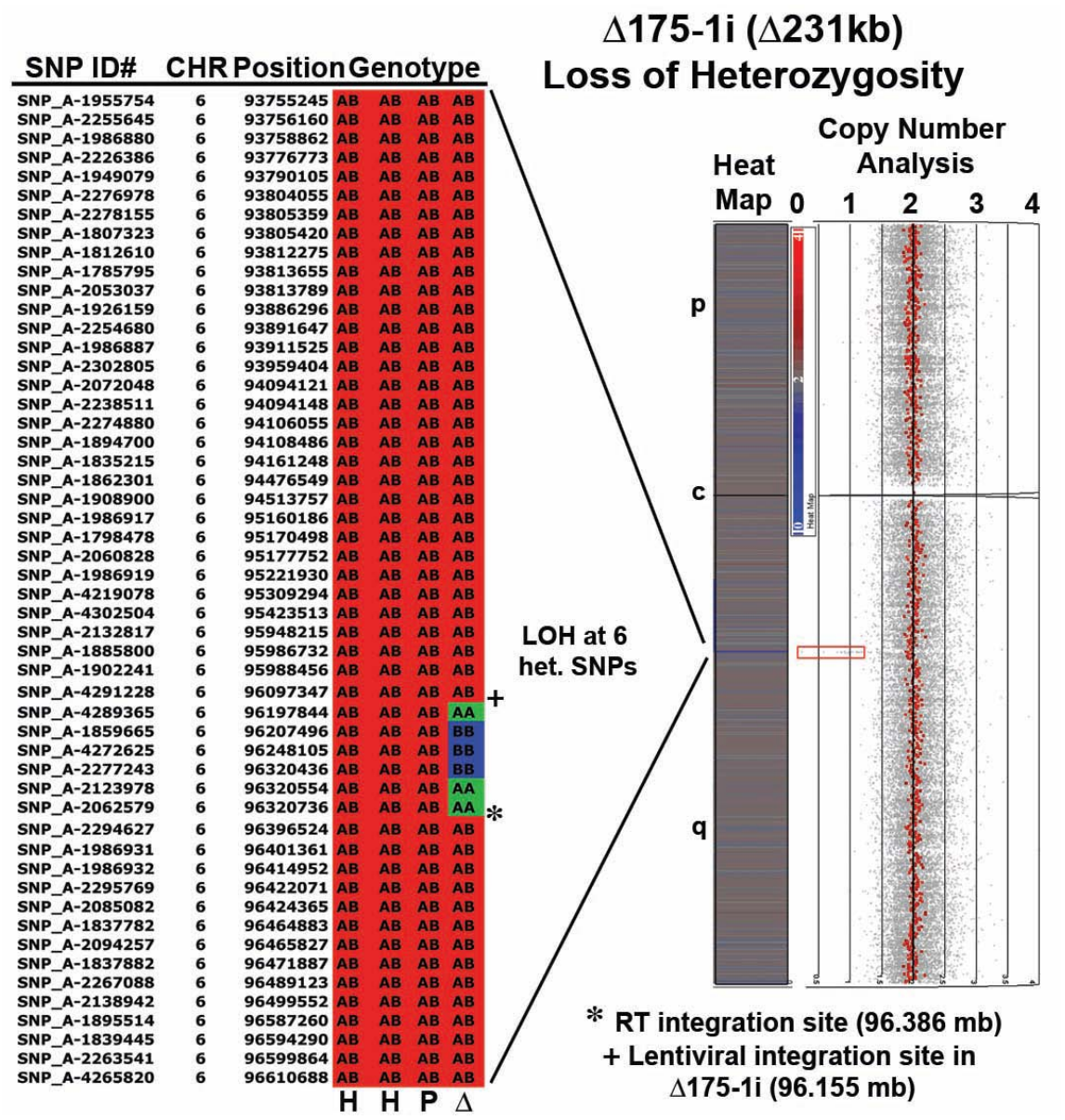


**Figure A1.21. SNP analysis of an ~231 kb chromosome 6 deletion.**

Affymetrix 500K SNP arrays were performed on two independent isolates of genomic DNA from the starting cell line HTD114 (H), and single genomic DNA isolates of P175 (P) and Δ175-1i (Δ). Δ175-1i is predicted to have an ~231 kb deletion. The left panel shows the heterozygous SNPs, defined as AB calls in both runs of the HTD114 DNA, comparing the genotype of HTD114 to P175 and Δ175-1i. The genotype calls were imported into Excel and color-coded, with the AA genotype in green, the BB genotype in blue, and the AB genotype in red. The SNP ID#, chromosome (CHR), position and genotype calls are indicated. The left panel shows an ~2.8 Mb region of chromosome 6. The genomic locations of the original *lox*P-3'RT and lentiviral integration sites are shown (* and +, respectively). The right panel shows the copy number and Heat Map analysis for a pair-wise comparison between P175 and Δ175-1i for chromosome 6. The positions of the short (p) and long (q) arms, as well as the centromere (c) are indicated. The red box highlights the region of chromosome 6 that shows decreased copy number and LOH. (Figure on next page).

**Figure A1.21. SNP analysis of an ~231 kb chromosome 6 deletion.**



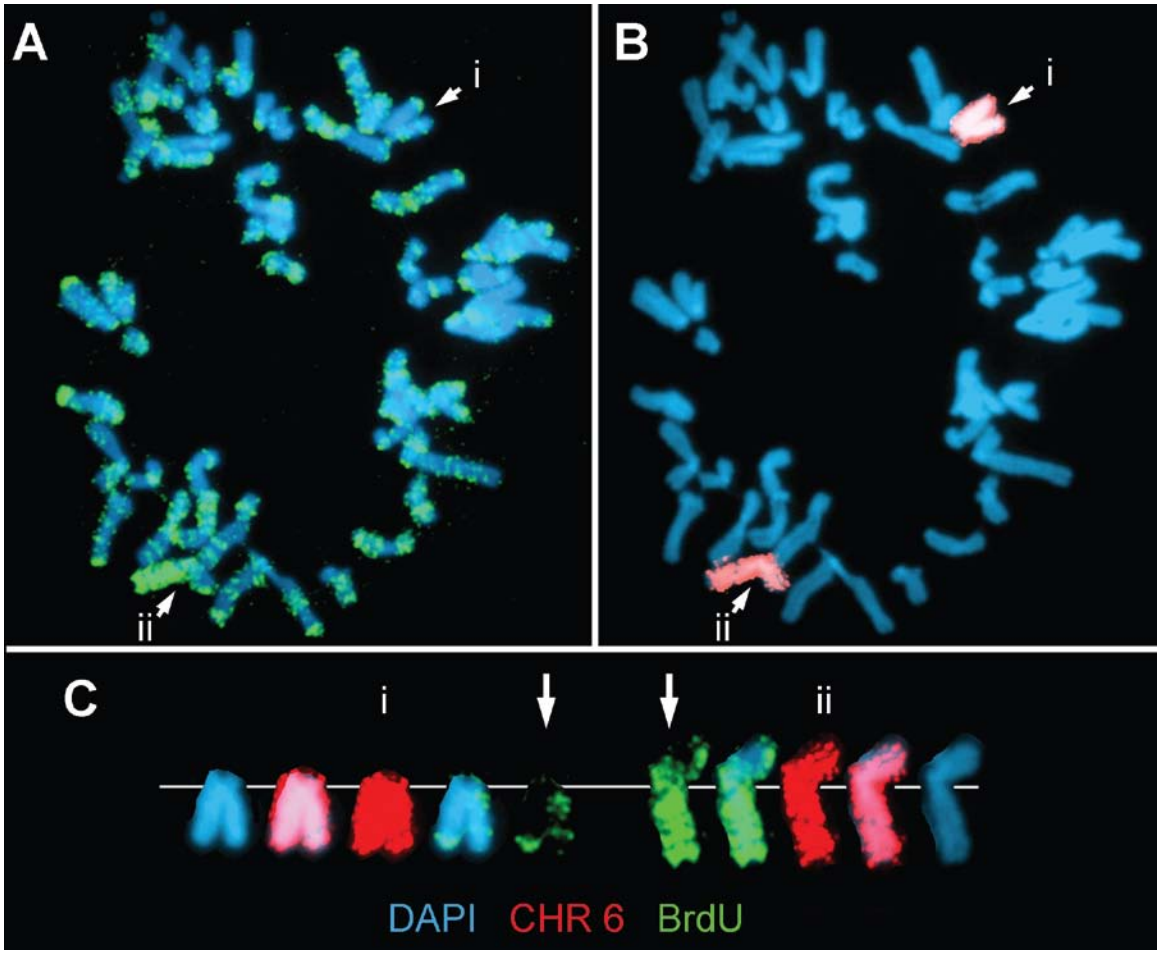
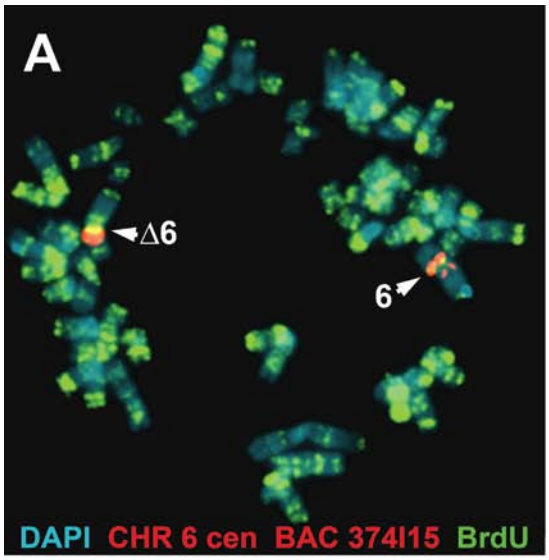
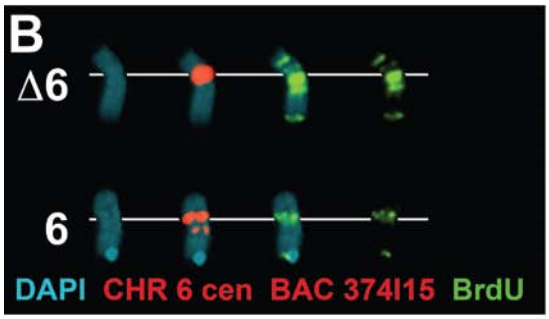


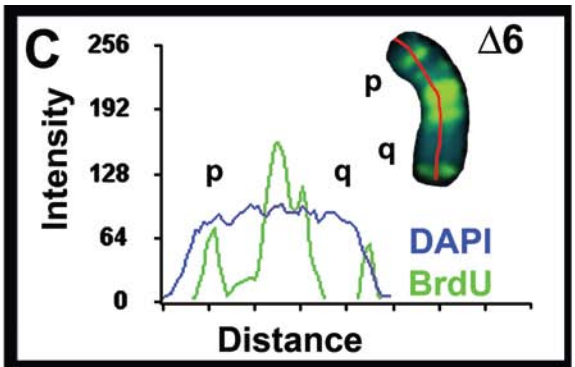

**Figure A1.22. Replication timing assay for chromosome 6 in Δ175-2c (~21.9 Mb deletion).** Cells were incubated with BrdU for 4.5 h followed by mitotic chromosome harvest. Fixed cells were dropped on microscope slides and processed for FISH using a chromosome 6 paint (red), stained for BrdU incorporation (green), and counterstained with DAPI (blue). (**A** and **B**) Mitotic chromosome spread showing the two chromosome 6s (**i** and **ii**), BrdU incorporation (green in panels **A** and **C**) and chromosome 6 paint (red in panels **B** and **C**). (**C**) The two chromosome 6s from the same cell as in (**A** and **B**), are aligned with their centromeres positioned on a white line. The arrows mark the FITC signal from the BrdU incorporation for comparison.

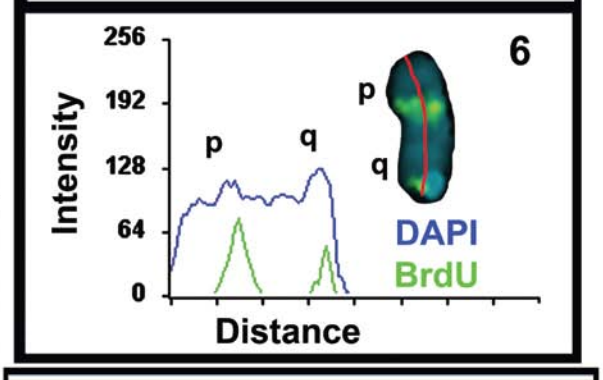
**Figure A1.23. Replication timing of chromosome 6 in ∆175-1i cells (~231 kb deletion).** (**A**) An example of a mitotic spread processed for BrdU incorporation (green) and FISH with chromosome 6 centromere (red) and BAC3741I5 (red) probes. The deleted chromosome 6 (∆6) lacks the BAC signal. (**B**) The two chromosome 6s were separated from the rest of the spread and the different fluors were overlaid in different combinations. (**C**) Intensity profile of the BrdU incorporation and DAPI staining for the two chromosome 6s shown in panels (**A**) and (**B**). The red line on each chromosome represents the path, starting on the short arm, used to determine the distance and location of the signal along the length of each chromosome, the short (p) and long (q) arms of the chromosome 6s. (**D**) Quantification of the total amount of BrdU incorporation and DAPI staining for the two chromosome 6s shown in panels (**A**-**C**). The values represent pixels. (**E**) Total BrdU incorporation in the chromosome 6s in ∆175-1i cells. The red (6) and blue (∆6) bars represent the two chromosome 6s from 7 different mitotic cells (Cell Number). (Figure on next page).

**Figure A1.23. Replication timing of chromosome 6 in Δ175-1i cells (~231 kb deletion).**

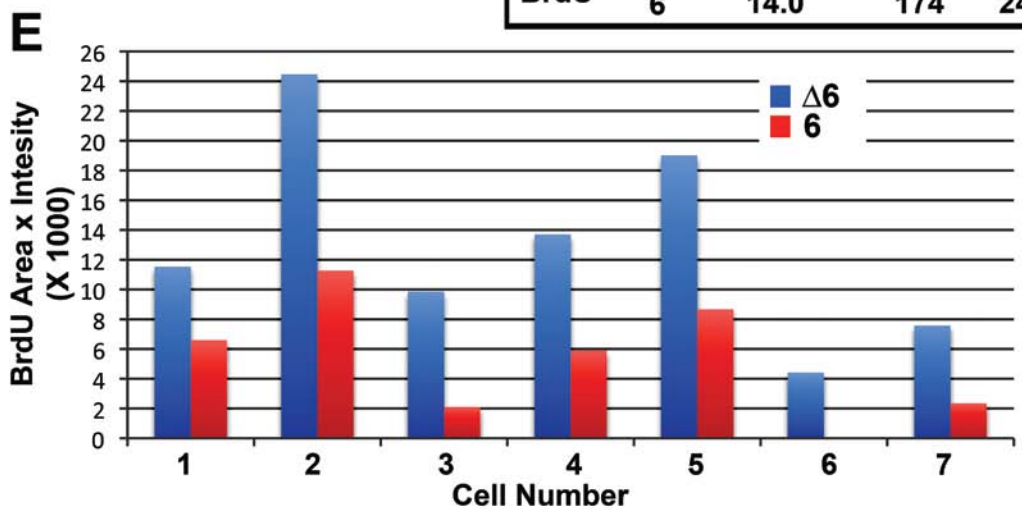



| D | CHR | Intensity | Area | Total |
|---|---|---|---|---|
| DAPI | Δ6 | 29.6 | 623 | 18441 |
| | 6 | 38.1 | 586 | 22327 |
| BrdU | Δ6 | 27.6 | 357 | 9858 |
| | 6 | 14.0 | 174 | 2436 |



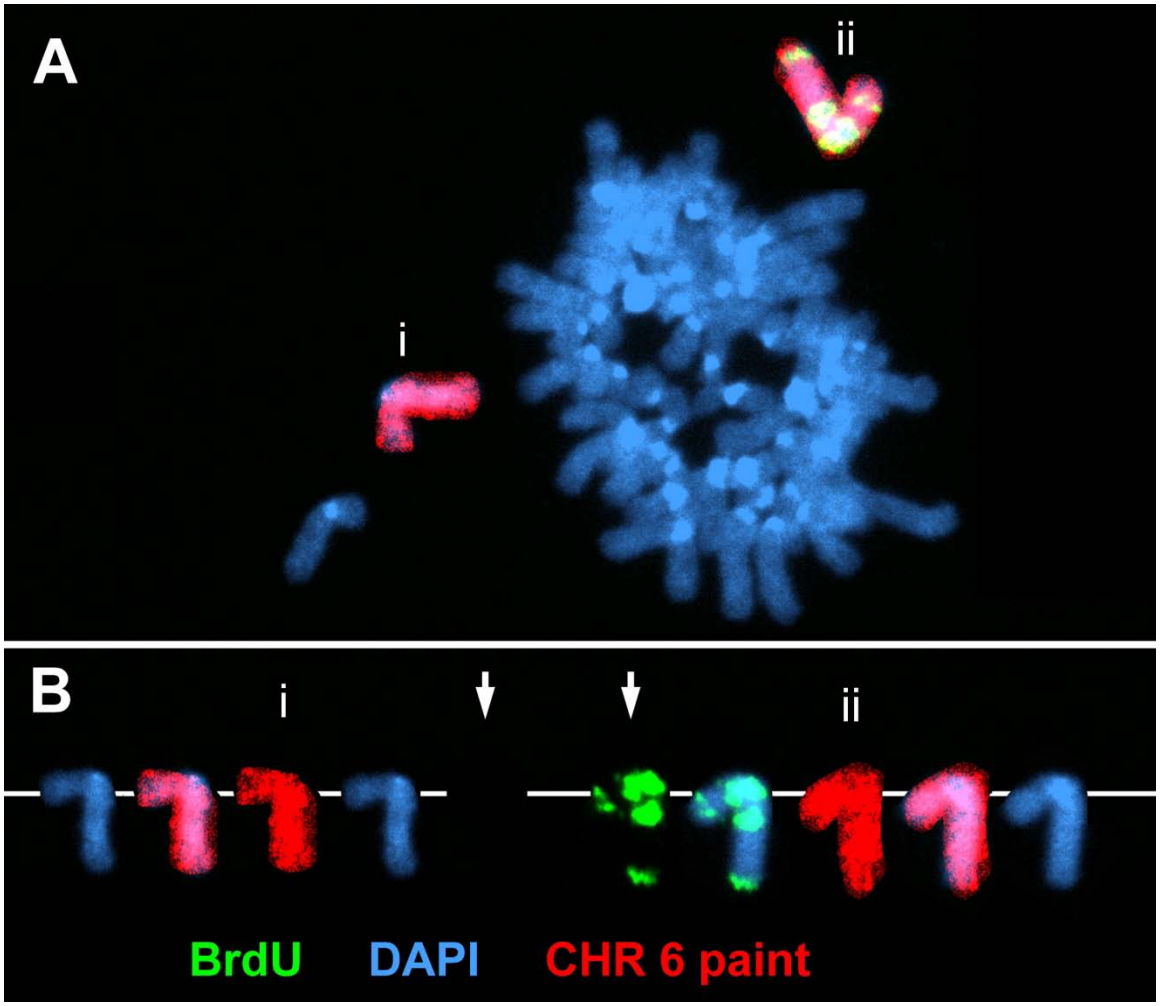



**Figure A1.24. Replication timing assay for chromosome 6 in ∆175-11d (~94 kb deletion).** Cells were incubated with BrdU for 4.5 h followed by mitotic chromosome harvest. (**A**) Mitotic chromosome spread showing the two chromosome 6s (**i** and **ii**), BrdU incorporation (green) and chromosome 6 paint (red). (**B**) The two chromosome 6s from the same cell as in (**A**), are aligned with their centromeres positioned on a white line with the different flours in different combinations. Arrows mark the BrdU only signal from both chromosomes for comparison.

**Figure A1.25. Replication timing of chromosome 6 in the distal deletion lines.** (**A**) An example of a mitotic spread from ∆175F-11c (~25.9 Mb deletion) processed for BrdU incorporation (green) and FISH with chromosome 6 centromere (red) and BAC CTD-2231M1 (red) probes. The deleted chromosome 6 (∆6) lacks the BAC signal. (**B**) The two chromosome 6s from two different mitotic cells were separated from the rest of the spread and the different fluors were overlaid in different combinations. (**C**-**E**) Total BrdU incorporation in the chromosome 6s in ∆175F-11c (~25.9 Mb deletion), ∆175F-11a (1.4 Mb deletion) and ∆175F-7a cells (**C**, **D**, and **E**, respectively). The red (6) and blue (∆6) bars represent the two chromosome 6s from different mitotic cells (Cell Number). (Figure on next page).

**Figure A1.25. Replication timing of chromosome 6 in the distal deletion lines.**

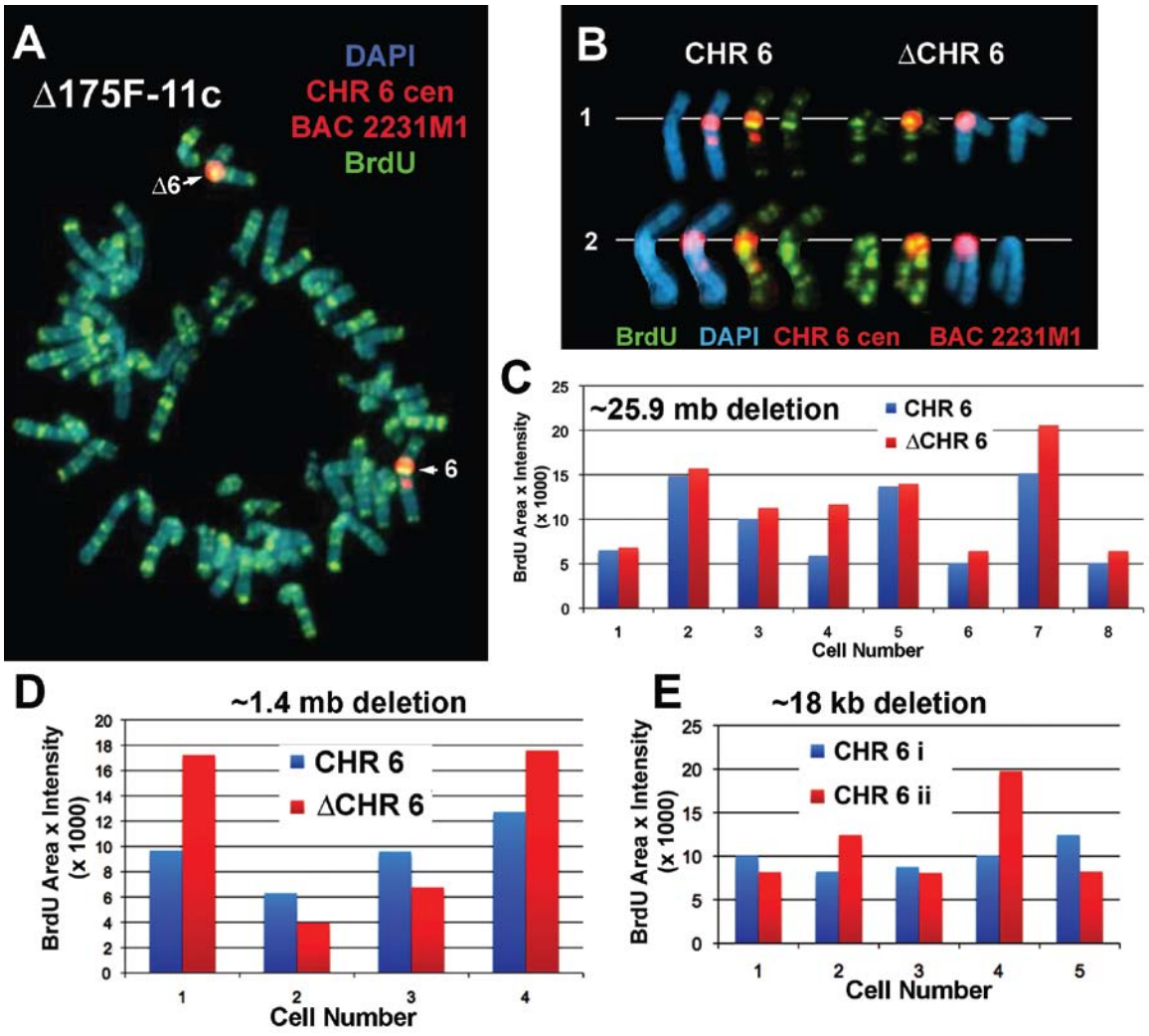


**Figure A1.26. Mono-allelic expression of *FHL5OST*.** (**A**) RT-PCR of two SNPs located within introns of *FHL5*, rs9486715 and rs9398148, and a SNP located telomeric to *FHL5*, rs11153085. PCRs were carried out on genomic DNA from P175 or RNA from P175 treated either with (+RT) or without (-RT) reverse transcriptase. Sequencing traces from the same PCR products are shown in the right panels for each SNP and each PCR. The arrows mark the locations of the heterozygous SNPs. (**B**) Exon-specific RT-PCR for expression of *FHL5*. PCRs were carried out, with primers designed to detect exons 6 and 7 from the spliced *FHL5* cDNA, on RNA from P175 either with (+RT) or without (-RT) reverse transcriptase. PCR on testes cDNA (+RT) yielded the correct size product, which was confirmed by sequencing (not shown). PCR products were not detected from P175 DNA, because the intron spanned by these two primers is >5 kb. Sequencing of the testes PCR product revealed properly spliced *FHL5* cDNA (not shown). (**C**) RT-PCR products are from the *FHL5* opposite strand in P175. Strand-specific reverse transcriptase reactions were carried out with primers designed to generate cDNA from either the *FHL5* or *FHL5* opposite strand on RNA from P175 or testes either with (+RT) or without (-RT) reverse transcriptase. Subsequent PCR with internal primers, designed to detect SNP rs94867115, indicated that the transcripts from P175 RNA were synthesized from the opposite strand from *FHL5*. Both products were generated from testes RNA, which served as control for these reactions. (Figure on next page).

**Figure A1.26. Mono-allelic expression of *FHL5OST*.**

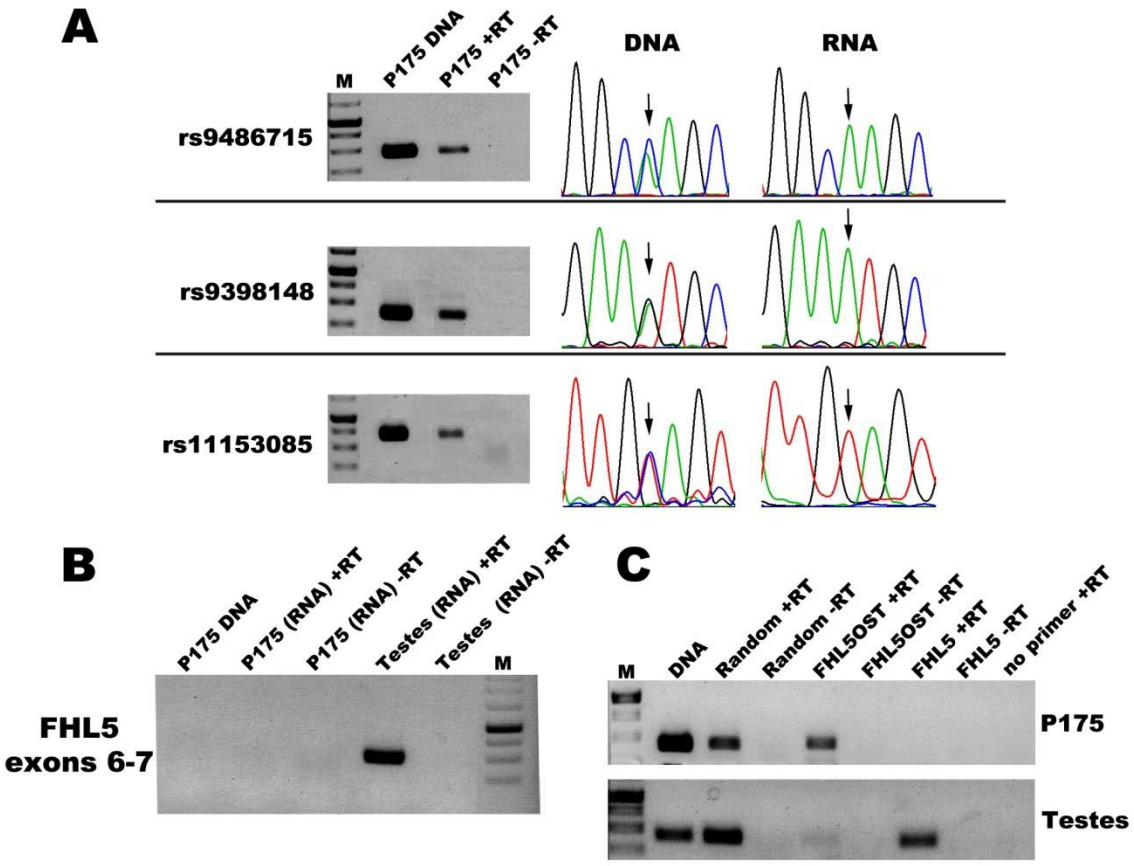


**Figure A1.27. RNA-DNA FISH for *ASAR6* in P175 cells.** (**A**-**R**) P175 cells were subjected to RNA-DNA FISH using the Fosmids G248P83419A4 (3419A4) (**A**-**F**), G248P86031A6 (6031A6) (**G**-**L**), and G248P89065D8 (9065D8) (**M**-**R**) as probes for the RNA (green) and BAC RP11-959I6 as the probe for the DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probes detected a positive signal in >90% of the cells for each probe. DNA was stained with DAPI. Arrows mark the location of RNA signal. Two sites of RNA hybridization (**E** and **N**) were detected in <5% of cells (see Fig. 2.4G). (Figure on next page).

**Figure A1.27. RNA-DNA FISH for *ASAR6* in P175 cells.**

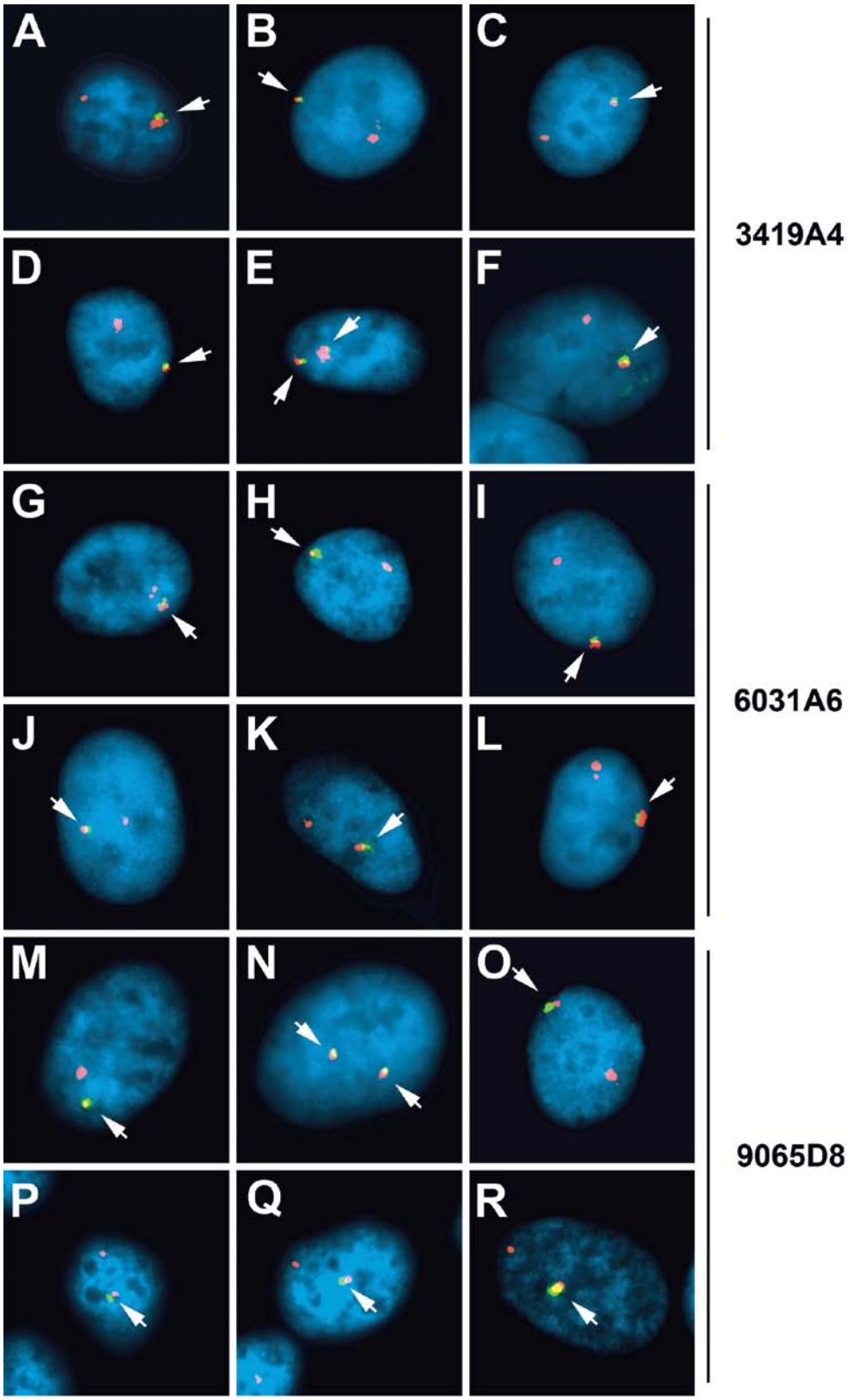


**Figure A1.28. RNA-DNA FISH with *FUT9* cDNA probe in P175 cells.** (**A-R**) P175 cells were subjected to RNA-DNA FISH using a *FUT9* cDNA as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in <10% of cells (see Figs 2.5F and 2.7D). (Figure on next page).

**Figure A1.28. RNA-DNA FISH with *FUT9* cDNA probe in P175 cells.**

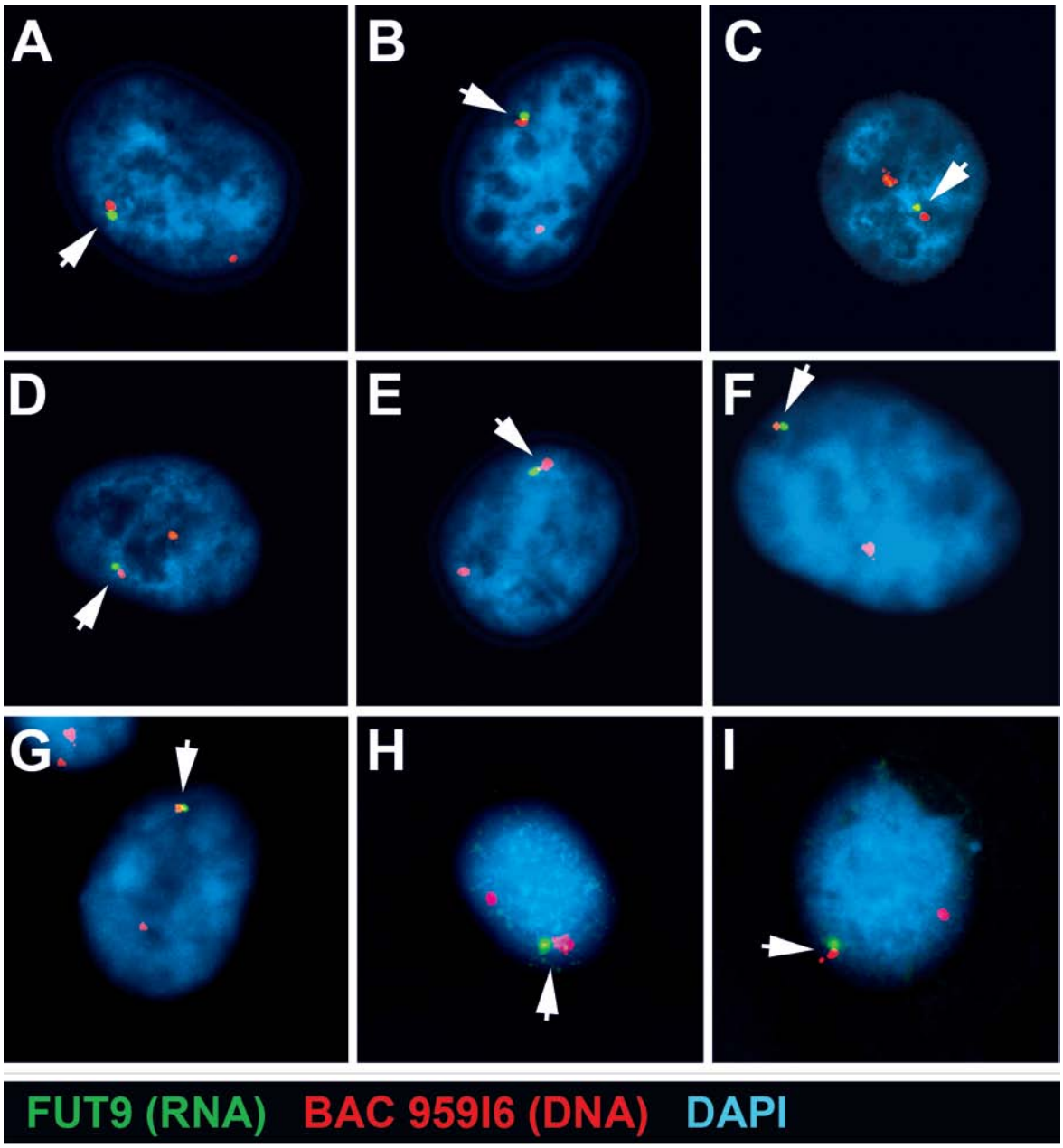


**Figure A1.29. RNA-DNA FISH with *FUT9* cDNA probe in ∆175-1i cells.**

(**A**-**R**) ∆175-1i cells were subjected to RNA-DNA FISH using a *FUT9* cDNA as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in >60% of cells (see Fig. 2.7C). (Figure on next page).

**Figure A1.29. RNA-DNA FISH with *FUT9* cDNA probe in Δ175-1i cells.**

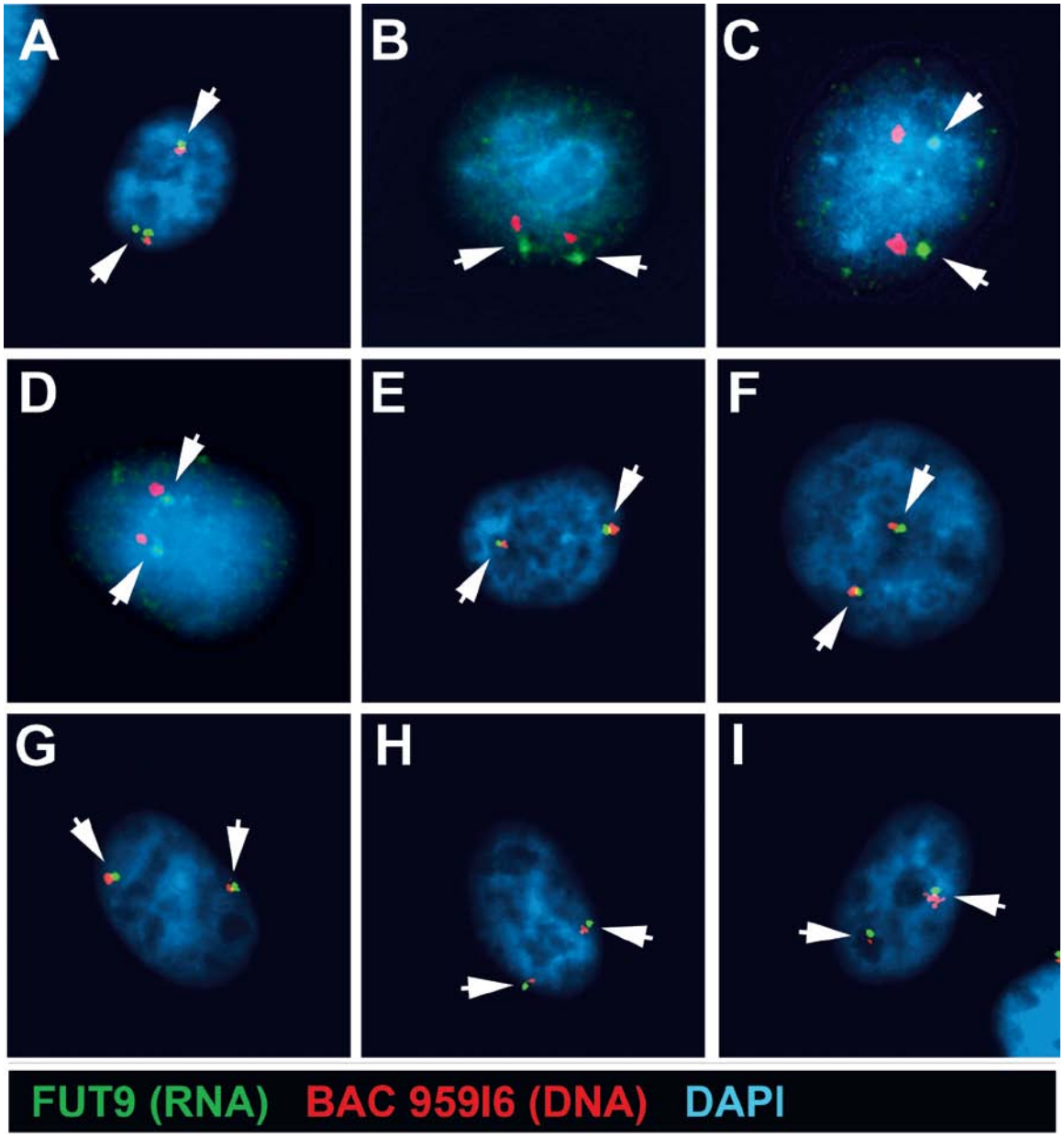


**Figure A1.30. RNA-DNA FISH with *FUT9* intronic probe in P175 cells.**

(**A**-**R**) P175 cells were subjected to RNA-DNA FISH using a 4.5 kb genomic fragment from the first intron of *FUT9* as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in <10% of cells (see Figs 2.4F and 2.7D). (Figure on next page).

**Figure A1.30. RNA-DNA FISH with *FUT9* intronic probe in P175 cells.**

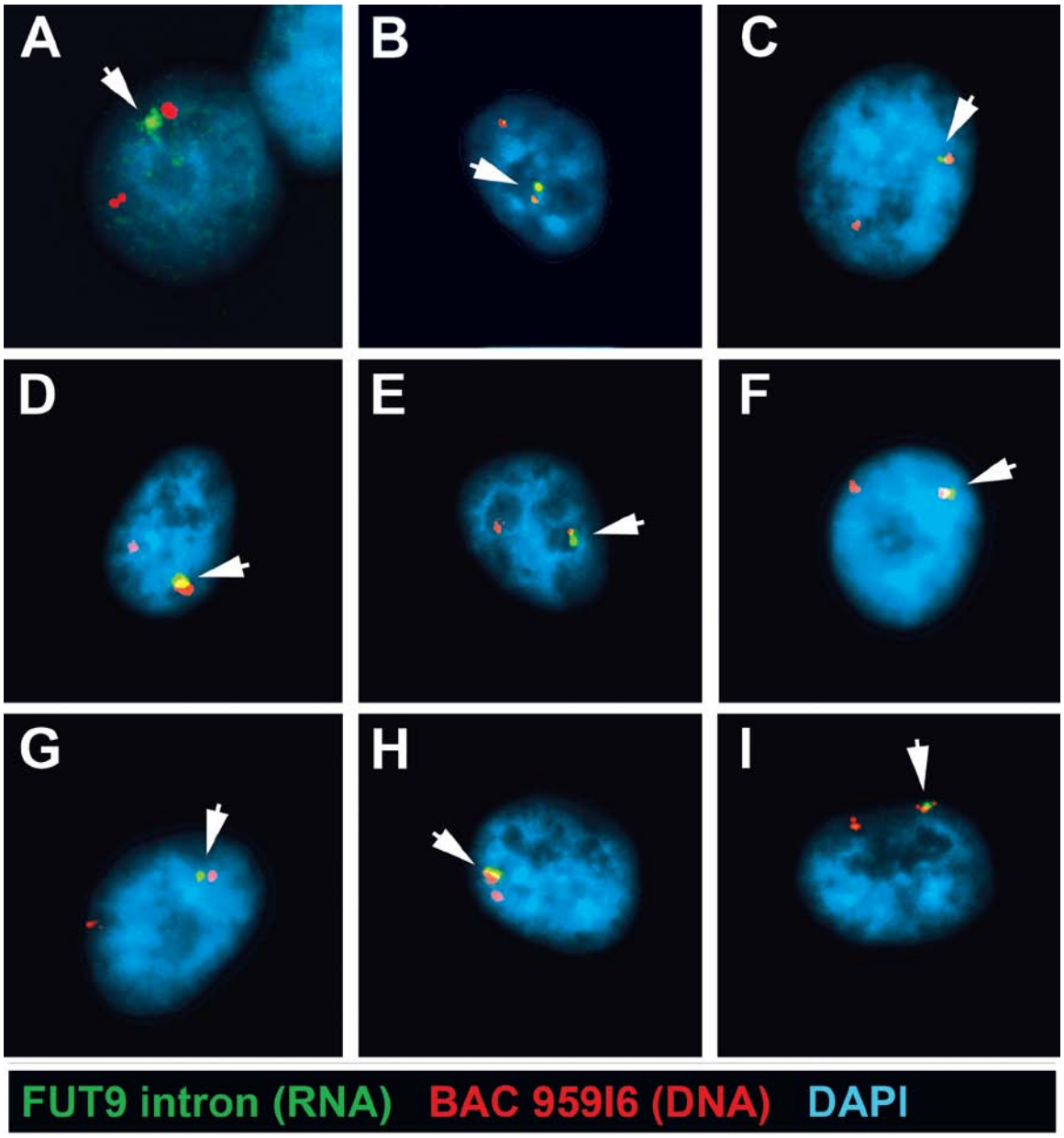


**Figure A1.31. RNA-DNA FISH with *FUT9* intronic probe in ∆175-1i cells.** (**A**-**R**) ∆175-1i cells were subjected to RNA-DNA FISH using a 4.5 kb genomic fragment from the first intron of *FUT9* as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in >55% of cells (see Fig. 2.7D). (Figure on next page).

**Figure A1.31. RNA-DNA FISH with *FUT9* intronic probe in Δ175-1i cells.**

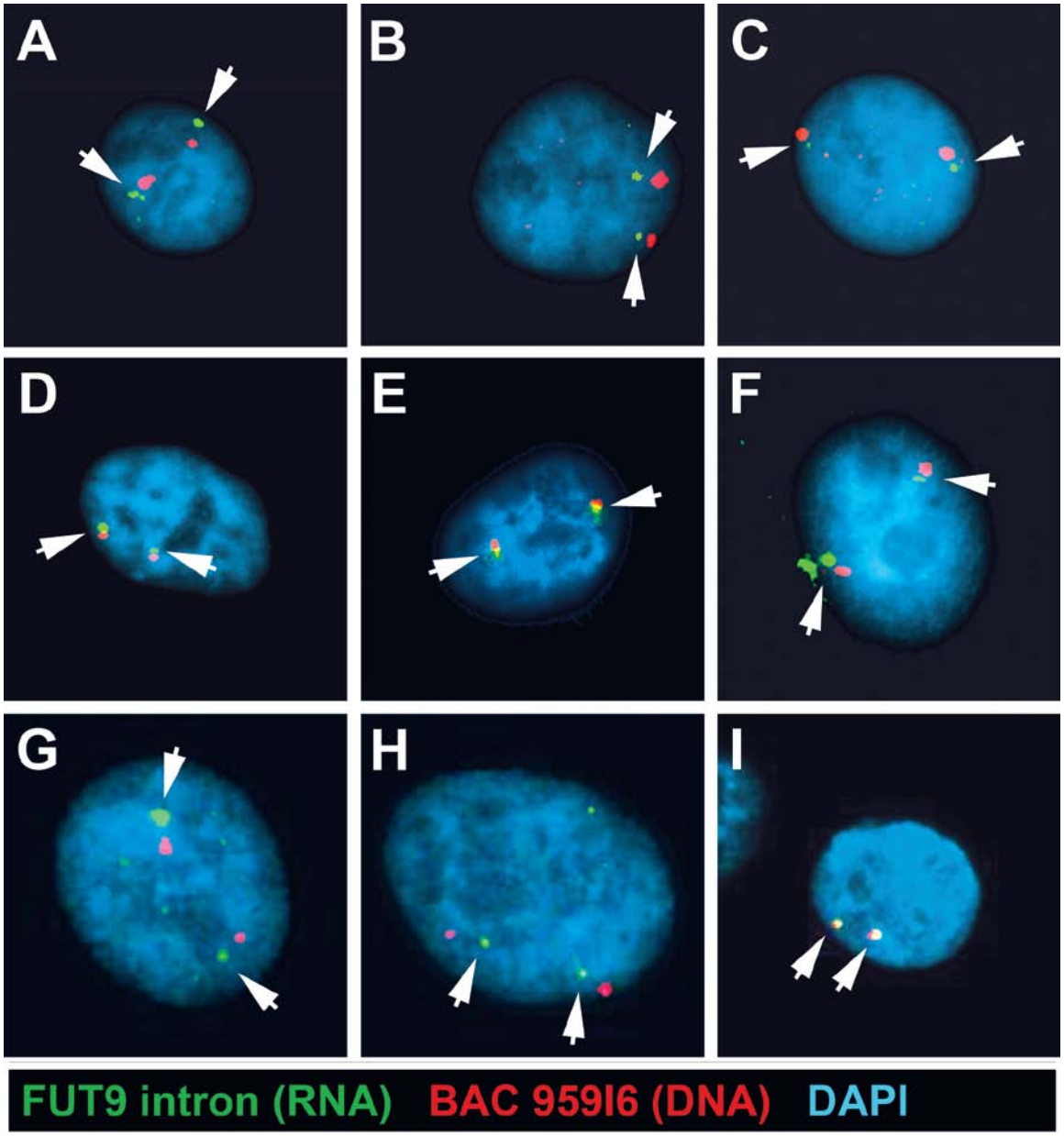


**Figure A1.32. RNA-DNA FISH for *MANEA* in P175 cells.** (**A**-**R**) P175 cells were subjected to RNA-DNA FISH using a *MANEA* cDNA as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signals. Two sites of RNA hybridization were detected in >70% of cells (see Fig. 2.5F). (Figure on next page).

**Figure A1.32. RNA-DNA FISH for *MANEA* in P175 cells.**

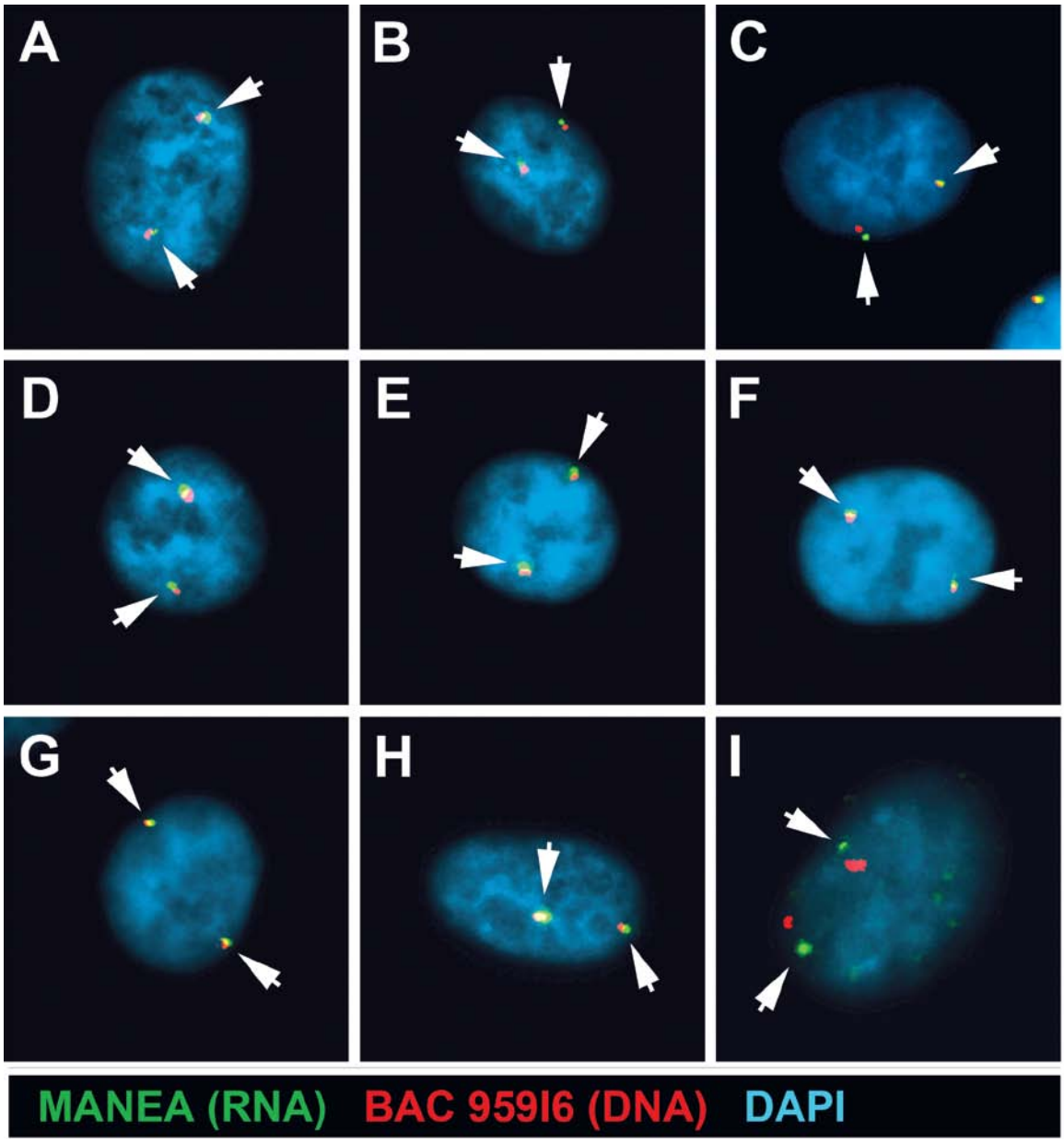


**Figure A1.33. RNA-DNA FISH for *MANEA* in Δ175-1i cells.** (**A**-**R**) Δ175-1i cells were subjected to RNA-DNA FISH using a *MANEA* cDNA as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signals. Two sites of RNA hybridization were detected in >75% of cells (see Fig. 2.7F). (Figure on next page).

**Figure A1.33. RNA-DNA FISH for *MANEA* in Δ175-1i cells.**

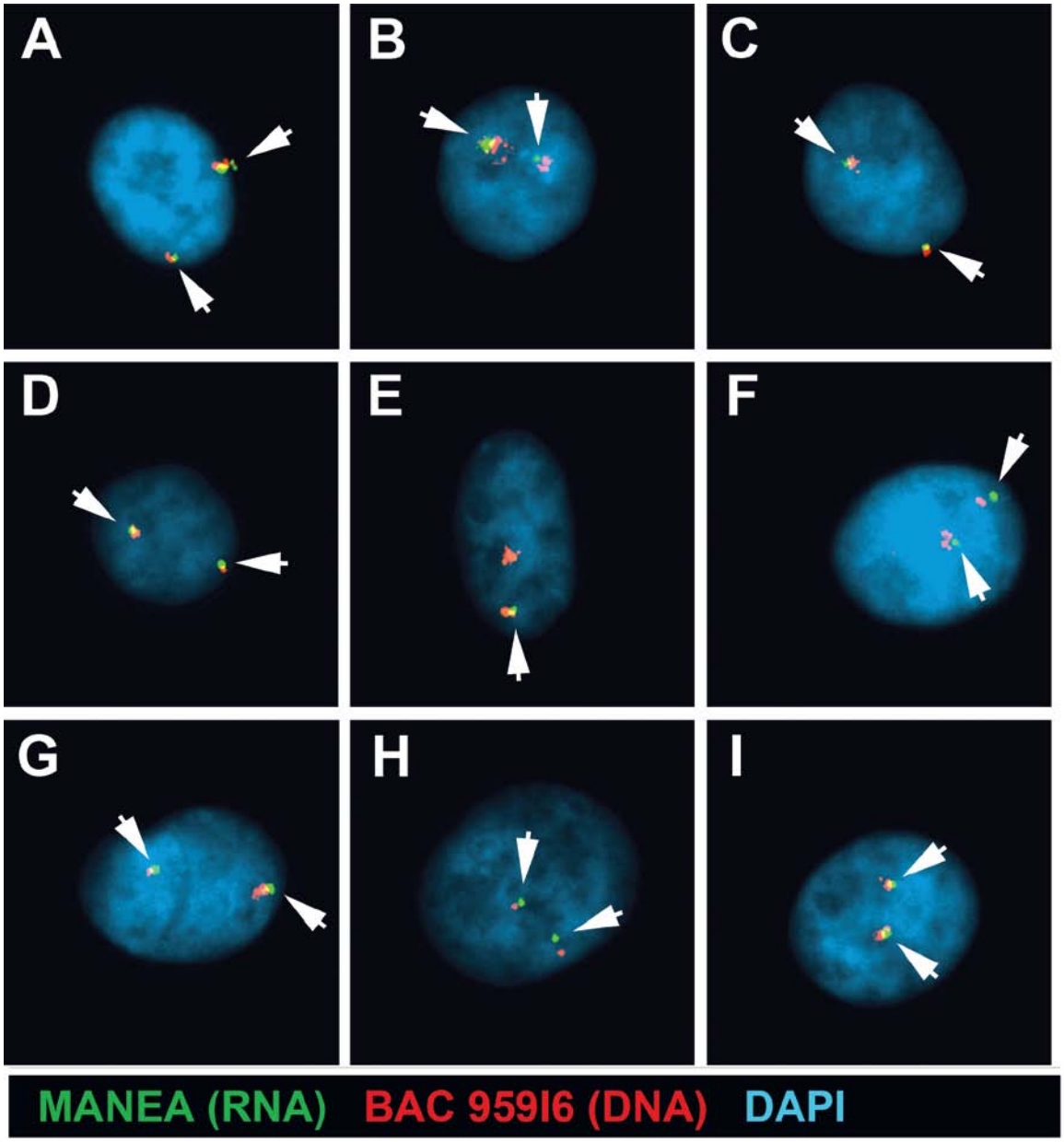


**Figure A1.34. RNA-DNA FISH for *FHL5OST* in P175 cells.** (**A**-**R**) P175 cells were subjected to RNA-DNA FISH using Fosmid G248P86054G4 as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in <15% of cells (see Fig. 2.7E). (Figure on next page).

**Figure A1.34. RNA-DNA FISH for *FHL5OST* in P175 cells.**

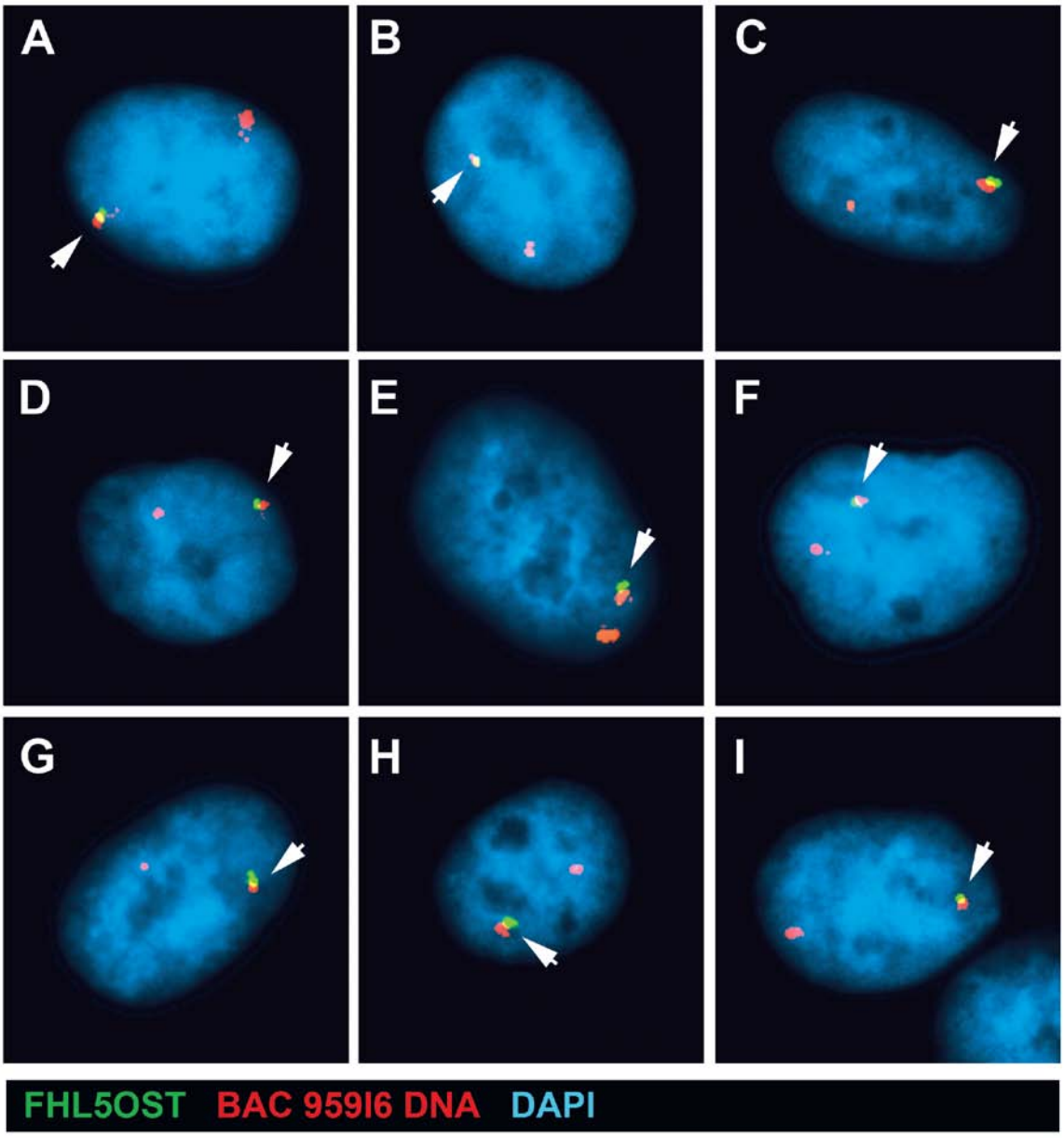


**Figure A1.35. RNA-DNA FISH for *FHL5OST* in Δ175-1i cells.** (**A**-**R**) Δ175-1i cells were subjected to RNA-DNA FISH using Fosmid G248P86054G4 as probe for RNA (green) and BAC RP11-959I6 as the probe for DNA (red). In regions of the slides where the FISH worked well, the RNA FISH probe detected a positive signal in >90% of the cells. DNA was stained with DAPI. Arrows mark the location of the RNA signal. Two sites of RNA hybridization were detected in >50% of cells (see Fig. 2.7E). (Figure on next page).

**Figure A1.35. RNA-DNA FISH for *FHL5OST* in Δ175-1i cells.**

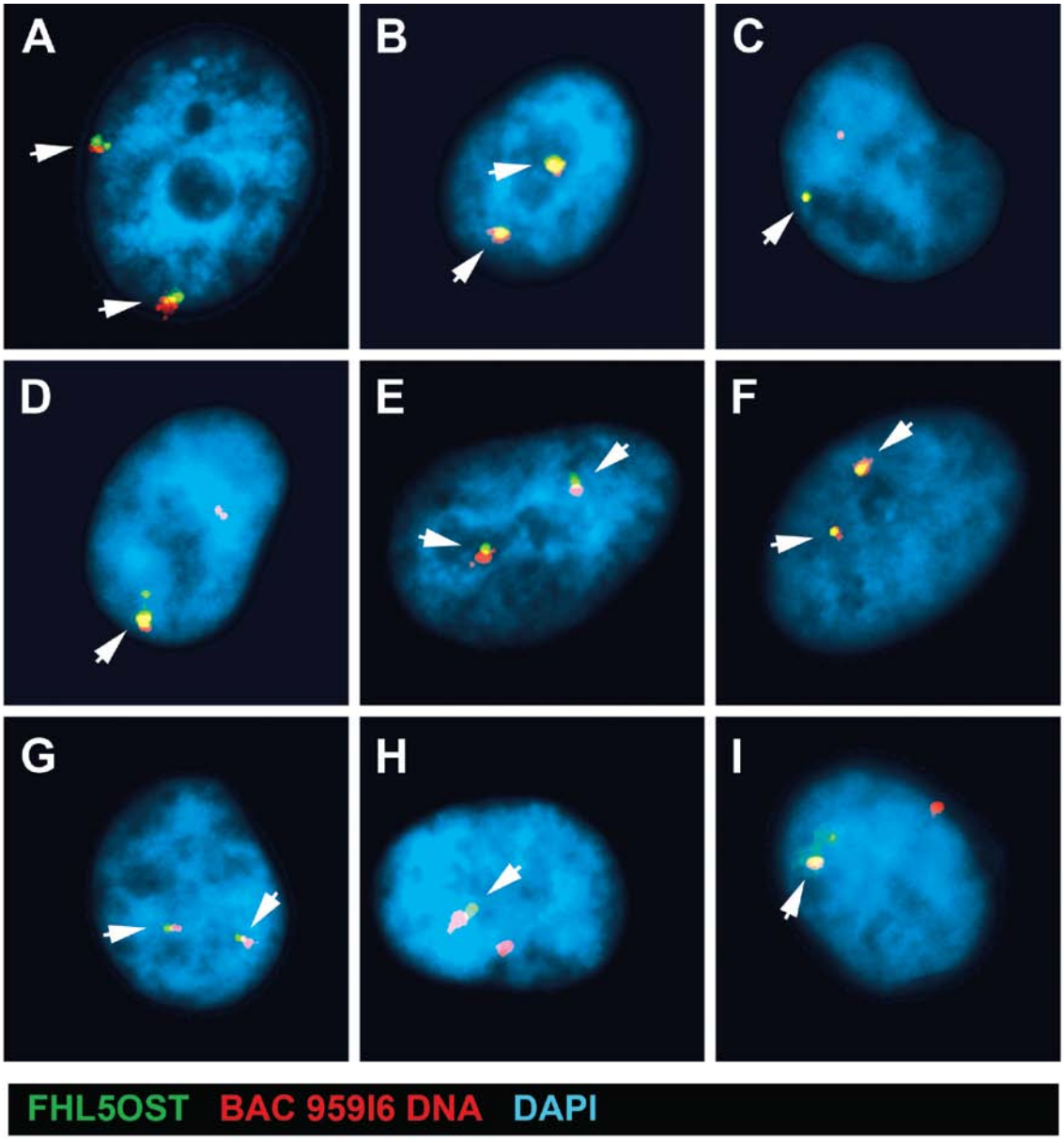


**Figure A1.36. DNA FISH replication assay.** Low passage primary human skin fibroblast cells were pulse-labeled with BrdU for 60 minutes, harvested by trypsinization, fixed in 3:1 methanol:acetic acid, dropped on slides, and processed for DNA FISH using BAC RP11-374I15 as probe (red) (**C**, **F**, and **I**). BrdU was detected using an FITC labeled anti-BrdU antibody (green) (**B**, **E**, and **H**). DNA was stained with DAPI (**A**, **D**, and **G**). Approximately 35% of cells were BrdU positive. Arrows mark the location of the DNA signals (**C**, **F**, and **I**). Single dots (S) represent un-replicated loci and double dots (D) represent replicated loci. The frequency of the SS, DD and SD patterns were determined in 200 cells and are shown as percent (%). (Figure on next page).

**Figure A1.36. DNA FISH replication assay.**

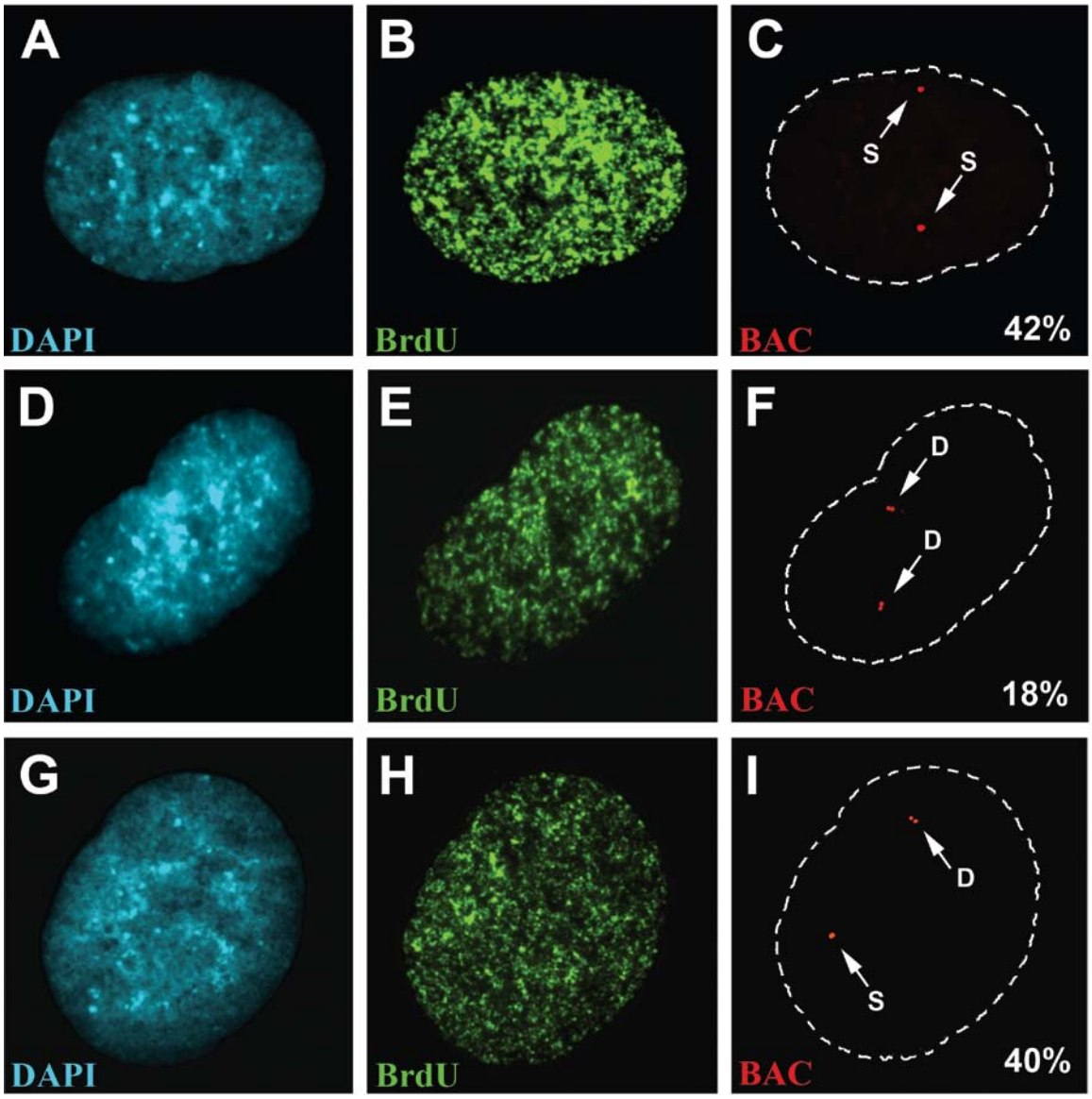
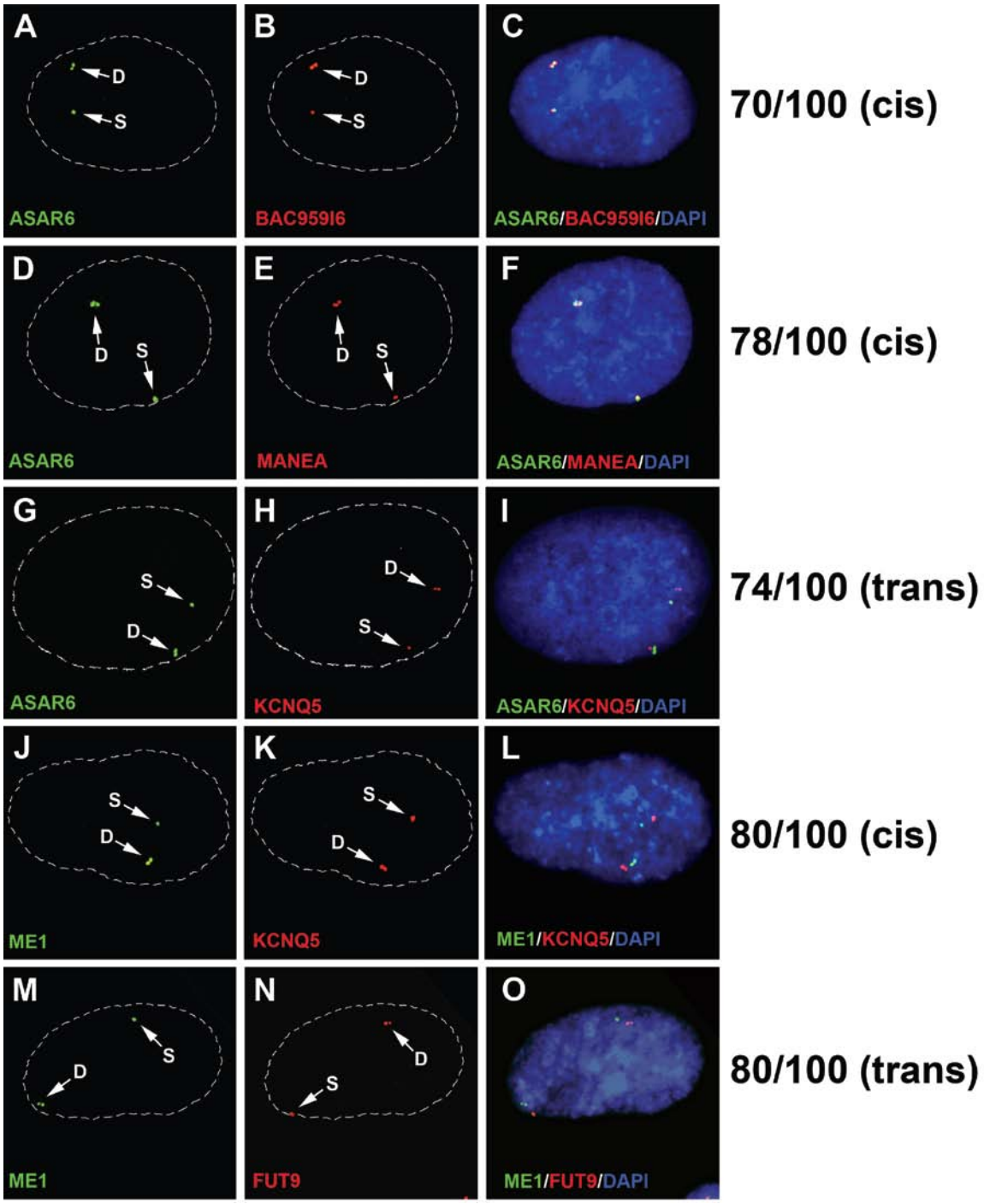


**Figure A1.37. Two-color DNA FISH assay for coordination in asynchronous replication.** Low passage primary human skin fibroblasts were harvested by trypsinization, fixed in 3:1 methanol:acetic acid, dropped on slides, and processed for DNA FISH using BAC RP11-374I15 as probe (*ASAR6*; green; **A**-**I**) plus BAC959I6 (red; **A**-**C**), Fosmid G248P8476B7 (*MANEA*; red; **D**-**F**) or BAC RP11-106C24 (*KCNQ5*; red; **G**-**I**). FISH was also carried out with BAC RP11-703G14 as probe (*ME1*; green; **J**-**O**) plus BAC RP11-106C24 (*KCNQ5*; red; **J**-**L**) or BAC RP11-75N19 (*FUT9*; red; **M**-**O**). DNA was stained with DAPI. Arrows mark the location of the FISH signals. Single dots (S) represent un-replicated loci and double dots (D) represent replicated loci. 100 cells that simultaneously displayed the SD pattern for both probes were scored for the *cis* or *trans* configuration. (Figure on next page).

**Figure A1.37. Two-color DNA FISH assay for coordination in asynchronous replication.**



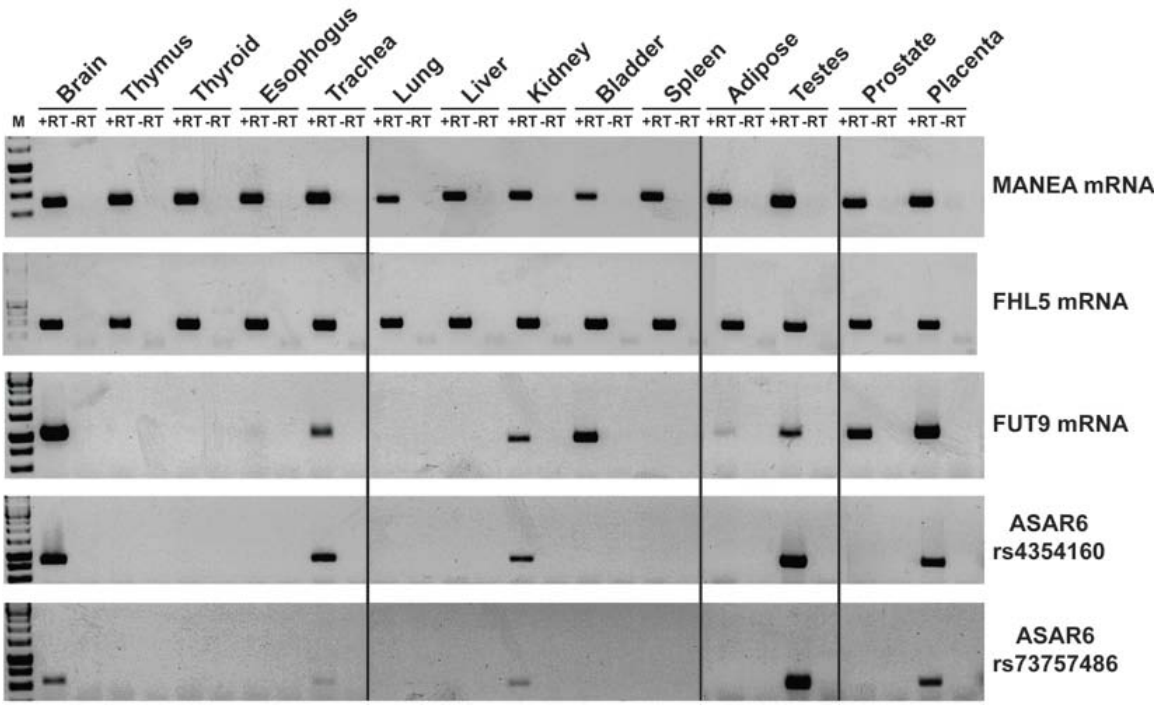



**Figure A1.38. Expression of *MANEA*, *FHL5*, *FUT9* and *ASAR6* in primary human tissues.** Reverse transcriptase reactions, either with (+RT) or without (-RT) reverse transcriptase, on total RNA, were subjected to PCR with primers designed to span *MANEA FHL5*, and *FUT9* exons and primers designed to amplify the rs2224951 and rs4354160 SNP locations within *ASAR6* (see Tables A1.2 and A1.3). PCRs were carried out on first strand cDNA generated from total RNA isolated from the indicated tissues. Due to the large number of samples, multiple gels per primer set were run and the images inverted and combined using Photoshop. The black lines indicate the partition of individual gels.

**Table A1.1. Lentiviral integration sites in P175.** LAM-PCR [107] was used to determine the lentiviral 5' LTR-genome junctions. Integration sites are shown in base pairs along chromosome 6. P175 has *lox*P-3'RT cassette junctions in chromosome 6 at 96,386,321 and 96,386,392 bp.

| Chromosome 6 Lentiviral Integration Sites in P175 | | Deletion (base pairs): | |
|---|---|---|---|
| **Clone** | **Integration Site** | **Proximal** | **Distal** |
| **Δ175-4a** | **65,869,673** | **30,416,648** | |
| **Δ175-6b** | **66,114,310** | **30,272,011** | |
| **Δ175-2c** | **74,519,565** | **21,866,756** | |
| **Δ175-6c** | **75,369,866** | **21,016,455** | |
| **Δ175-1h** | **76,238,275** | **20,148,046** | |
| **Δ175-2d** | **79,806,616** | **16,579,705** | |
| **Δ175-3f** | **92,457,124** | **3,929,197** | |
| **Δ175-4c** | **93,150,816** | **3,235,505** | |
| **Δ175-2g** | **94,268,844** | **2,117,477** | |
| **Δ175-3e** | **96,153,003** | **233,318** | |
| **Δ175-1i** | **96,155,530** | **230,791** | |
| **Δ175-32a** | **96,238,021** | **148,300** | |
| **Δ175-11d** | **96,291,854** | **94,467** | |
| **Δ175-31h** | **96,303,977** | **82,344** | |
| **Δ175-23a** | **96,309,971** | **76,350** | |
| **loxP-3'RT** | **96,386,321-96,386,392** | | |
| **Δ175F-7a** | **96,404,917** | | **18,525** |
| **Δ175F-12a** | **96,539,178** | | **152,786** |
| **Δ175F-8a** | **97,757,638** | | **1,371,246** |
| **Δ175F-11a** | **97,809,711** | | **1,423,319** |
| **Δ175F-7e** | **100,958,279** | | **4,571,887** |
| **Δ175F-11c** | **122,283,187** | | **25,896,795** |

## Table A1.2. Mono-allelic expression of chromosome 6q16.1.

Genotyping and expression analysis on total RNA isolated from HTD114, P175, and two chromosome 6 mono-chromosomal hybrids [L(Hyg)-1 and L(Neo)-38]. PCR primers were designed to amplify both homozygous and heterozygous SNPs present in HTD114 DNA. The SNP genotypes were determined using the Affymetrix 500K SNP array (Figs A1.20 and A1.21) and/or by direct sequencing of PCR products.

| | | | Cell Lines | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | HTD114 | | P175 | | L(Hyg)-1 | | L(Neo)-38 | | |
| DNA Feature | Location (Chr.6)* | Region | Genotype | Expressed | Genotype | Expressed | Genotype | Expressed | Genotype | Expressed | Assay(s)† |
| rs7758062 | 96,139,599 | MANEA | C/C | + | C/C | + | | + | | + | r,h |
| rs9374586 | 96,141,896 | MANEA | T/T | | T/T | + | | + | | + | r,h |
| exons 2-3 | 96,141,286-96,151,394 | MANEA | | | | + | | + | | + | r,h |
| exons 2-3 | 96,141,041-96,151,345 | MANEA | | | | + | | + | | + | r,h |
| rs9489578 | 96,171,771 | ASAR6 | G/G | + | G/G | + | | + | | - | r,h |
| rs6934550 | 96,204,543 | ASAR6 | A/G | + | A/G | + (A) | A | + (A) | G | – | r,s,h,as |
| rs9375155 | 96,207,496 | ASAR6 | A/G | + | A/G | + (G) | G | + (G) | A | – | r,h,s |
| rs4629688 | 96,213,935 | ASAR6 | C/T | + | C/T | + (T) | | + | | – | r,h,s |
| rs4354160 | 96,319,715 | ASAR6 | G/G | + | G/G | + | | + | | – | r,h |
| rs4645429 | 96,320,436 | ASAR6 | A/G | + | A/G | + (G) | G | + (G) | A | – | r,s,h,as |
| rs73757486 | 96,321,629 | ASAR6 | C/C | + | C/C | + | | + | | – | r,h |
| rs9389160 | 96,325,215 | ASAR6 | A/A | – | A/A | + | | + | | – | r,h |
| rs9321478 | 96,330,812 | ASAR6 | A/A | + | A/A | + | | + | | – | r,h |
| rs28789131 | 96,333,352 | ASAR6 | A/A | + | A/A | + | | + | | – | r,h |
| rs9399157 | 96,338,401 | ASAR6 | A/A | – | A/A | + | | + | | – | r,h |
| rs62417886 | 96,340,518 | ASAR6 | A/A | – | A/A | + | | + | | – | r,h |
| rs13206934 | 96,341,725 | ASAR6 | C/C | + | C/C | + | | + | | – | r,h |
| rs57663635 | 96,342,884 | ASAR6 | G/G | + | G/G | + | | + | | – | r,h |
| rs58828340 | 96,344,427 | ASAR6 | | + | | + | | + | | – | r,h |
| rs9494487 | 96,343,838 | ASAR6 | G/G | + | G/G | + | | + | | – | r,h |
| rs62417889 | 96,346,038 | ASAR6 | | + | | + | | + | | – | r,h |
| rs73549956 | 96,348,040 | ASAR6 | | – | | + | | + | | – | r,h |
| rs13205985 | 96,352,261 | ASAR6 | T/T | + | T/T | + | | + | | – | r,h |
| rs9389466 | 96,352,465 | ASAR6 | A/A | + | A/A | + | | + | | – | r,h |
| rs4330541 | 96,353,864 | ASAR6 | T/T | + | T/T | + | | + | | – | r,h |
| rs10872471 | 96,355,124 | ASAR6 | | + | | + | | + | | – | r,h |
| rs11968441 | 96,355,242 | ASAR6 | A/A | + | A/A | + | | + | | – | r,h |
| rs4489163 | 96,357,056 | ASAR6 | | + | | + | | + | | – | r,h |
| rs13203882 | 96,358,585 | ASAR6 | A/A | + | A/A | + | | + | | – | r,h |
| rs34875992 | 96,361,889 | ASAR6 | G/G | – | G/G | + | | + | | – | r,h |
| rs73543073 | 96,365,807 | ASAR6 | | + | | + | | + | | – | r,h |
| rs4499938 | 96,369,610 | ASAR6 | T/T | + | T/T | + | | + | | – | r,h |
| rs4840081 | 96,371,876 | ASAR6 | T/T | – | T/T | + | | + | | – | r,h |
| rs34592682 | 96,372,712 | ASAR6 | | + | | + | | + | | – | r,h |
| rs4516948 | 96,373,551 | ASAR6 | G/G | + | G/G | + | | + | | – | r,h |
| rs10457692 | 96,376,214 | ASAR6 | C/C | – | C/C | + | | + | | – | r,h |
| rs12204006 | 96,386,171 | ASAR6 | | – | | + | | + | | – | r,h |
| rs9390771 | 96,572,701 | FUT9 | A/A | | A/A | + | | – | | + | r,h |
| rs9390772 | 96,572,822 | FUT9 | A/A | | A/A | + | | – | | + | r,h |
| rs60053258 | 96,572,844 | FUT9 | T/T | | T/T | + | | – | | + | r,h |

Table A1.2 continued on next page…

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs1325079 | 96,572,999 | FUT9 | C/C | | C/C | + | | – | | + | r,h |
| rs6925780 | 96,582,526 | FUT9 | A/A | | A/A | + | | – | | + | r,h |
| rs6925964 | 96,582,615 | FUT9 | A/A | | A/A | + | | – | | + | r,h |
| rs9498737 | 96,587,260 | FUT9 | A/G | | A/G | + (G) | | – | | + | r,h,s |
| rs9386298 | 96,594,290 | FUT9 | A/T | | A/T | + (T) | | – | | + | r,h,s |
| rs2224951 | 96,599,864 | FUT9 | A/G | | A/G | + (A) | | – | | + | r,h,s |
| rs2248645 | 96,599,931 | FUT9 | A/G | | A/G | + (A) | | – | | + | r,h,s |
| rs2248641 | 96,599,954 | FUT9 | A/G | | A/G | + (A) | | – | | + | r,h,s |
| rs2248548 | 96,600,076 | FUT9 | C/T | | C/T | + (T) | | – | | + | r,h,s |
| rs4304170 | 96,618,578 | FUT9 | C/T | | C/T | + (C) | | – | | + | r,h,s,as |
| rs2387151 | 96,618,639 | FUT9 | G/T | | G/T | + (G) | | – | | + | r,h,s,as |
| rs4002794 | 96,618,666 | FUT9 | A/G | | A/G | + (G) | | – | | + | r,h,s,as |
| rs4598084 | 96,636,820 | FUT9 | C/T | | C/T | + | | – | | + | r,h |
| rs7754532 | 96,640,396 | FUT9 | C/T | | C/T | + (T) | | – | | + | r,h,s |
| rs12207766 | 96,653,696 | FUT9 | C/T | | C/T | + | | – | | + | r,h |
| rs9499225 | 96,669,708 | FUT9 | A/G | | A/G | + | | – | | + | r,h |
| rs4515397 | 96,711,946 | FUT9 | C/T | + | C/T | + (C) | T | – | C | + (C) | r,h,s |
| rs4380761 | 96,712,354 | FUT9 | A/G | + | A/G | + (A) | G | – | A | + (A) | r,h,s |
| rs9373720 | 96,763,370 | FUT9 | | | | + | | – | | + | r,h |
| rs17726532 | 96,764,292 | FUT9 | A/G | + | A/G | + (G) | A | – | G | + (G) | r,h,s |
| rs9399845 | 96,766,398 | FUT9 | | | | + | | – | | + | r,h |
| rs7749759 | 96,768,632 | FUT9 | | | | + | | – | | + | r,h |
| exons 2-3 | 96,668,695-96,758,742 | FUT9 | | | | + | | – | | + | r,h |
| | | | | | | | | | | | |
| rs9373918 | 97,096,060 | KIAA0076 | A/G | + | A/G | + (A/G) | G | + | A | + | r,h,s |
| | | | | | | | | | | | |
| rs2273621 | 97,165,274 | FHL5OST‡ | A/G | | A/G | + (A) | G | | A | | r,s |
| rs11153082 | 97,166,387 | FHL5OST | A/G | | A/G | + (A) | G | | A | | r,s |
| rs9486715 | 97,166,490 | FHL5OST | A/C | | A/C | + (A) | C | | A | | r,s |
| rs6902835 | 97,169,726 | FHL5OST | C/T | | C/T | + (C) | T | | C | | r,s |
| rs9373985 | 97,170,243 | FHL5OST | G/C | | G/C | + (C) | | | | | r,s |
| rs9398148 | 97,170,276 | FHL5OST | A/G | | A/G | + (A) | | | | | r,s |
| rs9398152 | 97,170,445 | FHL5OST | G/T | | G/T | + (T) | | | | | r,s |
| rs11153085 | 97,173,076 | FHL5OST | | | | + | | | | | r |
| rs4839692 | 97,225,981 | FHL5OST | | | | + | | | | | r |
| exons 2-5 | 97,117,838-97,160,585 | FHL5 | | | | – | | | | | r |
| exons 6-7 | 97,165,315-97,170,606 | FHL5 | | | | – | | | | | r |

*UCSC Genome Browser, Mar. 2006 Assembly (hg18)

blank spaces = not determined
() = Sequence from RT-PCR product

† r = RT-PCR
h = RT-PCR in hybrid cell lines
s = RT-PCR and sequencing
as = allele-specific RT-PCR

‡FHL5OST is also known as ENSG00000246665 from Ensembl.

## Table A1.3. Primers used in various PCR assays.

**PCR Primers**

| DNA Feature | Gene Region | Forward | Reverse |
|---|---|---|---|
| rs7758062 | MANEA | CCTTTTTGCCCTTTGAAGTATG | TGACAGTGCAAGAAAACGTCTC |
| rs9374586 | MANEA | TGAAAAAGCCCAAACACCTATC | GACATTGGCTCCAGCTTTTATC |
| mRNA exon 2-3 | MANEA | AAACCTTCCAAAGCCTCTGAAC | TCAGTAGGTTCTCCATTTTCATC |
| mRNA exon 2-3 | MANEA | CAAAGTTTCGGAGAAGGACTTG | GGGCTAGTACACCAATTGAAGC |
| rs9489578 | ASAR6 | TTGTACATGTGTTCCATGATTTG | AAAAGAATTGGTTCTGCTCTCC |
| rs6912187 | ASAR6 | CCGACAGAACCCAAACTCAT | GGTTCCCTTATCAGCCTCCTAC |
| rs6934550 | ASAR6 | CCTACCTTGCTTTGAGTTGTCC | GGTTCCCTTATCAGCCTCCTAC |
| rs9375155 | ASAR6 | CCAGAGACCCAGAGATATCGAC | AACCATCAAACTTAGGGTCACATAG |
| rs4629688 | ASAR6 | TCAGAGAAGCTTGCACAGAAAG | TGGGTTAGCATGTAGTCAGACC |
| rs9482760 | ASAR6 | TTTCCAGATGTGATGACTGTCC | TCTATATCCACAGGGTGCATTG |
| rs4568459 | ASAR6 | GCCACACATTTTACAGTTTTGC | TCCTCTTGTATTTCAGGCATTG |
| rs4645429 | ASAR6 | TGTGGTATTCACGTAGCCAGTC | CTGGTCCTCATCCAGAGCTTAC |
| rs73757486 | ASAR6 | ACGCAGCAGATCGAAGTATATG | AAATCCCTAATGGGGAATCAAC |
| rs9389160 | ASAR6 | TAAGTTCAAACCCTCCAGCAAC | GAGTGGACTTAATTGGCTGCTC |
| rs9321478 | ASAR6 | AGCAAGCAAGCCAGAAGATTAC | GGGAAAGGATTCCCTATTCAAC |
| rs28786474 | ASAR6 | CAGCCATGAATGAGAGTTTCTG | ATTGGAGAAAATCCTTGCACTC |
| rs9399157 | ASAR6 | TCACCAAGGAATAACACCAATG | AATCCAGCAGCACTAAAACCTC |
| rs62417886 | ASAR6 | GTACTCCCTTCACAATCCCTTG | CCAAATGCCTATGTTGAAGTCC |
| rs13206934 | ASAR6 | CCTGTAAGTTGCTGTTGGTCAG | ACAAATAATTCTCGCACAAACG |
| rs57663635 | ASAR6 | CAGGAGGTATTCGTTCTCCAAG | GAGGGGAAGATGGAAATTGAG |
| rs58828340 | ASAR6 | TGTCCTCTGGGATATTTCTTGG | TCCTCTGTCTGTGTAGAGGGAAG |
| rs9494487 | ASAR6 | TACTGATGCCACTCATCCAAAC | TCCAAGAAATATCCCAGAGGAC |
| rs62417889 | ASAR6 | TGCAACATGACCAAGTGGATAG | TATGTGTGTGTGTGTGTGTTGG |
| rs73549956 | ASAR6 | GAAACACCAATTAGCCAGTTCC | ATTTTGCATCCCCACTTAGATG |
| rs13205985<br>rs9389466 | ASAR6 | TTTGCATTTTGAAGGATTGTTG | TTCTCCCACACCATTTCTTTTC |
| rs4330541 | ASAR6 | GACCAGACCAAAGAAGAAGAGAAC | GAAAAATACTGCGTGATCTAGCC |
| rs10872471 | ASAR6 | ATTCGCAATAACAGTGCACAAG | GGAACAAAGTCAGAGGCATCAC |
| rs11968441 | ASAR6 | GAGGATTTCCCCTGTATTTTCC | GTGCAACCCCTACCATAATTTC |
| rs4489163 | ASAR6 | GGAGTCTTTGAAATTTTCTGTGG | GAACAGACCAACAACCTGTACG |
| rs13203882 | ASAR6 | TTGAAAGGTTTCCCAACTGTG | TCTCCTTTTAACCCTTGCTTTG |
| rs34875992 | ASAR6 | TCACAATGATCTTGAGCTCTGC | GTTCAGCCATCCTGGAAAATAG |
| rs73543073 | ASAR6 | CCTATGTGGAGTCTTCCAATCC | ACCGAATGCCTTTTCTCCTAAG |
| rs4499938 | ASAR6 | TGTTTCAAGCACTGTTGTGATG | TCCTGCCACTGTTATTTAGCTTC |
| rs4840081 | ASAR6 | ACCAGCATCAACTAATGCAAAC | AAAAGTGTCCCTCCCTTATTGG |
| rs34592682 | ASAR6 | CATCCTCCCAAGACTAAACCAG | TCAGTTTGCCAGTATTTTGCTG |
| rs4516948 | ASAR6 | CTATCTTTGCTTGCAGATGACG | GTCTTCTTTTGAGGACGGTTTG |
| rs10457692 | ASAR6 | GCGAAGCCTTTTTGTTTCAG | TCCATTCCGTGGTTCTGTATC |
| rs12204006 | ASAR6 | GATTGCCTCCAAAGTCTCTGAC | TTGCTTATGAGTCCTGTGGATG |
| rs9390771<br>rs9390772<br>rs60053258<br>rs1325079 | FUT9 | GTGGATAGCCGCATTAAAAGTC | AATCCTGGCAGATCTTCTTGAC |
| rs6925964 | FUT9 | TTCCTTCTGTGCAGTGAACTTG | AGCCAGTCATTTCTTGACCTTC |
| rs9498737 | FUT9 | TTTAATGGCTGCTAGGAAGCTC | AGAACAGGGAGTTCTGAAAAGG |
| rs9386298 | FUT9 | AAGCATTACCTCCTCCTCACAG | CCAAGAGACATGGGCTTCTAAC |
| rs2224951<br>rs2248645<br>rs2248641<br>rs2248548 | FUT9 | TGCTGGATTAAAATGACTGGTG | TGCTATCTTGGATGACAACTGC |

Table A1.3 continued on next page…

| | | | |
|---|---|---|---|
| **rs4304170<br>rs2387151<br>rs4002794** | **FUT9** | TCCTGAAGATGCCACCTATACC | CTATGCCCCTCATTCAGTTCTC |
| **rs4598084** | **FUT9** | CCTTCTTCCAGTTCAGGTCTTG | CAAATAGCCAATGGCAGAAATAG |
| **rs7754532** | **FUT9** | CCAAGACCGCACCTTCTAATAC | GTCAGTTGGAAAGCTGGGTATC |
| **rs12207766** | **FUT9** | TTCAAATGTGCTGTTGAAATCC | TCTGTTCTCCCAAATTGTTGTG |
| **rs9499225** | **FUT9** | TCGCTTCTCAAGCCTTACTCTC | TGTAGGACAATTTGGGTTTGTG |
| **rs4515397** | **FUT9** | CAATTCTGCCCATGACACCTCAGA | CCATGGTTGCCTTCAAATGTCA |
| **rs4380761** | **FUT9** | GAGTGAACAATACTCCTGGATTTGGGA | ACAAATGTATTCCTTTCATTTGCATCA |
| **rs9373720** | **FUT9** | GGATTTGGATCACGTTAACATTC | CCTAGGGTAGCACTCACATTCC |
| **rs17726532** | **FUT9** | AGGGAGTCATTGGAAATACCTG | CTAAACAGGTGGGATGAAAACC |
| **rs9399845** | **FUT9** | TGTTAGCCAGCCTGTGATTAAG | CTGCCTTTCTATTTCCTGAACC |
| **rs7749759** | **FUT9** | TATGCATGATTCACGTCCTTTC | TTAGTGTTGTGCAGACCTGACC |
| **FUT9 mRNA exon 2-3** | **FUT9** | ACGTGCTTCCCATGATATGTTC | CCAAAATCGTGGAAGATTTACAG |
| **rs9373918** | **KIAA0076** | TACATACCCAAAGCACCACTTG | TGATTGACATGATTCCACACAC |
| **rs2273621** | **FHL5OST** | CACCAGTTGTGGCATAGAAGAC | TGCACCCACTTGTGTTTATCTC |
| **rs11153082<br>rs9486715** | **FHL5OST** | ATCCCTTTCCTCCTTCTTCATC | AGGGAGGTGATCTAGCTGAGTG |
| **rs6902835** | **FHL5OST** | AGACAGAGATCAGACCCTCACC | TGCCATAATCAACAGAAGAAGG |
| **rs9398148<br>rs9373985<br>rs9398152** | **FHL5OST** | GGTGCCAAGTTTATCTGCTTTC | CAGAGTGCAGTATTTTGCAAGG |
| **rs11153085** | **FHL5OST** | CCTCCTTAGGTCTTTGCGTAAC | TTGCATACTCATCTTTGCCAAC |
| **rs4839692** | **FHL5OST** | CATGGATTAAAAAGTAAGGGATGG | AGAATGTGTGTTTCGCTGTTTG |
| **FHL5 mRNA exon 2-5** | **FHL5** | GGTGACTTTCTGTTCCTTTTGG | GCTTTGTCCCTATAGGTTGTCG |
| **FHL5 mRNA exon 6-7** | **FHL5** | TTATGCCAACAAGTGTGTAGCC | CAGAGTGCAGTATTTTGCAAGG |

**Allele-Specific Primer Sets**

| SNP | Gene Region | Common Primer | Allele 1 Primer | Allele 2 Primer |
|---|---|---|---|---|
| **rs4645429** | **ASAR6** | CTGGTCCTCATCCAGAGCTTAC | AAAGGCTAGTTTAGTAATTCAGACA | AAAGGCTAGTTTAGTAATTCAGACG |
| **rs6934550** | **ASAR6** | CCGACAGAACCCAAACTCAT | TCAATCAACTGGGAACATTTATTT | TCAATCAACTGGGAACATTTATTC |
| **rs2224951** | **FUT9** | TGGCACATGGAACAACATAAA | GAGCTAAAGAACCAGCAACAATCTA | GAGCTAAAGAACCAGCAACAATCTG |
| **rs4002794** | **FUT9** | GCAGATGCACCAAAGTATCTCA | GTTCTGGATTTCCTTACACTAACCG | GTTCTGGATTTCCTTACACTAACCC |

**FHL5/FHL5OST Strand-specific RT-PCR Primers**

| | |
|---|---|
| **Forward RT Primer** | CTCTGCGTTGTTTTGCTCATAG |
| **Reverse RT Primer** | ATGAGGGCAAATTGTAGCTTTG |
| **Forward PCR** | GGTGCCAAGTTTATCTGCTTTC |
| **Reverse PCR** | CAGAGTGCAGTATTTTGCAAGG |

**Inverse PCR Primers**

| | Forward Primer | Reverse Primer |
|---|---|---|
| **First Round PCR** | ACTTTGTACACGTCCCACACAC | CCAAGGGTACCAATAGGAAGTG |
| **Nested PCR** | TGTGCTACTGGAGATCAATAAACC | TCCTAGGGTGTTTCTCAAGCTC |

**RT Junction Primers**

| | Vector Primer | Genome Primer |
|---|---|---|
| **Proximal Junction** | GTATTACCGCCTTTGAGTGAGC | GCCAGGAATTAAGTCAAGAAGG |
| **Distal Junction** | AGGAGTGGCCTTTTTCTTTTTC | TGTGCTACTGGAGATCAATAAACC |

Table A.3 continued on next page…

## Lentiviral Integration Junction Primers

| Clone | Genome Junctions | Lentivirus Junction |
|---|---|---|
| Δ175-23a | TTTCCCAACTGACATGATTTTG | TCTTCGTTGGGAGTGAATTAGC |
| Δ175-31h | GCTCTAATTGAACATATAATCGATGG | |
| Δ175-11d | GATGATACGTGAGGAAACAGAGG | |
| Δ175-32a | GTGAGCCTTCTGAAAAGCATTC | |
| Δ175-1i | ATTGTGTGCTGACAAATTGGAG | |
| Δ175-3e | TGCAGAAAGAGGTTTAAAGAATCTG | |
| Δ175-2g | GTAAGAGCCCCAAGCTGAAAC | |
| Δ175-4c | ATTTAAATGACACCAAAATCAGG | |
| Δ175-3f | TCTTCCACCTTTCACCAAAATC | |
| Δ175-2d | TATGGCGTTCCATGTACTTTTG | |
| Δ175-1h | CGGATAATTTTGCATATCCTGTC | |
| Δ175-6c | TGAATACTGTTGTGGGAACAGC | |
| Δ175-2c | TATTGGGGAAAGAAGAGCTCAG | |
| Δ175-6b | CAGTATTGCCAGCAATGGATAG | |
| Δ175-4a | TGGCATATCATGACAGCTTTTC | |
| Δ175F-7a | CGTGCAAACTGGAAACACTTC | |
| Δ175F-12a | TCCAAATTCATTCAGTCTGTGG | |
| Δ175F-8a | AGAGACTTGCTTCAAGGAGTGG | |
| Δ175F-11a | AATGAGCAAAAGTTGGTTCCTC | |
| Δ175F-7e | ACCCTTTCACCATCTGTTCTTC | |
| Δ175F-11c | AAGTTTTGAAGTGGCTTTTTGG | |

**Predicted mono-allelic gene PCR Primers**

| SNP | Gene Region | Forward Primer | Reverse Primer |
|---|---|---|---|
| **rs6458511** | **CYP39A1** | GCCAGGATTTCATAAATGGTTC | ATCTAGCCCACAGAAAATGGTC |
| **rs6913639** | **HCRTR2** | AAAAGCGTATTTCTGGAATTGG | AAAAAGGTCTGAAAATGCAAGC |
| **rs3122155** | **HCRTR2** | CTGAGACTCAAGGCATGATCC | AGTAAAAATACGGGGTGAGAGC |
| **rs9475663** | **COL21A1** | CCCTCTTCCCATTGAGTAAGTC | TTATCGGTTGCATTGATTTTTG |
| **rs6901838** | **COL21A1** | GGAAGTCTACAAGCACACTTGG | GTTCATGTAACACTGCCCTTGC |
| **rs239554** | **TTK** | TCTTGAGGTATTGTGGCCTCTC | AACATTTATGGGAGCCATTGAG |
| **rs239577** | **TTK** | AAGGGGAGACTGGAGAAAAGTC | TGAGGAACTGTGTGTTTTGAGG |
| **rs239579** | **TTK** | GGATTGTAAGGGAACTGAAAACC | TTTGGCAATTATGAAAATGCTG |
| **rs9294265** | **UBE2CBP** | AAGCAGTCCAATGAGGAAGAAG | GATGAAAAGGCACCAGTAGTCC |
| **rs9449554** | **UBE2CBP** | CATGCACACACACTCACAAATC | AACCCATTCATCTGTGGAAATC |
| **rs9449559** | **UBE2CBP** | CCAGGTGCTCATTTTCTTTAGC | GAATTCACCTTTCTGCACCTTC |
| **rs2300465** | **ME1** | GACCTTGGAAACTCATGTACCC | GCTTGCCATCTACCTCAGAAAG |
| **rs4706990** | **ME1** | TAGGGAAATGCTGGTTAAAAGG | GAGTGCTGACATCTGACATTGAG |
| **rs9294274** | **ME1** | CATAACTCCCCTCAGCATTCTC | TAAACTCCTTCGCAACCTTCTC |
| **rs1059307** | **SNORD50** | CTACTCGTCCACACTCAGAACG | TGTGCAATTCGGATTAATGAAG |
| **rs9444704** | **MDN1** | TGAACCCTCATGAGTCACTGTC | CACACATGTTCATTGGTGTTTG |
| **rs7452753** | **ASSC3** | GAATTACAGCTTTGCCTTGGTC | TCACGGGGAAAATAATCAAATG |
| **rs11155596** | **ASSC3** | CACATCTAAACAAAGGCTGATG | TCCTGTGTAACTTGCAGAATGG |
| **rs9390685** | **ASSC3** | TAGGATGAACCCAAGAACATGC | GCCAGGTTTGTGAGGAATATTAG |
| **rs9386257** | **ASSC3** | ACAGTTGCCAAATTTCTTACGC | TTGCTCTTCCTAGACCCGTTAC |
| **rs4840157** | **ASSC3** | CTAGGTTTCCCATTCAAAGGTG | TAATCTGTGCTAGGCCCTTTTC |
| **rs3777980** | **ROS1** | ATTGCAAAAATGACAAGCTTCC | TGAGGGTTGGCTAAGAAGAGAG |
| **rs9320600** | **ROS1** | AAAATGATACAATGGGCTTTGG | TGTGCTGCAGAAAACAACAAAC |
| **rs12194883** | **PHACTR2** | AGGGAACTGGATATGTGGATTG | TCCATCGAAAATCACTTGACAC |
| **rs10484829** | **PHACTR2** | TCACTCTGAACTCAGCTTGGAC | ATGTCGCCTCTCCTCAGTTAAG |
| **rs9390356** | **FBXO30** | CCAAGGACCACACTTTAGGAAC | CATGGAAACCATCTCTCTCTCC |
| **rs10872580** | **FBXO30** | AACCAGCCTGTACTCAGGTAGC | TGTGGTGTACAGGAACAAAACC |
| **rs719787** | **RAB32** | AAAAATGGGGAGTTGTTCAGTG | ACAGAAAAATCCAATTCCCTTG |
| **rs11155475** | **RAB32** | TGTGAGCACTTAATCTGACACG | CTTAACCAGCGTTTCACCTTTC |

## Table A1.4. Probes used in asynchronous replication assay.

| Locus | Probe | Position | REF |
|---|---|---|---|
| | | | |
| MANEA | Fosmid:G248P8476B7 | 6:96144961-96182585 | |
| ASAR6 | BAC:RP11-374I15 | 6:96168652-96344099 | |
| FUT9 | BAC:RP11-75N19 | 6:96502279-96663444 | Allen et.al.[87] |
| BAC959I6 | BAC:RP11-959I6 | 6:96687004-96881427 | |
| FHL5 | Fosmid:G248P86054G4 | 6:97123343-97162459 | Allen et.al.[87] |
| | | | |
| HTR1E | BAC:RP11-152M1 | 6:87597666-87763359 | Allen et.al.[87] |
| KCNQ5 | BAC:RP11-106C24 | 6:73640760-73789607 | Gimelbrant et. al.[86] |
| ME1 | BAC:RP11-703G14 | 6:84052213-87763359 | Gimelbrant et. al.[86] |
| FRK | BAC:RP11-819F13 | 6:116319345-116492199 | Gimelbrant et. al.[86] |
| | | | |
| PTK6 | BAC:RP11-95N13 | 20:61564623-61727698 | |
| LARP | BAC:CTD-2546N4 | 5:154075536-154286044 | Ensminger & Chess.[100] |
| C9orf43 | BAC:CTD-2167I10 | 9:115153143-115242032 | Ensminger & Chess.[100] |

The references associated with predicted or known mono-allelic genes, as well as synchronously replicating loci are indicated in the REF column.

**Table A1.5. Expression of predicted or known mono-allelic genes.**

Expression analysis was done by RT-PCR on total RNA isolated from P175 and two chromosome 6 mono-chromosomal hybrids [L(Hyg)-1 and L(Neo)-38]. PCR primers were designed to amplify SNPs present in HTD114 DNA. The SNP genotypes were determined using the Affymetrix 500K SNP array (Figs A1.20 and A1.21). Genes were predicted or known to be mono-allelically expressed according to the provided reference. (Table on next page).

## Table A1.5. Expression of predicted or known mono-allelic genes.

| DNA Feature | SNP Location* | Gene Region | Cell Lines: P175 Detected | Cell Lines: L(Hyg)-1 Detected | Cell Lines: L(Neo)-38 Detected | Assay(s) | Reference |
|---|---|---|---|---|---|---|---|
| rs6458511 | 46,678,742 | CYP39A1 | — | — | — | r,h | 1 |
| rs6913639 | 55,141,614 | HCRTR2 | — | — | — | r,h | 1 |
| rs3122155 | 55,189,315 | HCRTR2 | — | — | — | r,h | 1 |
| rs9475663 | 56,287,451 | COL21A1 | — | — | — | r,h | 1 |
| rs6901838 | 56,303,235 | COL21A1 | — | — | — | r,h | 1 |
| rs239554 | 80,776,234 | TTK | + | + | + | r,h | 1 |
| rs239577 | 80,787,733 | TTK | + | + | + | r,h | 1 |
| rs239579 | 80,790,628 | TTK | + | + | + | r,h | 1 |
| rs9294265 | 83,665,958 | UBE2CBP | + | + | + | r,h | 2 |
| rs9449554 | 83,666,767 | UBE2CBP | + | + | + | r,h | 2 |
| rs9449559 | 83,670,399 | UBE2CBP | + | + | + | r,h | 2 |
| rs2300465 | 83,984,854 | ME1 | G/T | + | + | r,h,s | 2 |
| rs4706990 | 84,164,784 | ME1 | A/T | + | + | r,h,s | 2 |
| rs9294274 | 84,167,909 | ME1 | C/T | + | + | r,h,s | 2 |
| rs1059307 | 86,444,607 | SNORD50 | — | — | — | r,h | — |
| rs9444704 | 90,512,150 | MDN1 | + | + | + | r,h | 3 |
| rs7452753 | 101,160,447 | ASSC3 | + | + | + | r,h | 1 |
| rs11155596 | 101,259,352 | ASSC3 | + | + | + | r,h | 1 |
| rs9390685 | 101,331,446 | ASSC3 | + | + | + | r,h | 1 |
| rs9386257 | 101,358,790 | ASSC3 | + | + | + | r,h | 1 |
| rs4840157 | 101,433,096 | ASSC3 | + | + | + | r,h | 1 |
| rs3777980 | 117,840,666 | ROS1 | — | — | — | r,h | 1 |
| rs9320600 | 117,851,338 | ROS1 | — | — | — | r,h | 1 |
| rs12194883 | 144,071,842 | PHACTR2 | C/T | + | + | r,h,s | 2 |
| rs10484829 | 144,125,492 | PHACTR2 | C/T | + | + | r,h,s | 2 |
| rs9390356 | 146,160,676 | FBXO30 | + | + | + | r,h | 1 |
| rs10872580 | 146,181,974 | FBXO30 | + | + | + | r,h | 1 |
| rs719787 | 146,891,370 | RAB32 | + | + | + | r,h | 1 |
| rs11155475 | 146,914,378 | RAB32 | + | + | + | r,h,s | 1 |

*UCSC Genome Browser, Mar. 2006 Assembly (hg18)

[1] Allen et al.[87]

[2] Gimelbrant et al.[86]

[3] Heap et al.[140]

## Erratum

Some changes have been made from the published version of this Supplemental Material. Various editing changes have been made for clarity, without any change in the meaning. The introduction section was added to this dissertation, but does not appear in the published version. Additionally, three columns have been removed from Table A1.5 for space considerations, resulting in no loss of information.

# APPENDIX II

Author Contributions

## Author Contributions by Figure

**Figure 2.1.** The t(6;9) translocatino was generated by Dan Pankratz (D.P.). DNA FISH and replication timing assays were performed by Leslie Smith (L.S.)

**Figure 2.2.** The chromosome 6 deletions were generated, characterized and confirmed by Mathew Thayer (M.T.) and Eric Stoffregen (E.S.). DNA FISH and replication timing assays were performed by L.S.

**Figure 2.3.** ∆175-5'-23 was generated by E.S. DNA FISH and replication timing assays were performed by L.S.

**Figure 2.4.** Allele-specific RT-PCR, RT-PCR, PCR, and sequencing was done by E.S. ∆175 clones were generated, characterized and confirmed by M.T. and E.S. RNA-DNA FISH was performed by L.S.

**Figure 2.5.** Allele-specific RT-PCR, RT-PCR, PCR, and sequencing was done by E.S. RNA-DNA FISH was performed by L.S. Subcloning of GM13130 and characterization of the subclones was performed by E.S.

**Figure 2.6.** DNA FISH assays and analysis of asynchronous replication timing were performed by L.S.

**Table 2.1.** DNA FISH assays and analysis of asynchronous replication timing were performed by L.S.

**Figure 2.7.** RNA-DNA FISH was performed by L.S. Δ175-1i was generated, characterized and confirmed by M.T. and E.S.

**Figure 3.1.** Creation of rAAV-5'-AP-loxP construct was done by E.S. from the 5'AP-loxP construct created by Daniel Stauffer (D.S.) Generation of rAAV, screening for, and characterization of, deletions was performed by E.S.

**Figure 3.2.** The ~46 kb deletion was generated by E.S. Replication timing assays were performed by L.S.

**Figure A1.1.** Engineering strategy is from [9-10].

**Figure A1.2.** Generation and characterization of balanced translocations is from [9-10].

**Figure A1.3.** Generation and characterization of the t(6;10) is from [9-10].

**Figure A1.4.** This Southern blot screen for alternative partner translocations was done by M.T.

**Figure A1.5.** The (6;17) alternative partner translocation was generated by D.P. G-banding, FISH, the replication timing assay and karyotyping of the translocation was performed by L.S.

**Figure A1.6.** The (6;7) alternative partner translocation was generated by D.P. G-banding, FISH, and the replication timing assay for the translocation were performed by L.S.

**Figure A1.7.** The (6;9) alternative partner translocation was generated by D.P. G-banding, FISH, and the DMC assay for the translocation were performed by L.S.

**Figure A1.8.** The (6;9) alternative partner translocation was generated by D.P. FISH and the replication timing assay for the translocation were performed by L.S.

**Figure A1.9.** The (6;8) alternative partner translocation was generated by D.P. G-banding, FISH, and the replication timing assay for the translocation were performed by L.S.

**Figure A1.10.** The (1;10) alternative partner translocation was generated by D.P. G-banding, FISH, and the replication timing assay for the translocation were performed by L.S.

**Figure A1.11.** The (4;10) alternative partner translocation was generated by D.P. G-banding, FISH, and the replication timing assay for the translocation were performed by L.S.

**Figure A1.12.** The (5;10) alternative partner translocation was generated by D.P. G-banding, FISH, and the replication timing assay for the translocation were performed by L.S.

**Figure A1.13.** The replication timing assay for chromosome 6 in P175 was performed by L.S.

**Figure A1.14.** The replication timing assay for chromosome 6 in P175 was performed by L.S.

**Figure A1.15.** The replication timing analysis of chromosome 10 in P175 cells was performed by L.S.

**Figure A1.16.** The replication timing analyses of chromosomes 1, 4, 5, 7, 8, 9, and 17 in P175 cells was performed by L.S.

**Figure A1.17.** Southern blot hybridization of chromosome 6 deletion clones was performed by M.T. Junction PCR on chromosome 6 deletion clones was performed by E.S.

**Figure A1.18.** Engineering scheme for chromosome 6 distal deletions by M.T.

**Figure A1.19.** Chromosome 6 distal deletions were generated, characterized and confirmed by M.T. and E.S.

**Figure A1.20.** SNP analysis was performed by D.P. and M.T.

**Figure A1.21.** SNP analysis was performed by D.P. and M.T.

**Figure A1.22.** Δ175-2c was generated, characterized and confirmed by M.T. and E.S. The replication timing assay was performed by L.S.

**Figure A1.23.** Δ175-1i cells was generated, characterized and confirmed by M.T. and E.S. The replication timing assay was performed by L.S.

**Figure A1.24.** Δ175-11d was generated, characterized and confirmed by M.T. and E.S. The replication timing assay was performed by L.S.

**Figure A1.25.** Chromosome 6 distal deletion lines were generated, characterized and confirmed by M.T. and E.S. The replication timing assays were performed by L.S.

**Figure A1.26.** PCR, RT-PCR, strand-specific PCR, and sequencing was performed by E.S.

**Figure A1.27.** RNA-DNA FISH was performed by L.S.

**Figure A1.28.** RNA-DNA FISH was performed by L.S.

**Figure A1.29.** Δ175-1i cells was generated, characterized and confirmed by M.T. and E.S. RNA-DNA FISH was performed by L.S.

**Figure A1.30.** RNA-DNA FISH was performed by L.S.

**Figure A1.31.** Δ175-1i cells was generated, characterized and confirmed by M.T. and E.S. RNA-DNA FISH was performed by L.S.

**Figure A1.32.** RNA-DNA FISH was performed by L.S.

**Figure A1.33.** Δ175-1i cells was generated, characterized and confirmed by M.T. and E.S. RNA-DNA FISH was performed by L.S.

**Figure A1.34.** RNA-DNA FISH for FHL5OST was performed by L.S.

**Figure A1.35.** Δ175-1i cells was generated, characterized and confirmed by M.T. and E.S. RNA-DNA FISH was performed by L.S.

**Figure A1.36.** DNA FISH replication assay was performed by L.S.

**Figure A1.37.** Two-color DNA FISH assay for coordination in asynchronous replication was performed by L.S.

**Figure A1.38.** RT-PCR was performed by E.S.

**Table A1.1.** LAM-PCR to determine the lentiviral integration sites was performed by Dr. Michael Harkey and the Fred Hutchinson Cancer Research Center Clonal Analysis Core. Confirmation of the integration sites was done by E.S. Mapping of the loxP-3'RT integration site was done by E.S.

**Table A1.2.** All assays were performed by E.S. with some assistance from Nathan Donley (N.D.)

**Table A1.3.** All primer design was done by E.S.

**Table A1.4.** Probes were selected by L.S.

**Table A1.5.** All assays were performed by E.S.

# REFERENCES


1. Nichols, W.W., A. Levan, P. Aula, and E. & Norrby. (1964) Extreme chromosome breakage induced by measles virus in different in vitro systems. Hereditas, **51**(380).
2. Nichols, W.W., Levan, A., Aula, P., et al. (1965) Chromosome damage associated with the measles virus in vitro. Hereditas, **54**(1), 101-118.
3. Kato, H. and Sandberg, A.A. (1968) Chromosome pulverization in human cells with micronuclei. J Natl Cancer Inst, **40**(1), 165-179.
4. Miles, C.P. and O'Neill, F. (1969) 3H labeling patterns of permanent cell line chromosomes showing pulverization or accentuated secondary constrictions. J Cell Biol, **40**(2), 553-561.
5. Ikeuchi, T., Weinfeld, H., and Sandberg, A.A. (1972) Chromosome pulverization in micronuclei induced by tritiated thymidine. J Cell Biol, **52**(1), 97-104.
6. Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1998) Genetic instabilities in human cancers. Nature, **396**(6712), 643-649.
7. Heng, H.H., Bremer, S.W., Stevens, J., et al. (2006) Cancer progression by non-clonal chromosome aberrations. J Cell Biochem, **98**(6), 1424-1435.

8. Smith, L., Plug, A., and Thayer, M. (2001) Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations. Proc Natl Acad Sci U S A, **98**(23), 13300-13305.

9. Breger, K.S., Smith, L., Turker, M.S., et al. (2004) Ionizing radiation induces frequent translocations with delayed replication and condensation. Cancer Research, **64**(22), 8231-8238.

10. Breger, K.S., Smith, L., and Thayer, M.J. (2005) Engineering translocations with delayed replication: evidence for cis control of chromosome replication timing. Hum Mol Genet, **14**(19), 2813-2827.

11. Stoler, D.L., Chen, N., Basik, M., et al. (1999) The onset and extent of genomic instability in sporadic colorectal tumor progression. Proc Natl Acad Sci U S A, **96**(26), 15121-15126.

12. Phear, G., Bhattacharyya, N.P., and Meuth, M. (1996) Loss of heterozygosity and base substitution at the APRT locus in mismatch-repair-proficient and -deficient colorectal carcinoma cell lines. Mol Cell Biol, **16**(11), 6516-6523.

13. Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1997) Genetic instability in colorectal cancers. Nature, **386**(6625), 623-627.

14. Thompson, S.L., Bakhoum, S.F., and Compton, D.A. (2010) Mechanisms of chromosomal instability. Curr Biol, **20**(6), R285-295.

15. Marder, B.A. and Morgan, W.F. (1993) Delayed chromosomal instability induced by DNA damage. Mol Cell Biol, **13**(11), 6667-6677.

16. Grosovsky, A.J., Parks, K.K., Giver, C.R., et al. (1996) Clonal analysis of delayed karyotypic abnormalities and gene mutations in radiation-induced genetic instability. Mol Cell Biol, **16**(11), 6252-6262.

17. Huang, L., Snyder, A.R., and Morgan, W.F. (2003) Radiation-induced genomic instability and its implications for radiation carcinogenesis. Oncogene, **22**(37), 5848-5854.

18. Geigl, J.B., Obenauf, A.C., Schwarzbraun, T., et al. (2008) Defining 'chromosomal instability'. Trends Genet, **24**(2), 64-69.

19. Stephens, P.J., Greenman, C.D., Fu, B., et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell, **144**(1), 27-40.

20. Pflumm, M.F. (2002) The role of DNA replication in chromosome condensation. Bioessays, **24**(5), 411-418.

21. DePamphilis, M.L. (2003) The 'ORC cycle': a novel pathway for regulating eukaryotic DNA replication. Gene, **310**, 1-15.

22. Sclafani, R.A. and Holzen, T.M. (2007) Cell cycle regulation of DNA replication. Annu Rev Genet, **41**, 237-280.

23. Mechali, M. (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol, **11**(10), 728-738.

24. Huberman, J.A. and Riggs, A.D. (1966) Autoradiography of chromosomal DNA fibers from Chinese hamster cells. Proc Natl Acad Sci U S A, **55**(3), 599-606.

25. Wohlschlegel, J.A., Dwyer, B.T., Dhar, S.K., et al. (2000) Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science, **290**(5500), 2309-2312.

26. Fujita, M. (2006) Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. Cell Div, **1**, 22.

27. Gilbert, D.M., Takebayashi, S.I., Ryba, T., et al. (2010) Space and Time in the Nucleus: Developmental Control of Replication Timing and Chromosome Architecture. Cold Spring Harb Symp Quant Biol.

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

29. Lucas, I. and Feng, W. (2003) The essence of replication timing: determinants and significance. Cell Cycle, **2**(6), 560-563.

30. Schwaiger, M. and Schubeler, D. (2006) A question of timing: emerging links between transcription and replication. Curr Opin Genet Dev, **16**(2), 177-183.

31. O'Keefe, R.T., Henderson, S.C., and Spector, D.L. (1992) Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. J Cell Biol, **116**(5), 1095-1110.

32. Nakamura, H., Morita, T., and Sato, C. (1986) Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. Exp Cell Res, **165**(2), 291-297.

33. Nakayasu, H. and Berezney, R. (1989) Mapping replicational sites in the eucaryotic cell nucleus. J Cell Biol, **108**(1), 1-11.

34. Barr, M.L. and Bertram, E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature, **163**(4148), 676.

35. Ferguson, B.M. and Fangman, W.L. (1992) A position effect on the time of replication origin activation in yeast. Cell, **68**(2), 333-339.

36. Raghuraman, M.K., Brewer, B.J., and Fangman, W.L. (1997) Cell cycle-dependent establishment of a late replication program. Science, **276**(5313), 806-809.

37. Ofir, R., Wong, A.C., McDermid, H.E., et al. (1999) Position effect of human telomeric repeats on replication timing. Proc Natl Acad Sci U S A, **96**(20), 11434-11439.

38. Belmont, A.S. (2006) Mitotic chromosome structure and condensation. Curr Opin Cell Biol, **18**(6), 632-638.

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

40. Koshland, D. and Strunnikov, A. (1996) Mitotic chromosome condensation. Annu Rev Cell Dev Biol, **12**, 305-333.

41. Johnson, R.T. and Rao, P.N. (1970) Mammalian cell fusion: induction of premature chromosome condensation in interphase nuclei. Nature, **226**(5247), 717-722.

42. Laird, C., Jaffe, E., Karpen, G., et al. (1987) Fragile sites in human chromosomes as regions of late-replicating DNA. Trends in Genetics, **3**, 274-281.

43. Durkin, S.G. and Glover, T.W. (2007) Chromosome fragile sites. Annu Rev Genet, **41**, 169-192.

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

45. Kastan, M.B. and Bartek, J. (2004) Cell-cycle checkpoints and cancer. Nature, **432**(7015), 316-323.

46. Clarke, P.R. and Allan, L.A. (2009) Cell-cycle control in the face of damage--a matter of life or death. Trends Cell Biol, **19**(3), 89-98.

47. Weinert, T. (1998) DNA damage checkpoints update: getting molecular. Curr Opin Genet Dev, **8**(2), 185-193.

48. Casper, A.M., Nghiem, P., Arlt, M.F., et al. (2002) ATR regulates fragile site stability. Cell, **111**(6), 779-789.

49. Syljuasen, R.G. (2007) Checkpoint adaptation in human cells. Oncogene, **26**(40), 5833-5839.

50. Sandell, L.L. and Zakian, V.A. (1993) Loss of a yeast telomere: arrest, recovery, and chromosome loss. Cell, **75**(4), 729-739.

51. Toczyski, D.P., Galgoczy, D.J., and Hartwell, L.H. (1997) CDC5 and CKII control adaptation to the yeast DNA damage checkpoint. Cell, **90**(6), 1097-1106.

52. Yoo, H.Y., Kumagai, A., Shevchenko, A., et al. (2004) Adaptation of a DNA replication checkpoint response depends upon inactivation of Claspin by the Polo-like kinase. Cell, **117**(5), 575-588.

53. Syljuasen, R.G., Jensen, S., Bartek, J., et al. (2006) 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, **66**(21), 10253-10257.

54. Chang, B.H., Smith, L., Huang, J., et al. (2007) Chromosomes with delayed replication timing lead to checkpoint activation, delayed recruitment of Aurora B and chromosome instability. Oncogene, **26**(13), 1852-1861.

55. Meraldi, P., Honda, R., and Nigg, E.A. (2004) Aurora kinases link chromosome segregation and cell division to cancer susceptibility. Curr Opin Genet Dev, **14**(1), 29-36.

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

57. Gupta, P.K., Shao, C., Zhu, Y., et al. (1997) Loss of heterozygosity analysis in a human fibrosarcoma cell line. Cytogenet Cell Genet, **76**(3-4), 214-218.

58. Zheng, B., Sage, M., Sheppeard, E.A., et al. (2000) Engineering mouse chromosomes with Cre-loxP: range, efficiency, and somatic applications. Mol Cell Biol, **20**(2), 648-655.

59. Lucchesi, J.C., Kelly, W.G., and Panning, B. (2005) Chromatin remodeling in dosage compensation. Annu Rev Genet, **39**, 615-651.

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

61. Chow, J.C., Yen, Z., Ziesche, S.M., et al. (2005) Silencing of the mammalian X chromosome. Annu Rev Genomics Hum Genet, **6**, 69-92.

62. Chow, J. and Heard, E. (2009) X inactivation and the complexities of silencing a sex chromosome. Curr Opin Cell Biol, **21**(3), 359-366.

63. Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature, **190**, 372-373.

64. Wutz, A. and Gribnau, J. (2007) X inactivation Xplained. Curr Opin Genet Dev, **17**(5), 387-393.

65. Trojer, P. and Reinberg, D. (2007) Facultative heterochromatin: is there a distinctive molecular signature? Mol Cell, **28**(1), 1-13.

66. Morishima, A., Grumbach, M.M., and Taylor, J.H. (1962) Asynchronous duplication of human chromosomes and the origin of sex chromatin. Proc Natl Acad Sci U S A, **48**, 756-763.

67. Lee, J.T. (2010) The X as Model for RNA's Niche in Epigenomic Regulation. Cold Spring Harbor Perspectives in Biology, **2**(9).

68. Senner, C.E. and Brockdorff, N. (2009) Xist gene regulation at the onset of X inactivation. Curr Opin Genet Dev, **19**(2), 122-126.

69. Chaumeil, J., Le Baccon, P., Wutz, A., et al. (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev, **20**(16), 2223-2237.

70. Bacher, C.P., Guggiari, M., Brors, B., et al. (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol, **8**(3), 293-299.

71. Xu, N., Tsai, C.L., and Lee, J.T. (2006) Transient homologous chromosome pairing marks the onset of X inactivation. Science, **311**(5764), 1149-1152.

72. Xu, N., Donohoe, M.E., Silva, S.S., et al. (2007) Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat Genet, **39**(11), 1390-1396.

73. Zhao, J., Sun, B.K., Erwin, J.A., et al. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science, **322**(5902), 750-756.

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

75. Huang, S. (2000) Review: perinucleolar structures. J Struct Biol, **129**(2-3), 233-240.

76. Chow, J.C., Ciaudo, C., Fazzari, M.J., et al. (2010) LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell, **141**(6), 956-969.

77. Bailey, J.A., Carrel, L., Chakravarti, A., et al. (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci U S A, **97**(12), 6634-6639.

78. Lyon, M.F. (1998) X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet, **80**(1-4), 133-137.

79. Gartler, S.M. and Riggs, A.D. (1983) Mammalian X-chromosome inactivation. Annu Rev Genet, **17**, 155-190.

80. Hansen, R.S. (2003) X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis. Hum Mol Genet, **12**(19), 2559-2567.

81. Wang, Z., Willard, H.F., Mukherjee, S., et al. (2006) Evidence of influence of genomic DNA sequence on human X chromosome inactivation. PLoS Comput Biol, **2**(9), e113.

82. Diaz-Perez, S., Ouyang, Y., Perez, V., et al. (2005) The element(s) at the nontranscribed Xist locus of the active X chromosome controls chromosomal replication timing in the mouse. Genetics, **171**(2), 663-672.

83. Diaz-Perez, S.V., Ferguson, D.O., Wang, C., et al. (2006) 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, **174**(3), 1115-1133.

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

85. Zakharova, I.S., Shevchenko, A.I., and Zakian, S.M. (2009) Monoallelic gene expression in mammals. Chromosoma, **118**(3), 279-290.

86. Gimelbrant, A., Hutchinson, J.N., Thompson, B.R., et al. (2007) Widespread monoallelic expression on human autosomes. Science, **318**(5853), 1136-1140.

87. Allen, E., Horvath, S., Tong, F., et al. (2003) High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. Proc Natl Acad Sci U S A, **100**(17), 9940-9945.

88. Pauler, F.M., Koerner, M.V., and Barlow, D.P. (2007) Silencing by imprinted noncoding RNAs: is transcription the answer? Trends Genet, **23**(6), 284-292.

89. Kalantry, S., Purushothaman, S., Bowen, R.B., et al. (2009) Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. Nature, **460**(7255), 647-651.

90. Mattick, J.S. and Makunin, I.V. (2006) Non-coding RNA. Hum Mol Genet, **15 Spec No 1**, R17-29.

91. Kanduri, C. (2011) Kcnq1ot1: A chromatin regulatory RNA. Semin Cell Dev Biol.

92. Mohammad, F., Pandey, R.R., Nagano, T., et al. (2008) Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. Mol Cell Biol, **28**(11), 3713-3728.

93. Pandey, R.R., Mondal, T., Mohammad, F., et al. (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell, **32**(2), 232-246.

94. Peters, A.H., Kubicek, S., Mechtler, K., et al. (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell, **12**(6), 1577-1589.

95. Czermin, B., Melfi, R., McCabe, D., et al. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell, **111**(2), 185-196.

96. Cerrato, F., Dean, W., Davies, K., et al. (2003) Paternal imprints can be established on the maternal Igf2-H19 locus without altering replication timing of DNA. Hum Mol Genet, **12**(23), 3123-3132.

97. Gribnau, J., Hochedlinger, K., Hata, K., et al. (2003) Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. Genes Dev, **17**(6), 759-773.

98. Shemer, R., Hershko, A.Y., Perk, J., et al. (2000) The imprinting box of the Prader-Willi/Angelman syndrome domain. Nat Genet, **26**(4), 440-443.

99. Singh, N., Ebrahimi, F.A., Gimelbrant, A.A., et al. (2003) Coordination of the random asynchronous replication of autosomal loci. Nat Genet, **33**(3), 339-341.

100. Ensminger, A.W. and Chess, A. (2004) Coordinated replication timing of monoallelically expressed genes along human autosomes. Hum Mol Genet, **13**(6), 651-658.

101. Dutta, D., Ensminger, A.W., Zucker, J.P., et al. (2009) Asynchronous replication and autosome-pair non-equivalence in human embryonic stem cells. PLoS One, **4**(3), e4970.

102. Gimelbrant, A.A., Ensminger, A.W., Qi, P., et al. (2005) Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells. J Biol Chem, **280**(2), 1354-1359.

103. Tycko, B. (2010) Allele-specific DNA methylation: beyond imprinting. Hum Mol Genet, **19**(R2), R210-220.

104. Camargo, M. and Cervenka, J. (1982) Patterns of DNA replication of human chromosomes. II. Replication map and replication model. Am J Hum Genet, **34**(5), 757-780.

105. Cohen, S.M., Cobb, E.R., Cordeiro-Stone, M., et al. (1998) Identification of chromosomal bands replicating early in the S phase of normal human fibroblasts. Exp Cell Res, **245**(2), 321-329.

106. Mills, A.A. and Bradley, A. (2001) From mouse to man: generating megabase chromosome rearrangements. Trends Genet, **17**(6), 331-339.

107. Harkey, M.A., Kaul, R., Jacobs, M.A., et al. (2007) Multiarm high-throughput integration site detection: limitations of LAM-PCR technology and optimization for clonal analysis. Stem Cells Dev, **16**(3), 381-392.

108. Yeo, G.W., Coufal, N.G., Liang, T.Y., et al. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol, **16**(2), 130-137.

109. Selig, S., Okumura, K., Ward, D.C., et al. (1992) Delineation of DNA replication time zones by fluorescence in situ hybridization. EMBO J, **11**(3), 1217-1225.

110. Mostoslavsky, R., Singh, N., Tenzen, T., et al. (2001) Asynchronous replication and allelic exclusion in the immune system. Nature, **414**(6860), 221-225.

111. Azuara, V., Brown, K.E., Williams, R.R., et al. (2003) Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. Nat Cell Biol, **5**(7), 668-674.

112. Schlesinger, S., Selig, S., Bergman, Y., et al. (2009) Allelic inactivation of rDNA loci. Genes Dev, **23**(20), 2437-2447.

113. Krueger, C. and Morison, I.M. (2008) Random monoallelic expression: making a choice. Trends Genet, **24**(6), 257-259.

114. Ohlsson, R. (2007) Genetics. Widespread monoallelic expression. Science, **318**(5853), 1077-1078.

115. Tycko, B. Allele-specific DNA methylation: beyond imprinting. Hum Mol Genet, **19**(R2), R210-220.

116. Payer, B. and Lee, J.T. (2008) X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet, **42**, 733-772.

117. Brown, C.J. and Willard, H.F. (1994) The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. Nature, **368**(6467), 154-156.

118. Csankovszki, G., Panning, B., Bates, B., et al. (1999) Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet, **22**(4), 323-324.

119. Zhu, Y., Bye, S., Stambrook, P.J., et al. (1994) Single-base deletion induced by benzo[a]pyrene diol epoxide at the adenine phosphoribosyltransferase locus in human fibrosarcoma cell lines. Mutat Res, **321**(1-2), 73-79.

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

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

122. Vasileva, A. and Jessberger, R. (2005) Precise hit: adeno-associated virus in gene targeting. Nat Rev Microbiol, **3**(11), 837-847.

123. Vasileva, A., Linden, R.M., and Jessberger, R. (2006) Homologous recombination is required for AAV-mediated gene targeting. Nucleic Acids Res, **34**(11), 3345-3360.

124. Lieber, M.R., Ma, Y., Pannicke, U., et al. (2003) Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol, **4**(9), 712-720.

125. Heyer, W.D., Ehmsen, K.T., and Liu, J. (2010) Regulation of homologous recombination in eukaryotes. Annu Rev Genet, **44**, 113-139.

126. Russell, D.W. and Hirata, R.K. (1998) Human gene targeting by viral vectors. Nat Genet, **18**(4), 325-330.

127. Hirata, R.K. and Russell, D.W. (2000) Design and packaging of adeno-associated virus gene targeting vectors. J Virol, **74**(10), 4612-4620.

128. Hirata, R., Chamberlain, J., Dong, R., et al. (2002) Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. Nat Biotechnol, **20**(7), 735-738.

129. Sauer, B. and McDermott, J. (2004) DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. Nucleic Acids Res, **32**(20), 6086-6095.

130. Anastassiadis, K., Fu, J., Patsch, C., et al. (2009) Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. Dis Model Mech, **2**(9-10), 508-515.

131. Kotake, Y., Nakagawa, T., Kitagawa, K., et al. (2011) Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene, **30**(16), 1956-1962.

132. Pope, B.D., Hiratani, I., and Gilbert, D.M. (2010) Domain-wide regulation of DNA replication timing during mammalian development. Chromosome Res, **18**(1), 127-136.

133. Forrester, W.C., Epner, E., Driscoll, M.C., et al. (1990) 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, **4**(10), 1637-1649.

134. Cimbora, D.M., Schubeler, D., Reik, A., et al. (2000) 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, **20**(15), 5581-5591.

135. Simon, I., Tenzen, T., Mostoslavsky, R., et al. (2001) Developmental regulation of DNA replication timing at the human beta globin locus. EMBO J, **20**(21), 6150-6157.

136. Stoffregen, E.P., Donley, N., Stauffer, D., et al. (2011) An autosomal locus that controls chromosome-wide replication timing and mono-allelic expression. Human Molecular Genetics, **20**(12), 2366-2378.

137. Richardson, C. and Jasin, M. (2000) Frequent chromosomal translocations induced by DNA double-strand breaks. Nature, **405**(6787), 697-700.

138. Keim, M., Williams, R.S., and Harwood, A.J. (2004) An inverse PCR technique to rapidly isolate the flanking DNA of dictyostelium insertion mutants. Mol Biotechnol, **26**(3), 221-224.

139. Cailleau-Thomas, A., Coullin, P., Candelier, J.J., et al. (2000) FUT4 and FUT9 genes are expressed early in human embryogenesis. Glycobiology, **10**(8), 789-802.

140. Heap, G.A., Yang, J.H., Downes, K., et al. (2010) Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing. Hum Mol Genet, **19**(1), 122-134.