Identification and Characterization of *cis*-acting DNA Elements that Regulate Early Origin Activation

Thomas J. Pohl

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Bonita J. Brewer, Chair

M.K. Raghuraman, Co-Chair

Maitreya J. Dunham

Sue Biggins

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Identification and Characterization of *cis*-acting DNA Elements that Regulate Early Origin Activation

Thomas J. Pohl

Chair of the Supervisory Committee:
Professor Bonita J. Brewer

Molecular and Cellular Biology Program
Department of Genome Sciences

Duplication of a cell's genome is a central requirement for the propagation of life. This duplication, a highly regulated process known as DNA replication, occurs in S-phase and is initiated at genomic sites termed origins of replication (origins). In eukaryotes, origins are activated (initiate DNA synthesis) on average at different times in S-phase; some are activated in early S-phase while others are activated in late S-phase. This variation in origin activation time allows the genome to be replicated in a temporally staggered fashion. The biological significance and mechanisms behind temporal DNA replication have remained largely elusive. In the budding yeast *Saccharomyces cerevisiae*, the activation times of individual origins are not intrinsic to those origins but instead are governed by surrounding sequences and their local chromatin environment. My research, presented in this dissertation, identifies and characterizes DNA sequences in yeast that function to advance origin activation time. Through this work, I show that centromeres ensure their own early replication by inducing early activation of pericentric origins. I propose and investigate the hypothesis that there is selective pressure to

replicate centromeres in early S-phase. Furthermore, I have also helped to provide evidence for the existence of centromere independent mechanisms that promote early origin activation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There are many people whom I would like to acknowledge that have both mentored and supported me throughout my pursuit and attainment of this degree.

While attending the University of New Mexico (UNM) I had the pleasure of interacting and working with many important mentors and friends. I would like to thank Dr. Leah Larkin, who not only gave me my first job as an undergraduate researcher but also helped to solidify my interest in becoming a scientist. Leah introduced me to a very important mentor in my life, Dr. Maggie Werner-Washburne. Maggie, along with Lupe Atencio and Dr. Steve Phillips, ran the Initiatives for Maximizing Student Development (IMSD) program at UNM. I would like to thank the members of this program and specifically the aforementioned people for their guidance, friendship, and the support that they have given me over the years. Through the IMSD program, I met another very important mentor, Dr. Jac Nickoloff. Jac not only allowed me to conduct research in his lab but took me under his wing and helped me build foundational knowledge in how to think and write like a scientist. Thank you Jac for being a good mentor, role model, and friend.

Looking back on my time as a graduate student in the Brewer/Raghuraman lab at the University of Washington, I truly believe that I could not have asked for a better place to conduct my graduate studies. Over the years, both Bonny and Raghu have not been merely mentors but have become dear friends. As a member of this lab, I have gained valuable experience learning from Bonny's sharp wit and precision in conducting experiments as well as from Raghu's ability to think clearly and to both approach a problem and understand people's thought processes from multiple perspectives. I would like to thank Bonny and Raghu for their mentorship and

friendship, both of which have been instrumental in my development into the scientist that I am today.

Along with Bonny and Raghu, I am thankful for all the members of the Brewer/Raghuraman lab that I have had the pleasure to work with during my graduate studies: Gina Alvino, Sara Di Rienzi, Wenyi Feng, Liz Kwan, Helen Lai, Steven Lee, Kim Lindstrom, Kelsey Lynch, Anne Paul, and Joe Sanchez. I especially want to thank Gina Alvino and Wenyi Feng for showing me the ropes early on in my studies. I am grateful to have met such amazing people and am glad they have become part of my scientific family.

In addition to the members of Bonny and Raghu's lab I would like to thank MaryEllin Robinson for her involvement in both recruiting me to the University of Washington and her strong support though out my graduate career. I would also like to thank my committee members (Dr. Sue Biggins, Dr. Maitreya Dunham, Dr. Trisha Davis, Dr. Toshi Tsukiyama, and Dr. Celeste Burg) for their support, guidance, helpful suggestions on my research project, and for always being available to answer my questions or to simply chat.

On the non-academic front, I would like to thank my beautiful wife, Yanti Pohl, who has been by my side since our undergraduate years at the University of New Mexico. I could not have asked for a more intelligent and caring companion. I would also like to thank my family: Max Pohl (Dad), Debra Pohl (Mom), Lourie Pohl (Step-mom), Don Sieglitz (Step-Dad), Anita Pohl (Sister) along with her family, and Eddie Pohl (Brother) along with his family. You have all helped mold me into the person I am today. Your support has made life away from home a lot easier. Last but not least, I would like to thank all of my friends both back home in New Mexico and in Seattle, Washington for the good times and support.

DEDICATION

This dissertation is dedicated to

Caroline Esther Piro (nee Bernotte), (July 22, 1930)

Harold Edward Piro, (December 31, 1923 -- December 26, 1972)

Maria Arsenia Pohl (nee Mora), (May 15, 1917 -- December 17, 2006)

Macario Pohl, (March 14, 1914 -- March 12, 1998)

Chapter 1

Introduction to Eukaryotic Origin of Replication and Temporal Replication

DNA replication is both a highly complex and highly regulated process by which a cell's genomic content is precisely duplicated prior to the onset of cell division. In metazoans, misregulation of DNA replication has been shown to increase susceptibility to age related disorders and to be implicated in neurodegeneration, autoimmune diseases, and cancer[1-7]. Additionally, inefficient DNA replication can contribute to abnormal human development as exemplified by Bloom Syndrome[1,5,8] and recently, Meier-Gorlin and Wolf-Hirschhorn syndromes[9-11]. To gain an understanding of how efficient DNA replication is maintained and what the cellular consequences of defective DNA replication are, it is imperative to discern not only molecular processes that lead up to the initiation of DNA replication but also mechanisms that regulate those processes.

1.1: Origin Licensing and Initiation

In eukaryotes, DNA replication is initiated during S-phase at multiple genomic sites known as origins of replication (origins). Prior to S-phase, origins are primed or licensed to activate by loading of the pre-replicative complex (pre-RC)[12-14]. The ordered assembly of the pre-RC is conserved among eukaryotes and begins with the recruitment of an essential group of six related proteins collectively known as the origin recognition complex (ORC) to origin sequences[12,13,15-17]. ORC-origin association is followed by recruitment of the core replicative helicase component, Mcm2-7, to double stranded DNA via Cdc6 and Cdt1[12-14,18-21].

Although the biochemical processes leading to pre-RC assembly after ORC binding are conserved, there is wide variation among species in how ORC is recruited and bound to DNA. Most ORC subunits are of the AAA+ ATPase family. In *S. cerevisiae*[15] and *Drosophila*[22,23] ORC-DNA binding is ATP dependent. However ATP merely enhances DNA binding of human ORC[17] and is not required in *Schizosaccharomyces pombe*[24,25]. Additionally, all six ORC subunits necessitate ORC-origin interaction in *Drosophila*[23] whereas this interaction depends on Orc1-5 in *S. cerevisiae*[26]. In *S. pombe*, ORC-origin interaction relies primarily on the Orc4 subunit which contains an additional N-terminal domain composed of nine AT-hook DNA binding motifs that are not present in budding yeast or metazoans[27-29]. In contrast to *S. cerevisiae*, *Drosophila*, and *S. pombe*, where ORC is bound to DNA throughout the cell cycle, all six human ORC subunits bind DNA only during G1- and S-phase[12]. The variation in eukaryotic ORC-DNA contact is consistent with the observations that there is little consensus among origin sequences (described in more detail in chapters 1.2 and 1.3).

The formation of the pre-RC alone does not ensure that an origin will activate during the subsequent S-phase. In fact, only a subset of licensed origin activate in any given S-phase[30]. Furthermore, some origins are dormant and rarely fire despite forming a seemingly normal pre-RC[31]. Rather, initiation of DNA synthesis requires a complex interaction of multiple regulatory and replication fork proteins to associate with pre-RC components (reviewed in[12,13]). In all eukaryotes the assembly and activity of many of these regulatory/replication fork proteins are mediated by cyclin-dependent kinases (CDK) along with the DBF4-dependent kinase (DDK) composed of Cdc7 and its regulatory partner, Dbf4. DDK and CDK are required for both activation of the replicative helicase (GINS-Mcm2-7-Cdc45 complex)[32] and loading of the replication machinery (replisome) that is responsible for DNA synthesis[12,13,33,34]. Activation of

the helicase promotes unwinding of the origin DNA duplex resulting in the formation of a replication bubble. Concurrently, the pre-RC is disassembled; Cdc6 and Cdt1 dissociate while Mcm2-7 remains with replisomes that are formed at each replication fork. DNA synthesis is then free to proceed bidirectionally away from the origin. Eventually replicons, DNA replicated by the two forks from a single origin of replication, fuse resulting in duplication of the entire genome.

It is crucial that origins do not activate more than once during a single S-phase as it has been shown that re-initiation of origins can lead to an accumulation of DNA lesions that ultimately result in loss of both cell viability and genomic integrity[35,36]. Furthermore, re-replication has been shown to induce gene amplification[37], suggesting that breakdown of origin-firing control can have important implications for evolution and in the formation of cancers. To combat the potentially hazardous genomic insults caused by re-replication, eukaryotic cells have developed multiple overlapping regulatory mechanisms to insure that licensing (primarily Mcm2-7 loading) is restricted to late M phase and early G1.

In yeast, this delay of pre-RC formation is mediated by CDK activity. As cells enter S-phase CDK levels increase promoting origin initiation and also preventing re-establishment of the pre-RC through degradation of Cdc6 and Cdt1[38-42], nuclear export of Mcm2-7 and Cdt1[18,20,42,43], and inactivation of ORC[42]. In higher eukaryotes, reformation of the pre-RC is regulated by CDK activity and by the protein Geminin. CDK regulates the pre-RC components by targeting Cdt1 and ORC for degradation[44,45], inducing the dissociation of ORC and Mcm2-7 from DNA[46,47], and through exportation of Cdc6 form the nucleus[48,49]. In addition to CDK mediated degradation of Cdt1, Geminin prevents re-replication by interacting with Cdt1 to prevent loading of Mcm2-7[50-52].

1.2: Origin Structure in Budding Yeast

Eukaryotic origins were first identified in the budding yeast *S. cerevisiae* due to their ability to function as autonomously replicating sequences (ARS) promoting extra-chromosomal maintenance of plasmid DNA[53]. Following this initial discovery an extensive effort has been made to identify DNA sequences that are important for origin initiation. Studies conducted in various budding yeast species (*S. cerevisiae, Lachancea waltii, Lachancea kluyveri, and Kluyveromyces lactis*) have lead to the generalization that budding yeast origins (termed "point" origins) are small (<50 base pairs (bps)), primarily intergenic, found in nucleosome free regions, and characterized by the presence of an AT-rich ARS consensus sequence (ACS)[54-62]. Although the budding yeast ACS is generally AT-rich, one exception to this rule has been observed in a more distantly related species. Curiously, the origin recognition machinery in the budding yeast *Pichia pastoris* appears to recognize two classes of ARSs one that is AT-rich and another that has a distinctive GC-rich ACS (Ivan Liachko and Maitreya Dunham, personal communication). The molecular and evolutionary implications of this GC-rich ACS are yet to be determined.

The most detailed information on origin structure is derived from studies conducted in *S. cerevisiae* in which there are over 300 origins that are preferentially positioned in regions of convergent transcription[63-66]. All of these origins contain an 11-17 bp ACS that is required but not sufficient for origin activation[54,59-61]. In addition to the ACS, mutational analysis conducted on a subset of origins in *S. cerevisiae* has revealed that origins also contain anywhere from one to three B-elements (B1, B2, and B3) that are located up to 100 bps 3' of the T-rich ACS strand[56,59,67-69]. The B-elements act in *cis* to enhance origin efficiency (the percentage of cells activating that origin). Interestingly, B1 and B2 are functionally conserved despite having poor

sequence consensus[68,69]. The B1-element supplements the ACS in forming the ORC binding site[70-72] and, like all the B-elements, is dispensable if at least one other B-element is present.

Early studies examining the B2-element suggest that it functions as a DNA unwinding element (DUE), defined as sequence with low helical stability[68,73-77]. However, later studies indicate that the mechanism of B2-element function might be more complex than acting as a DUE alone. Rao and colleagues have shown that a three base pair mutation of the B2-element in *ARS307* dampens origin function while maintaining helical stability[68]. Furthermore, genetic screens performed on *ARS1* indicate that the B2-element contains a near perfect match to the ACS in inverted orientation[78,79] and functions as a weak binding site for an undetermined pre-RC component[79]. These data suggest that B2 may help to load the MCM complex.

Contrary to the B1 and B2 elements, the B3-element is clearly defined as the binding site for the essential transcription factor Abf1[80,81]. Interestingly, the B3-element can be replaced with other transcription factor binding sites without affecting ARS activity[67]. It has been proposed that the B3-element might function not through direct interaction with ORC but by modifying chromatin in a manner that is favorable to origin activity[71].

In addition to B-elements, some origins have been shown to have stimulatory sequences, termed C-elements that extend up to 200 bp upstream of the ACS. Though C-elements have not been extensively studied, examination of a limited number of origins has revealed that these elements are not essential for origin activation and that they contain binding sites for the transcription factors Abf1, Rap1, Mcm1, or Sum1[82-85]. Currently, the mechanism by which C-elements influence origin initiation is unknown. It is possible that the C-elements function in a similar manner as the B3-elements. This notion is supported by the findings that Abf1 and

Mcm1 can bind both B and C regions[80-82,85-87] and that B3-elements can be exchanged for other transcription factor binding sites while having no ill effect on origin activation[67].

L. waltii and *L. kluyveri* ARSs contain ACSs that are highly similar to that of *S. cerevisiae* in both size (13 and 9 bps, respectively) and at the sequence level[55,62]. Furthermore, the ACSs of these yeasts are essential but not sufficient for origin activation. An exhaustive search for B or C elements has not been conducted in either of these species; however, computational analysis of the ACS and flanking sequences have revealed the presence of auxiliary sequences that are comparable to the B1-element of *S. cerevisiae*. The importance of these auxiliary sequences on origin function has not been tested.

In contrast, *K. lactis* ARSs are not only located in transcriptionally convergent and divergent regions at equal frequency but the ACS also differs structurally in that it consists of a 39 to 50 bp sequence that is both necessary and sufficient for origin activity[57,58]. Interestingly, the ACS contains sequences that are similar to the ACS and B1 element in *S. cerevisiae*. Nonetheless, linker scanning mutation analysis indicates that *K. lactis* ARS function does not exclusively rely on these two elements[58]. Additionally, it was shown that only 5% of *K. lactis* ARSs function in *S. cerevisiae* whereas about 95% of *L. kluveri* ARSs will[62]. These findings suggest that the origin recognition machinery in *K. lactis* has evolved to recognize more specific origins that have incorporated the B-elements into a more complex ACS.

1.3: Origin Structure in Fission Yeast

The genome of the fission yeast *Schizosaccharomyces pombe* is duplicated using between 300 to 400 origins spaced roughly every 38 kb and positioned in intergenic regions with a preference for divergent transcription[63,88]. Although origins in *S. pombe* are AT-rich, they differ

from the origins in budding yeast in a variety of ways. *S. pombe* origins (termed "regional" origins) are at least 500 bp in length, do not contain a complex consensus motif, and are believed to be chosen in a stochastic manner[16,88-93]. Though these origins lack a complex consensus motif as observed in budding yeast, origins in *S. pombe* contain one or more 30 to 55 bp sequence modules that are essential but not sufficient for origin activation[89,94]. These modules have an AT content that is over 77%, contain essential stretches of consecutive adenine/thymine residues, and require at least 450 bp of AT-rich flanking sequence to stimulate origin activity. The high AT content, requirement of consecutive adenine/thymine residues, and lack of complex consensus sequence can be attributed to the finding that the *S. pombe Orc4* subunit contacts DNA through an N-terminal AT hook motif[27-29].

1.4: Origin Structure in Metazoans

Little is known about the sequences that are required for or enhance origin activation in metazoans. Currently, it is unclear if the metazoan origin recognition machinery binds a particular DNA sequence or functions in a sequence independent manner. Investigation of replication at the human β-globin domain[95,96], the human lamin B2 gene[97], and the Drosophila chorion locus[98] indicate that like yeast, metazoan origin activation depends at least in part on sequence. An origin responsible for the replication of the human β-globin domain contains sequences that are essential but not sufficient for origin activation[95,96]. Similar to *S. pombe*, these essential sequences are arranged in one or more independent modules ranging from 314 to 921 bps in length[96]. The human lamin B2 origin contains a 290 bp sequence that is both necessary and sufficient for origin activity[97].

In contrast, some mammalian origins are thought to activate over a broad region such as observed for the DHFR locus in Chinese hamster ovary cells[99-101]. Burhans et al., have shown that at this locus 85% of origin initiations occur within a 450 bp stretch of sequence[102]. Therefore, there appears to be preferred locations of origin initiation even at these more broad regions of replication initiation.

1.5: Temporal Replication

In 1960 it was shown that different regions of Chinese hamster chromosomes replicate at different times in S-phase[103] suggesting that initiation of DNA replication occurs in a staggered manner. It is now known that eukaryotes replicate their genome using hundreds to thousands of origins that, on average, activate at distinct times during S-phase; some are activated in early S-phase while others activate later. Although origins are dispersed throughout the genome, they are often clustered near other origins of similar activation times, thus establishing a temporal replication "program" characterized by large chromosomal domains that replicate at specific times during S-phase[63,104-107]. Curiously, only a subset of origins activate in any given S-phase, yet the temporal order is maintained from one cell cycle to the next. Questions such as what determines the subset of origins that activate in each cell cycle and how is the temporal program conserved between cell cycles remain unanswered, as the molecular mechanisms governing the temporal replication are highly complex involving differences in chromatin state and cis- or trans-acting elements including DNA-protein, and protein-protein interactions.

Temporal replication has been observed in every eukaryote in which replication has been studied and has been shown to have many conserved features. In both budding yeast and mammals, timing is imparted on origins in early G1 phase[108-110]. Although timing is established

at the same juncture that origins are licensed, origins themselves appears not to contain intrinsic timing information[111,112]. Rather, an origin's genomic location determines when that origin becomes active.

Like many other cellular processes, chromatin structure influences DNA replication, yet our current understanding of how replication is regulated by chromatin is scarce. In general, early replication occurs in euchromatic regions while heterochromatic regions replicate late. Replication timing and transcription are correlated in higher eukaryotes[113-118] but not in yeast[93,106]. Nevertheless, the mechanisms that mediate transcription and those that control changes in replication timing do not directly regulate each other[64,105,116,119], suggesting that in higher eukaryotes the two processes are linked through chromatin structure. In *S. cerevisiae,* deletion of the histone deacetylase Rpd3 advances the activation time of origins by increasing the acetylation of their flanking nucleosomes[120-122]. Inhibition of histone deacetylation yields similar results in mammalian cells[123]. It has been suggested that increased acetylation of nucleosomes near origins increases the accessibility of those origins to replication initiation and timing factors[119,124]. This suggestion is appealing because nucleosomes have been shown to hinder origin activity when formed over sequences that are important for origin initiation[125,126]. Consistent with this idea, the Human and *Drosophila* histone acetylase, Hbo1/Hat1 promotes pre-RC assembly[127,128].

Along with changes in chromatin state, certain DNA sequences and their associated protein complexes can also influence when an origin is active. A consistent theme of the replication landscape is that early replicating centromeres and late replicating telomeres are separated by intermittent early and late replicating domains[55,63,106,129,130]. The role that telomeres play in delaying origin activation time has garnered much attention from the scientific

community. Thus multiple studies have made substantial contributions to the understanding of how this process mechanistically occurs. *S. cerevisiae* telomeres were found to impart delayed timing on origins that are located up to 30 kb away[111,131]. This delay is controlled by the Ku complex[132] (a heterodimer consisting of Ku70 and Ku80 subunits), the silencing chromatin protein *Sir3*[133], and the TG_{1-3} repeat binding protein, Rif1[134]. A more recent study indicates that in *S. cerevisiae,* Rif1 also regulates the initiation of a subset of non-telomere associated origins by modulating DDK activity (Hiraga et al., submitted) however it is unclear if Rif1 influences origin activation time. In *S. pombe* the telomere binding proteins Rif1 and Taz1 not only contribute to delayed telomere replication but are also global regulators of origin activation time[135]. Though mammalian Rif1 lacks telomere maintenance activity, it too globally regulates DNA replication[136,137].

Less is known about the mechanisms that contribute to delayed activation of non-telomeric late origins. Timing studies performed on a limited number of origins in *S. cerevisiae* indicate that delayed activation time of at least a subset of non-telomeric origins is independent of the telomere[112]. Instead, the activation time of these origins relies on cryptic flanking DNA sequences that are located up to 6 kb away and function in an additive manner. Preliminary data suggest that initiation of these origins is not regulated through Rif1 (Hiraga et al., submitted). Likewise, delayed timing of a late activating origin in *S. pombe* was show to be additively imposed by three flanking DNA sequence elements[138]. However, these sequences are 10 bps in length and share a G-rich consensus motif. Similar G-rich sequences are associated with telomeres suggesting a possible overlap in mechanism.

Because DNA sequences that delay origin activation time have been identified, early activation has historically been considered to be an origin's default state. Therefore, little

research has been focused on identifying DNA sequence elements that confer early replication. The conservation of early centromere replication[116,129,130,139-141] suggests that centromeres might regulate replication dynamics in a way converse to that of telomeres. Centromeres might promote early origin activation instead of delaying activation time. Consistent with this idea Koren et al. have recently shown that spontaneous neocentromere formation in *Candida albicans* is correlated with early replication[129]. These data provide further evidence of the link between centromeres and early replication, yet it remains unknown if centromeres directly regulate early origin activation.

1.6: Research Goals

This dissertation strives to unravel the complexity that constitutes the replication landscape through the identification and characterization of factors that contribute to early replication. Specifically, this body of research addresses the following questions: (1) Do centromeres advance origin activation time? (2) What is the biological significance of early centromere replication? (3) Can DNA sequences stimulate early origin activation? The answers to these questions will reveal some of the mechanisms that regulate replication dynamics and thus provide insight into a fundamental process of life.

1.7: Chapter Layout

In this dissertation I utilize the budding yeast *Saccharomyces cerevisiae* to examine mechanisms that regulate early origin activation. In chapter 2, I answer the long-standing question of why centromeres reside in early replicating regions. I answer this question by showing that centromere function promotes early activation of nearby origins. I follow up on

this observation in chapter 3 by investigating the importance of early centromere replication. Here I show that early replication of centromeres helps prevent chromosome loss when the spindle checkpoint is compromised. In chapter 4, I characterize a DNA sequence element that can advance origin activation time independently of the centromere. In chapter 5, I present interesting results from some recent studies and discuss their implications in regard to replication timing and the studies that are presented in this dissertation. Chapters 2-4 have an introduction followed by results, discussion, and materials and methods sections. Chapters 2 and 4 have been published in primary literature and appear in this dissertation similar to their published form. Figures and tables that are relevant for each chapter follow the text. Supplemental materials are located in the appendices.

Chapter 2

Functional Centromeres Determine the Activation Time of Pericentric Origins of DNA Replication in *Saccharomyces cerevisiae*

2.1: Introduction

Centromere function, the ability to assemble a kinetochore and mediate chromosome segregation during meiosis and mitosis, is critical for the survival and propagation of eukaryotic organisms. Defects in centromere/kinetochore complexes lead to genome instability and susceptibility to cancer and cell death[142-144]. Therefore, it is important that properly functioning centromeres be established on both sister chromatids following replication of centromeric DNA and prior to the initiation of chromosome segregation. Centromeres in the budding yeast *Saccharomyces cerevisiae* replicate early in S-phase[106,107,130,145,146] and increasing evidence suggests that early centromere replication is conserved among fungi and is prevalent for at least a subset of centromeres in higher eukaryotes[116,129,139-141]. Yet surprisingly little is known about the mechanism that accomplishes this early replication. It has been hypothesized that early replication of centromere DNA provides sufficient time for the centromere-specific histone, CenH3, to be incorporated on both sister chromatids and to ensure subsequent microtubule attachment[147-149]. Consistent with this idea, Feng and colleagues showed that a delay in centromere DNA replication in the absence of the replication checkpoint leads to increased aneuploidy[147]. It is thought that this observed increase in aneuploidy is due to the lack of properly bi-oriented sister chromatids.

Currently, it is unclear if centromeres play an active role in their own early replication by influencing activation of nearby origins of replication (origins). An alternative possibility is that

centromeres have migrated to reside in early replicating regions of the genome over evolutionary time.

Indirect evidence supporting the idea that centromeres play an active role in their own early replication comes from a study examining epigenetic inheritance of centromeres in the pathogenic yeast *Candida albicans*[129]. The centromeres of *C. albicans* are considered regional centromeres akin to those of higher eukaryotes because they are not defined by any distinct DNA sequence. Instead, regional centromeres are defined by broad stretches of repetitive DNA sequences ranging from about 3 kilobases (kb) in *C. albicans* to megabases in humans. Regional centromeres that have been analyzed for replication time have been shown to contain origins within them[129,139,140,150]. Koren et al. showed that the *de novo* formation of an early activated origin within the neocentromeric region is responsible for the early replication of spontaneously formed neocentromeres in *C. albicans*[129]. The authors further showed that the origin recognition complex (ORC), which is essential for origin activation, is recruited to the neocentromere. Thus, centromeres of *C. albicans* appear to recruit at least a subset of the required replication initiation machinery. However, it is unclear from these studies whether centromere function directly influences origin activation time or if these regions merely provide a favorable environment for ORC recruitment, with early origin activation time being determined independently of centromere function.

Distinguishing between these possibilities is difficult because no distinct sequences for either centromeres or origins have been identified in *C. albicans*. Therefore, the function of the centromere responsible for replication initiation cannot effectively be separated from the kinetochore-binding portion of the centromere. For example, removal of centromere DNA in *C. albicans* results in removal of the origins contained within that sequence. Therefore,

determining if centromeres regulate origin activation time requires a situation in which the activity of the origin responsible for centromere replication can be separated from the centromere. The small sizes and well-defined sequences of centromeres and origins in *S. cerevisiae* provide us the opportunity to answer this question.

The origins of *S. cerevisiae* have been extensively studied and are defined by an 11 base pair (bp) consensus sequence that is necessary but not sufficient to initiate DNA synthesis[13]. Unlike *C. albicans*, the centromeres in *S. cerevisiae* are small, spanning only 125 bps and do not, with one possible exception (*CEN3*), contain potential origins[106,151-153]. However, all of the centromeres of *S. cerevisiae* reside in early replicating portions of the genome, suggesting that centromeres play a role in regulating origin activation time.

Early replication is not restricted to portions of the genome immediately flanking centromeres. In fact, early and late replicating blocks of DNA are dispersed throughout the entire genome[145,146]. The patterns in which temporal blocks are arranged indicate that origins can be crudely grouped into four classes with respect to their replication time and chromosomal position: centromere-proximal early, noncentromere-proximal early, telomere-proximal late, and nontelomere-proximal late[106,145,146]. Telomeres delay the activation times of nearby origins while late activation of non-telomeric origins is determined by an unknown mechanism involving unspecified DNA sequences located up to 14 kb away from the origin[111,112]. Plasmids containing the minimum required sequence for origin activation tend to replicate early, suggesting that early activation is the default state for origins[154]. However, in a plasmid construct, a DNA element immediately downstream of the *URA3* gene can advance origin activation time of an adjacent copy of *ARS1* through unknown means[155,156]. In light of these data, we hypothesize that there is

no common "default" activation time and that centromeres act as one of the early determinants for origin activation.

By measuring replication time locally as well as genome-wide in a strain with an ectopic centromere, we show that centromeres in *S. cerevisiae* act in *cis* to promote early activation on origins positioned as far as 19 kb away. Our results suggest that centromere-dependent early origin activation has a gradient effect such that the closer the origin is to the centromere the more profound is the timing effect on that origin. In addition, we show for the first time that early activation of centromere proximal origins is dependent on centromere function, suggesting that the ability of centromeres to establish proper kinetochore-to-microtubule attachments is important for regulating origin initiation.

2.2: Results

Proximity to centromeres influences time of origin activation

S. cerevisiae centromeres lie close to origins that undergo replication in early S-phase. To test the hypothesis that centromeres contribute to the early activation of adjacent origins, we obtained a strain in which the centromere on chromosome XIV (**Figure 2.1A**) had been relocated to a distal location on the left chromosomal arm[157] (**Figure 2.1B**). At this new location the centromere is positioned in a late replicating region near the potential origin *ARS1410*. We reasoned that if centromeres can influence origin activation time, then the origins that are closest to the moved centromeres would have the greatest chance of being affected.

The replication kinetics of three loci on chromosome XIV (*met2* or *MET2*, *ARS1410*, and *ARS1426*) in the rearranged and wild type (WT) strains were examined using a modified version of the Meselson-Stahl density transfer experiment[130]. Haploid cells grown in the presence of

dense ^{13}C and ^{15}N isotopes were arrested prior to S-phase. The cells were synchronously released into medium containing isotopically light carbon and nitrogen, and samples were collected at various times during the ensuing S-phase. Newly synthesized DNA was composed of light isotopes resulting in replicated DNA being hybrid or heavy-light (HL) in density whereas unreplicated DNA remained heavy-heavy (HH) in density. DNA was extracted from each cell sample, digested with restriction enzyme *Eco*RI, and subjected to ultracentrifugation in cesium chloride gradients. The gradients were then drip fractionated, and the kinetics with which the *Eco*RI restriction fragments containing each of the three loci of interest shifted from heavy to hybrid density were compared via slot blot analysis (**Figure 2.2A; also see Materials and Methods**).

WT cells entered S-phase by 40 minutes after release from alpha factor arrest, and most of the cells reached 2C DNA content between 120 and 140 minutes (**Figure 2.2B**). The percent replication of *MET2*, *ARS1410*, and *ARS1426* genomic fragments was calculated for each sample and plotted with respect to time (**Figure 2.2C**). The time of replication (T_{rep}) for each locus was calculated as the time it reached half maximal replication (see Materials and Methods). *ARS306*, one of the earliest known origins, and R11, a late replicating fragment on chromosome V, were used as timing standards for comparison. To facilitate comparison between cultures, these T_{rep} values were converted to replication indices[112] by assigning *ARS306* a replication index (RI) of 0 and R11 a RI of 1.0. Most other genomic loci have a RI between 0 and 1.0. The T_{rep} values for *MET2*, *ARS1410*, and *ARS1426* were then converted to RIs corresponding to the fraction of the *ARS306*-R11 interval elapsed when the T_{rep} for each locus was obtained. As previously observed, *MET2* replicated late in the WT strain (RI = 0.87), as did its nearest ARS, *ARS1410* (RI = 0.77), while *ARS1426* replicated early (RI =0.16) (**Figures 2.2C and 2.2D**).

Similar to WT cells, the cells with the relocated centromere entered S-phase by 40 minutes following release from alpha factor arrest and reached 2C DNA content between 120 and 140 minutes (**Figure 2.2B**). In contrast to the WT strain, both *met2* and *ARS1410* replicated early with respective RIs of 0.24 and 0.23 (**Figure 2.2C and 2.2D**). Consistent with *ARS1410* being the origin from which *met2* replicates, *ARS1410* maintained a slight timing advantage over *met2*. Meanwhile, *ARS1426* became later replicating (RI = 0.79) in the absence of its nearby centromere (**Figure 2.2C and 2.2D**). Similar results were obtained using an independent segregant (**see Figure S1**).

There are three explanations for the change in replication time of the *met2* locus: (1) the centromere advances the time of activation of *ARS1410*; (2) the centromere increases the efficiency (percentage of cells in which an origin is active) of *ARS1410*; and/or (3) insertion of the centromere created a new origin at the site. To distinguish among these possibilities we examined the replication of *ARS1410*, *ARS1426*, and *met2* or *MET2* by two-dimensional (2D) gel electrophoresis[158]. The presence of bubble-arcs indicated that *ARS1410* is indeed a functional origin in both the WT and rearranged strains (**Figure 2.3A**). Highly efficient origins display a more intense bubble-arc signal relative to the Y-arc[159]. Based on the similarity of bubble- to Y-arc ratios we conclude that the centromere has no obvious effect on the efficiency of *ARS1410* (**Figure 2.3A**). A similar result was obtained for *ARS1426* (**Figure 2.3A**) supporting the idea that the observed timing changes are not due to changes in efficiency of existing origins. No bubble-arc was observed for the *MET2* locus in the WT strain or the *met2* locus in the rearranged strain (**Figure 2.3B**), indicating that an origin was not created by insertion of the centromere into the *MET2* locus. Together these data suggest that the presence of a nearby centromere or some feature of its pericentric DNA induces early activation of origins.

Centromere function is required for early activation of nearby origins

DNA sequences that determine origin activation time have remained largely elusive. However, previous work has shown that sequences flanking a subset of origins on chromosome XIV delay the activation of those origins[112]. These sequences, coined "delay elements", can reside up to 14 kb from their affected origins. Therefore, it is possible that by integrating the centromere into the *MET2* locus in the rearranged strain, a delay element responsible for making *ARS1410* late activating was disrupted or pushed out of its effective range, thereby causing the origin to fire early. Although this scenario would explain the change in the replication times of *met2* and *ARS1410*, the observation that *ARS1426* became later replicating when the centromere was removed from its endogenous position argues in favor of the timing changes being a consequence of centromere proximity. Alternatively, it is conceivable that there is an uncharacterized sequence element, distinct from centromeric sequence, residing in pericentric DNA that is promoting early activation of nearby origins. The existence of such an element is formally possible as a cryptic sequence residing at the 3' end of the *URA3* gene was shown to advance the activation time of nearby origins on a plasmid and in an artificial chromosomal setting[155,156]. We tested these possibilities as described below.

At the sequence level, *S. cerevisiae* centromeres are composed of three essential elements: CDEI, CDEII, and CDEIII. CDEI and CDEIII are defined by essential consensus sequences whereas CDEII is a 78-86 bp AT-rich sequence that separates CDEI and CDEIII[142,152,153]. CDEIII has been found to be the element most important for centromere function as it is the binding site for essential inner kinetochore proteins, notably members of the CBF3 complex[142,153,160]. To ask if a functional centromere is required for early activation of

nearby origins we engineered a strain to have a non-functional centromere with a mutated CDEIII motif integrated at the *MET2* locus while the functional centromere remained in its endogenous position (**Figure 2.4A**). This strain was also subjected to flow cytometry (**Figure 2.4B**) and replication timing analysis as described above (**Figure 2.4C**). Unlike the dramatic replication timing change observed when we introduced a functional centromere, introducing the mutated centromere caused no replication timing change at *met2* and *ARS1410* (RI of 0.81 and 0.74, respectively; **Figure 2.4D**). Therefore, we conclude that centromere function is needed to effect a timing change on nearby origins. Furthermore, the late replication of this region was not due to inactivation of *ARS1410* through insertion of the mutated centromere as indicated by 2D gel analysis (**Figure 2.4E**). Together, these data demonstrate that functional centromeres actively advance the activation times of origins over a distance of 11.5 kb.

Centromeres determine the replication time of their local genomic environment

Upon finding that centromeres advance the activation time of origins to a distance of at least 11.5 kb, we sought to determine how far the centromere's effect extends along the chromosome. Raghuraman et al. presented statistical evidence that the regions of chromosomes within 25 kb of centromeres replicated earlier than the genome average, raising the possibility that centromeres could influence origin activation time over this larger distance[106]. Interestingly, at least one early replicating origin can be found within 12.8 kb of every centromere[147], suggesting that centromeres may be able to influence the activation time of origins over at least this distance. A recent study using an S-phase cyclin mutant[146] showed that early replicating domains that include a centromere can be well over 100 kb, implying that the range over which centromeres can regulate origin activation time might be quite broad.

To determine the range over which a centromere can influence replication time we performed a genome wide analysis of replication in the WT and rearranged strains. Cells were grown in dense medium (see Materials and Methods) and timed samples were collected following release into S-phase in light medium. To obtain a genome-wide view, the HH and HL DNAs from each timed sample were labeled with different fluorophores, cohybridized to microarray slides, and replication profiles were generated (**Figures 2.5A, S2.2, and S2.3**). Peak locations in the profile correspond to the locations of origins while the timed sample in which the peaks first appear gives an indication of the time at which the corresponding origins become active during S-phase[106].

Replication profiles generated for the WT strain (**Figure S2.2**) were consistent with previous studies[145]. Chromosome XIV displayed a strong peak at *ARS1426* at the earliest time (40 minutes) in the time course and a less well-defined peak at *ARS1410* (**Figure 2.5A**). Conversely, replication profiles in the rearranged strain displayed a shallow and late appearing peak at *ARS1426* while the peak at *ARS1410* appeared strong and early (**Figures 2.5A and S2.3**) confirming the observations made by slot blot analysis of individual restriction fragments. To directly compare replication profiles from the two strains, the percent replication values from the 40, 45, and 65 minute samples were normalized by conversion to Z-scores (see Materials and Methods) and superimposed on the same axes (**Figures 2.5B, S2.4, S2.5, and S2.6**).

Comparison of the Z-scores on chromosome XIV (**Figure 2.5B**) indicates that centromeres have a drastic influence over the activation times of their closest origins, suggesting that centromeres, mechanistically, operate locally. Z-score profiles for WT and rearranged chromosomes display a prominent early appearing peak centered about 19 kb to the left of the endogenous centromere. *ARS1425* is a potential origin located between *ARS1424* and the

endogenous centromere[65]. However, 2D gel analysis demonstrated that *ARS1424* is the origin likely responsible for this peak as *ARS1425* is not active in either strain (data not shown). Replication timing analysis of *ARS1424* by slot blot hybridization indicates that this origin is influenced by the centromere to a far lesser extent than *ARS1426* (WT RI = 0.11, Rearranged RI = 0.27). That this mild effect on *ARS1424* is not reflected in the Z-score overlays is likely due to the increased resolution of slot blot analysis compared to microarray analysis.

We then asked if the moved centromere influenced the activation times of origins located on other chromosomes. The replication kinetics of *ARS1*, a well-characterized origin on chromosome IV, were examined by slot blot hybridization and found to be similar between the two strains (WT RI = 0.68, Rearranged RI = 0.66) indicating that the activation time of *ARS1* is not influenced by centromere position on chromosome XIV. The remaining profiles were examined by Z-score comparisons looking for possible trans effects of centromere position. The profiles were strikingly similar (**Figures S2.4, S2.5, and S2.6**). Only one other location, corresponding to *ARS1531* on chromosome XV, showed a timing difference as drastic as that seen for *ARS1410* and *ARS1426* (**Figures 2.6A, S2.4, S2.5, S2.6**). A prominent peak of high percent replication corresponding to *ARS1531* was present in the WT strain while no peak was observed in the rearranged strain. However, further examination using 2D gel electrophoresis revealed that although *ARS1531* is not an active origin in the isolate of the rearranged strain used for microarray analysis it is active in the WT strain as well as in the rearranged independent segregant used in the slot blot analysis of timing (**Figure 2.6B**; for biological replicates kinetic data see **Figure S2.1**). These data suggest that the apparent timing change at this chromosome location is likely to be a consequence of a polymorphism affecting the ability of *ARS1531* to function as an origin rather than from any long-range effect of centromere function. Sequencing

of *ARS1531* revealed that this origin in YTP16 contains a mutation in the essential ARS consensus sequence (ACS)[65] converting it from ATATTTATATTTAGA to ATACTTATATTTAGA (**Figure 2.6B**). A change of the T to a C at this position has been shown to entirely abolish origin activity in *ARS307*[61]. *ARS1531* does not contain this substitution in either YTP12 or YTP15. To confirm that this basepair change is responsible for the observed difference in origin activity, *ARS1531* was cloned from YTP12 and YTP16 and tested for its ability to produce transformants in YTP12, YTP16, and YTP19. The "T" version of *ARS1531* resulted in the formation of many robust colonies in the three strains tested whereas the "C" version of *ARS1531* did not (data not shown).

2.3: Discussion

In this study we investigated the long-standing question of why centromeres replicate early in S-phase. We considered two possibilities: (1) that some component required for centromere function is also involved in early origin activation, or (2) that evolution has favored the migration of centromeres to early replicating regions. It has been hypothesized that early origin initiation in *S. cerevisiae* is the default state for origin timing[154], suggesting a more passive mechanism for early centromere replication. However, the observation that early centromere replication is conserved[129,139,140] in conjunction with the identification of a DNA element that is not associated with centromeres but capable of advancing origin activation time[155,156] indicates that establishment of early origin activation time is more complex than previously thought. Consistent with this idea, a recent study shows that the centromeres of *C. albicans* can alter the replication times of the loci in which they reside by allowing the formation of a *de novo* early firing origin[129]. Because the neocentromere in *C. albicans* creates a new origin, it is unclear

whether *C. albicans* centromeres also directly influence origin activation time. These results raised the question of whether the relocation of a centromere would have a similar effect in organisms with point centromeres. In this study, we took advantage of the well-characterized centromeres and origins in *S. cerevisiae* to effectively separate the centromere from origin function and address these questions.

We show that centromeres in *S. cerevisiae* are capable of advancing the replication time of genomic regions in which they reside by inducing early activation of their adjacent origins at distances of 11.5 and 6.8 kb (compare *ARS1410* and *ARS1426* in WT and rearranged strains, **Figures 2.2C and 2.2D**). We also show that early activation of *ARS1410* and *ARS1426* depends on centromere function. We find that centromere-mediated early origin activation requires an intact CDEIII region, suggesting that early origin activation is dependent on at least some portion of the DNA-protein complex normally formed at the centromere. Thus, centromeres and at least a subset of the kinetochore proteins they assemble participate as *cis*-acting regulatory elements of origin firing time. Furthermore, our 2D gel results indicate that centromeres do not affect origin efficiency, suggesting that the mechanisms responsible for centromere-mediated early origin activation are distinct from those that determine efficiency.

S. cerevisiae centromeres are known to reside in large early replicating portions of the genome spanning as much as 100 kb and containing multiple origins[106,107,145,146]. In light of the finding that centromeres regulate the activation times of their closest origins, we were interested in determining over what distance centromeres exert their effect. Our genome-wide replication timing analysis indicates that the centromere's influence on origin activation time is severely diminished at a distance of ~19 kb (**Figure 2.5B**) indicating that not all early replicating origins are under the centromere's influence. This result implies that there are at least two distinct

mechanisms by which origins can fire early. In contrast to what has been observed in *C. albicans*, we see no evidence that the centromeres of *S. cerevisiae* create new origins.

The genome is not randomly organized within the nucleus but particular genomic regions co-localize or cluster into functional foci during processes such as DNA replication[161,162]. In particular, centromeres in *S. cerevisiae* cluster throughout the cell cycle[163]. Therefore, it is plausible that centromeres, their neighboring origins, as well as other portions of the genome that interact with them, are clustered in G1 phase when timing decisions are made. This clustering could provide a way for centromeres to mediate replication time through a *trans*-acting mechanism. To determine if the mechanism responsible for centromere mediated early origin activation is capable of acting in *trans*, we examined the genome wide replication timing data for other timing changes occurring in the genome as a result of the repositioned centromere on chromosome XIV. We looked for differences in the Z-score profiles between matched S-phase samples, demanding that they persist over the course of S-phase (See **Figures S2.4, S2.5, S2.6**) to be considered significant.

Upon inspection, only one other location centered on *ARS1531* displayed a timing difference of at least the same magnitude as observed for *ARS1410* or *ARS1426* (compare **Figures 2.6A, S2.4, S2.5, S2.6**). 2D gel analysis indicates that this timing difference is not a result of relocating the centromere position on chromosome XIV and that it is likely due to a polymorphism affecting the ability of *ARS1531* to function in this particular isolate. Other than these three locations, the replication profiles are remarkably similar, suggesting that the mechanisms by which centromeres influence origin activation time are restricted to relatively limited adjacent regions. However, it is possible that centromeres contribute to smaller timing differences observed on other chromosomes such as that observed on chromosome XV at ~850

kb (**Figure 2.6A, S2.4, S2.5, S2.6**). Because the timing change at this location was only present in two of the three Z-scored samples, we did not consider it to be a significant timing difference.

Here we show direct evidence of a molecular link between the establishment of the kinetochore and replication initiation machinery. Although the mechanism of centromere-mediated early origin activation is unknown, we show that such a mechanism is dependent on at least some of the protein components associated with the kinetochore that require an intact CDEIII region. We also show that the mechanism is capable of affecting initiation time of origins that reside up to ~19 kb away from the centromere. We propose four possible models, which are not mutually exclusive (**Figure 2.7**): (1) The nuclear environment near the microtubule organizing center (MTOC) is particularly enriched in replication initiation factors; (2) The tension exerted by the microtubule is translated along the nearby DNA, altering its chromatin structure, thereby influencing the accessibility of the imbedded origins to initiation factors; (3) Proteins within the kinetochore directly (or indirectly) interact with initiation factors, recruiting them to nearby origins; and (4) The C-loop architecture of the pericentric chromatin, ensures the origins within the C-loop will be at the periphery of the chromatin mass and are therefore more exposed to initiation factors. Studies examining the concentration of replication initiation proteins near MTOCs have not been conducted; however, nuclear pore complexes are enriched near MTOCs in *S. cerevisiae*[164,165], the significance of which is unknown. While it is tempting to invoke localization to the nuclear periphery in the vicinity of MTOCs as a potential link between replication timing and centromeres, there is as yet no clear causal connection between nuclear localization and replication timing. Late origins tend to dwell at the nuclear periphery while early origins tend to be internal to the nucleus[166]. However, tethering an early origin to the nuclear periphery is not sufficient to alter its replication time[167] and conversely,

delocalizing telomeres from the nuclear periphery does not advance their replication time[168]. Although experiments that test for protein-protein, protein-DNA, and DNA-DNA interactions of kinetochore/centromeric components have been reported, we are not aware of any experiments that bear directly on the second model (tension-mediated promotion of early origin activation).

Kinetochores are multiprotein complexes composed of over 60 proteins[153,169]. Some of these proteins have been shown to interact both genetically and physically with proteins involved in regulating DNA replication[170-172]. However, as we are interested in molecular components that specify origin initiation time and these components have yet to be identified, it is not straightforward to determine which interacting protein candidates should be singled out for further study. A systematic way to parse through the various kinetochore components would be to determine if members of the inner-, mid-, or outer-kinetochore complexes are required for early origin initiation through the use of temperature sensitive alleles. Though this method poses it own set of challenges, it may prove to be fruitful in uncovering some of the mechanisms by which centromeres regulate origin activation time.

In vitro studies of protein-DNA interactions have shown that the CBF3 complex (composed of inner kinetochore DNA binding proteins) induces a 55^{o} bend in centromeric DNA[173]. In concordance with this finding, more recent work from the Bloom lab has shown that pericentric DNA adopts a particular tertiary structure coined the C-loop[174]. The C-loop is characterized by an intramolecular loop centered on the centromere resembling a hairpin of duplex DNA. C-loop formation is dependent on a subset of inner kinetochore proteins that physically bind centromere DNA, namely those that bind the CDEIII region[174,175]. Cohesin, which is enriched in pericentromeric regions[176-178], is thought to stabilize the C-loop following centromere replication[174]. Additionally, this study reports that the C-loop is present in G1-phase

during alpha factor arrest, suggesting that neither the presence of a sister chromatid nor pericentric cohesin is essential for the maintenance of the structure. Interestingly, the intramolecular interaction of the C-loop is reported to extend more than ~11.5 kb and less than 25 kb from the centromere, a distance similar to the findings reported in this study over which centromeres can regulate origin activation time. Anderson et al. hypothesized that proteins involved in the C-loop formation provide an essential geometry required for centromere position and accessibility for microtubule binding[175]. In light of these data, it is tempting to hypothesize that this geometry provides a favorable environment for early origin initiation by increasing the accessibility of nearby origins to replication initiation machinery.

Although the DNA binding components of the inner kinetochore are not well conserved, in human cells a DNA binding component of the inner kinetochore, CENP-B, has been shown to induce a ~59° bend in its target sequence[179], suggesting that regional centromeres also adopt a particular configuration. Therefore, this model can also explain how centromeres can vary greatly at the sequence level yet still affect timing of their associated origins.

2.4: Materials and Methods

Yeast strains: Strain and vector genotypes can be found in Table S2.1. Primer sequences can be found in Table S2.2. All *S. cerevisiae* strains used in this study were derived from a cross of CH1870 obtained from[157] and S288C. CH1870 is a mating type alpha strain derived from the S288C background. The *MET2* locus of this strain (located on chromosome XIV) was disrupted by an insertion of centromere VII sequence and *LEU2*. The endogenous centromere on chromosome XIV of CH1870 was also replaced with a *URA3* selectable marker[157].

The *ura3-52* allele in YTP13 (derived from a CH1870 to S288C cross) was restored to *URA3* by gene replacement, resulting in YTP15 (WT strain in this study, Table S2.1). Restoration was confirmed by Southern blot analysis. Independent segregants of the rearranged strains (YTP12 and YTP16) were obtained from separate tetrads from a CH1870 to S288C cross as met-, leu+, and ura+ progeny. Three different PCR reactions using primer pairs 71:72, 72:88, and 71:126 were used to confirm that these strains contained the desired inserted sequence at the *MET2* locus (data not shown). The desired insertion was further confirmed by Southern blot analysis on genomic DNA that was digested with XbaI (data not shown).

Mutant centromere at the met2 locus: YTP19 was constructed to have the same *met2::CEN7.LEU2* cassette as YTP12 and YTP16 save for a 3 bp mutation in CDEIII. The mutation of CDEIII in the centromere sequence integrated at *MET2* was made non-functional through site directed mutagenesis[180]. This strain was constructed as follows: Two halves of *met2*::*CEN7.LEU2* were PCR amplified from YTP12 using primer pairs 71:133, and 72:134. Primer pair 71:133 was used to amplify the 5' end of *met2* with primer 71 hybridizing upstream of the BsmI restriction site located in *met2* and primer 133 hybridizing downstream of the EcoRV restriction site located in the *LEU2* gene. Primer pair 72:134 was used to amplify the 3' end of *met2* with primer 72 hybridizing downstream of *met2* and primer 134 hybridizing upstream of the EcoRV restriction site located in the *LEU2* gene. The PCR products were sequentially digested with either BsmI or AatII followed by an EcoRV digest. The two fragments were then inserted into a pUC18 vector that contained KanMX and *ARS228* through a tri-molecular ligation reaction, creating plasmid pTP18.

The centromere on plasmid pTP18 was made non-functional, creating pTP19, through site directed mutagenesis using primers 147 and 148[180]. These primers were used to mutate the

Centromere DNA Element III (CDEIII) from TCCGAA to TCTAGA, introducing an XbaI site into the sequence. The most important bases for centromere function are the middle CG in TCCGAA[142,173]. The mutated centromere sequence on pTP19 was Sanger sequenced to ensure that only CGA had been changed. The lack of centromere activity in pTP19 was confirmed by a plasmid stability assay in which serially diluted cells were spotted to selective medium over 24 generations (data not shown).

The 4025 bp AatII/BsmI fragment of plasmid pTP19 containing *met2*::*cen7*.*LEU2* was transformed into YTP15[181,182]. Transformants were selected on synthetic dropout (SD) medium plates lacking leucine. Cells were then replica-plated to SD medium lacking methionine. Colonies that were prototrophic for leucine and auxotrophic for methionine were further screened similar to the rearranged strains mentioned earlier using PCR primer pairs 71:72 and 72:88 and through Southern blot analysis on XbaI digested genomic DNA.

Two-dimensional agarose gel electrophoresis: Origin activity was analyzed by standard 2D agarose gel electrophoresis techniques performed on total genomic DNA obtained from either asynchronous or synchronous S-phase cells[183-185]. Asynchronous samples were used for 2D gels conducted on *ARS1426*, *ARS1531*, and *met2::CEN7*. Synchronized samples were used for 2D gels conducted on *ARS1410* and *MET2*. For asynchronous samples, cells were collected in early log phase. For synchronized samples, cells were arrested in G1-phase with alpha factor (final concentration 3 μM) and released synchronously into S-phase by the addition of pronase (0.15 mg/mL). Cells were then collected every 2 minutes and pooled. Genomic DNA was harvested similar to asynchronous samples. In the first dimension, DNA was separated in 0.4% agarose for 18-20 hours at 1V/cm. The second dimension was run for 3-3.5 hours in 1.1% agarose containing Ethidium Bromide (0.3 μg/mL) at ~5-6 V/cm at 4°C.

Flow cytometry: Cells were harvested by mixing with 0.1% sodium azide in 0.2 M EDTA and then fixed with 70% ethanol. Flow cytometry was performed as previously described[186] upon staining cells with Sytox Green (Molecular Probes, Eugene, OR). The data were analyzed with CellQuest software (Becton-Dickinson, Franklin Lakes, NJ).

Density transfer experiments: Density transfer experiments were performed as described in[130] with slight modifications in cell synchronization and sample preparation for microarray analysis[106]. Cells were grown overnight at 23°C in a 5 mL culture of dense medium containing ^{13}C-glucose at 0.1% (w/v) and ^{15}N-ammonium sulfate at 0.01% (w/v). Cells were then diluted into a larger vessel in dense medium and allowed to reach an optical density of 0.16 (~$2x10^6$ cells/mL). Cells were arrested in G1 by incubating with 3 μM alpha factor until at least 95% of cells were unbudded based on microscopic analysis. Cells were then filtered and transferred to medium containing ^{12}C-glucose (2%), ^{14}N-ammonium sulfate (0.5%), and alpha factor. Cells were then synchronously released from the G1 arrest through the addition of pronase at a concentration of 0.15 mg/mL. Samples were collected every 10 minutes. Cell samples were treated with a mixture of 0.1% odium azide and 0.2 M EDTA then pelleted and frozen at -20°C. Genomic DNA was extracted from pelleted cells, digested with EcoRI, and the DNA fragments were separated based on density by ultracentrifugation in cesium chloride gradients. Gradients were drip fractionated and slot blotted to nylon membrane where they were hybridized with probes of interest. Unreplicated DNA is HH in density while replicated DNA is HL in density.

Replication kinetics and analysis: To construct replication kinetic curves, slot blots were hybridized to sequences of interest, and the percent replication in each timed sample was calculated as described in[106,130,146]. A detailed protocol can be found at http://fangman-brewer.genetics.washington.edu/density_transfer.html. The time at which the kinetic curves

reach half maximal defines the time of replication (T_{rep}) for each individual locus within the population of cycling cells. However, because not all G1 cells will enter or complete S-phase, T_{rep} for the probed regions in this study were calculated based on the plateau of the replication kinetic curve for a genomic probe consisting of EcoRI digested total genomic DNA from a strain lacking mitochondrial DNA.

To compare timing differences between strains, T_{rep} values were converted to replication indices in the following manner. First the T_{rep} values of two timing standards, *ARS306* (a known early replicating region) and R11 (a known late replicating region near *ARS501*), were obtained and assigned the values of 0 and 1, respectively. Next the T_{rep} values of regions of interest were assigned a replication index corresponding to the fraction of the *ARS306*-R11 interval elapsed at the time at which the T_{rep} for each locus was obtained. Replication index is calculated as $(T_{rep}(X) - T_{rep}(ARS306))/(T_{rep}(ARS306) - T_{rep}(R11))$ where X is the fragment of interest.

For microarray analysis, slot blots were hybridized with a genomic DNA probe to indentify fractions containing either HH or HL DNA. Once identified, the HH and HL fractions were separately pooled and differentially labeled with cyanine (Cy3 or Cy5) conjugated dUTP (Perkin Elmer)[145]. Timed samples were hybridized to high density Agilent yeast ChIP-to-chip 4x44K arrays according to the manufacturer's recommendations. The algorithms for analyzing microarray data are described in[147]. An 18 kb window was used for overall smoothing.

To facilitate direct comparison of replication profiles, percent replication values obtained from microarray analysis were converted to Z-score values. Z-score values were calculated using the following formula: $Z = (X-\mu)/\sigma$ where X = the percent replication value for a given probe, μ = the genomic average percent replication for a given sample, and σ = the standard

deviation of the distribution of X. The 55 minute samples were not used because the genomic percent replication values were not well matched between the two strains.

2.5: Notes

This chapter has been published: Pohl, T.J., Brewer, B.J., and Raghuraman, M.K. (2012). *PLoS Genetics*. 8(6): e1002677. doi:10.1371/journal.pgen.1002677. PMID: 22589733

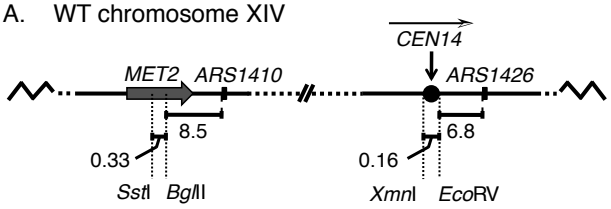
Supplemental information is located in Appendix A.

Acknowledgements

We thank Sue Biggins, Trisha Davis, members of the Brewer/Raghuraman and the Dunham laboratories for helpful discussions. We also thank Gina Alvino and Wenyi Feng for expert technical assistance and critical comments on the manuscript. This research was supported by the National Institute of General Medical Sciences grant (18926) to BJB and MKR and by the NIGMS training grant at the University of Washington to TJP.

FIGURE 2.1

A. WT chromosome XIV



B. Rearranged chromosome XIV

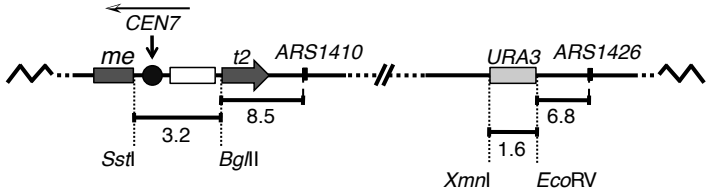


Figure 2.1: A schematic diagram of chromosome XIV in wild type (WT) and rearranged strains. (A) In the WT strain, the *Bgl*II restriction site in *MET2* is located 8.5 kb to the left of *ARS1410*. Centromere XIV resides in its endogenous position located 6.8 kb to the left of *ARS1426*. **(B)** In the rearranged strain the endogenous centromere was replaced with a *URA3* selectable marker while a functional centromere was integrated along with *LEU2* (open box) into the *MET2* locus such that the centromere was positioned ~11.5 kb from *ARS1410*. The white and black arrowhead above each centromere indicates the direction of the centromere DNA elements CDEI, CDEII, CDEIII.

FIGURE 2.2

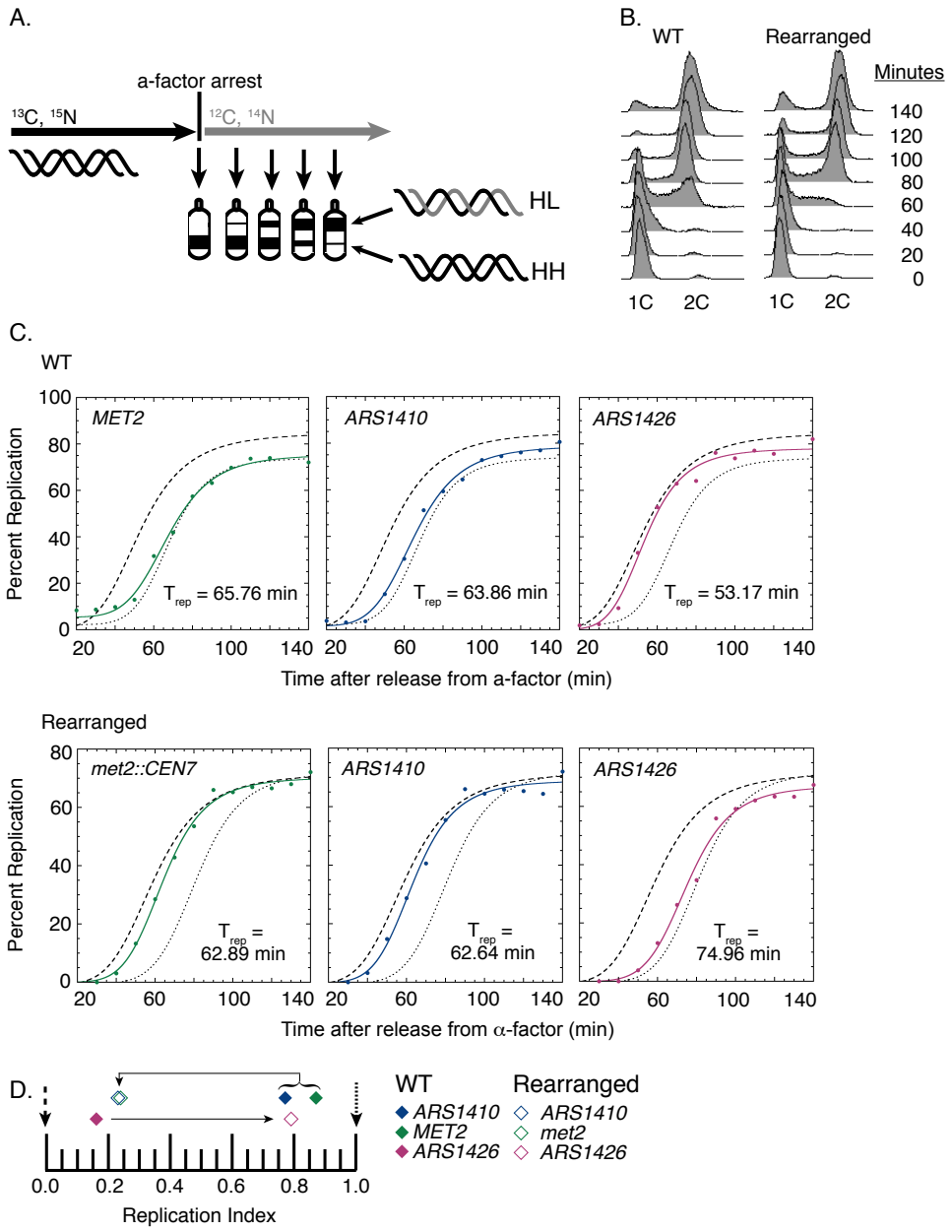

A.
a-factor arrest
13C, 15N
12C, 14N
HL
HH
B.
WT
Rearranged
Minutes
140
120
100
80
60
40
20
0
1C
2C
1C
2C
C.
WT
MET2
ARS1410
ARS1426
Percent Replication
Trep = 65.76 min
Trep = 63.86 min
Trep = 53.17 min
Time after release from a-factor (min)
Rearranged
met2::CEN7
ARS1410
ARS1426
Trep = 62.89 min
Trep = 62.64 min
Trep = 74.96 min
Time after release from α-factor (min)
D.
0.0
0.2
0.4
0.6
0.8
1.0
Replication Index
WT
ARS1410
MET2
ARS1426
Rearranged
ARS1410
met2
ARS1426

Figure 2.2: Replication time of native and relocated centromeres on chromosome XIV. (A) Cartoon depiction of experimental setup. Cells were grown in medium containing heavy carbon (^{13}C) and nitrogen (^{15}N) isotopes. Upon genome saturation with the heavy isotopes, cells were arrested by the addition of alpha factor and released synchronously in medium containing light carbon (^{12}C) and nitrogen (^{14}N) isotopes. The cells were then collected over the next 140 minutes and their DNA was extracted, digested with *Eco*RI, and separated via ultra centrifugation in cesium chloride gradients such that unreplicated DNA resides lower in the gradient than newly replicated DNA. DNA samples were then collected and analyzed through drip fractionation. **(B)** S-phase progression of WT (left) and rearranged (right) cells as measured by flow cytometry. Cells from both strains entered S-phase by 40 minutes and achieved 2C DNA content by 140 minutes as indicated by the peak shift from 1C to 2C DNA content. **(C)** Replication kinetic curves for *met2* or *MET2*, *ARS1410*, and *ARS1426* in WT (top panel) and rearranged (bottom panel) cells. The kinetic curves for *ARS306* and *R11* are shown as dashed and dotted lines, respectively. T_{rep} is the time of half-maximal replication for each locus (see Materials and Methods). **(D)** Replication indices for *met2* or *MET2* (green), *ARS1410* (blue), and *ARS1426* (magenta) in WT (solid diamonds) and rearranged (empty diamonds) strains. *ARS306* (black dashed arrow) and *R11* (black dotted arrow) were used as early and late timing standards, respectively. In the WT strain, *MET2*, *ARS1410*, and *ARS1426* had replication indices of 0.87, 0.77, and 0.16 respectively. In the rearranged strain, *met2*, *ARS1410*, and *ARS1426* had replication indices of 0.24, 0.23, and 0.79, respectively. Direction of the black arrows indicates the direction of the shift in replication index for each locus between WT and rearranged strains.

FIGURE 2.3

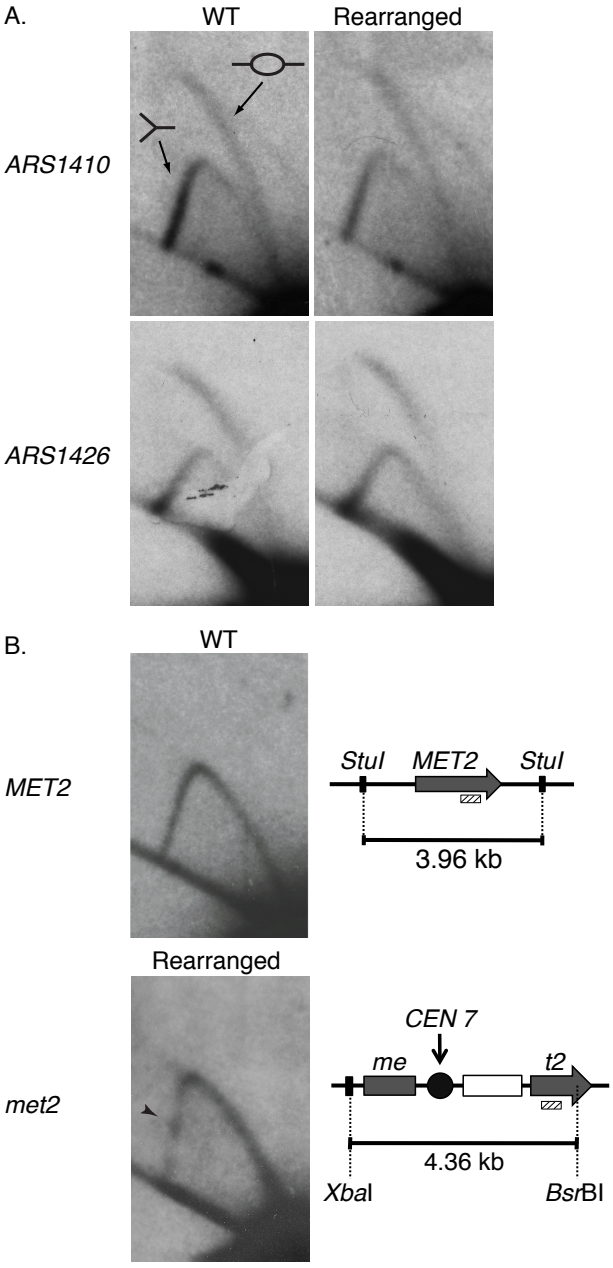

A.
WT
Rearranged
ARS1410
ARS1426
B.
WT
MET2
StuI
MET2
StuI
3.96 kb
Rearranged
met2
CEN 7
me
t2
4.36 kb
XbaI
BsrBI

Figure 2.3: 2D gel analysis of *ARS1410*, *ARS1426*, and *MET2* and *met2*. DNA fragments containing a functional origin are detected as a bubble arc (depicted by the bubble fragment) while fragments that are passively replicated are detected as a Y-arc (depicted as a Y shaped fragment). **(A)** *ARS1410* and *ARS1426* are functional origins in the WT and rearranged strains. **(B)** 2D gel analysis of the *MET2* or *met2* locus in the WT and the rearranged strains. WT DNA was digested with the restriction enzyme StuI giving a fragment of 3.96 kb centered on *MET2* (grey arrow). Rearranged DNA was double digested with XbaI and BsrBI resulting in a 4.36 kb fragment harboring most of *met2*, the integrated centromere (black circle), and the *LEU2* marker (white rectangle). The absence of a bubble arc when probed for the 3' end of *MET2* and *met2* (hashed rectangle) indicates that an origin is not present on either DNA fragment. The centromere in the rearranged construct was detected as a pause site (black arrowhead) visualized as a dot of relatively increased intensity on the descending Y-arc.

FIGURE 2.4

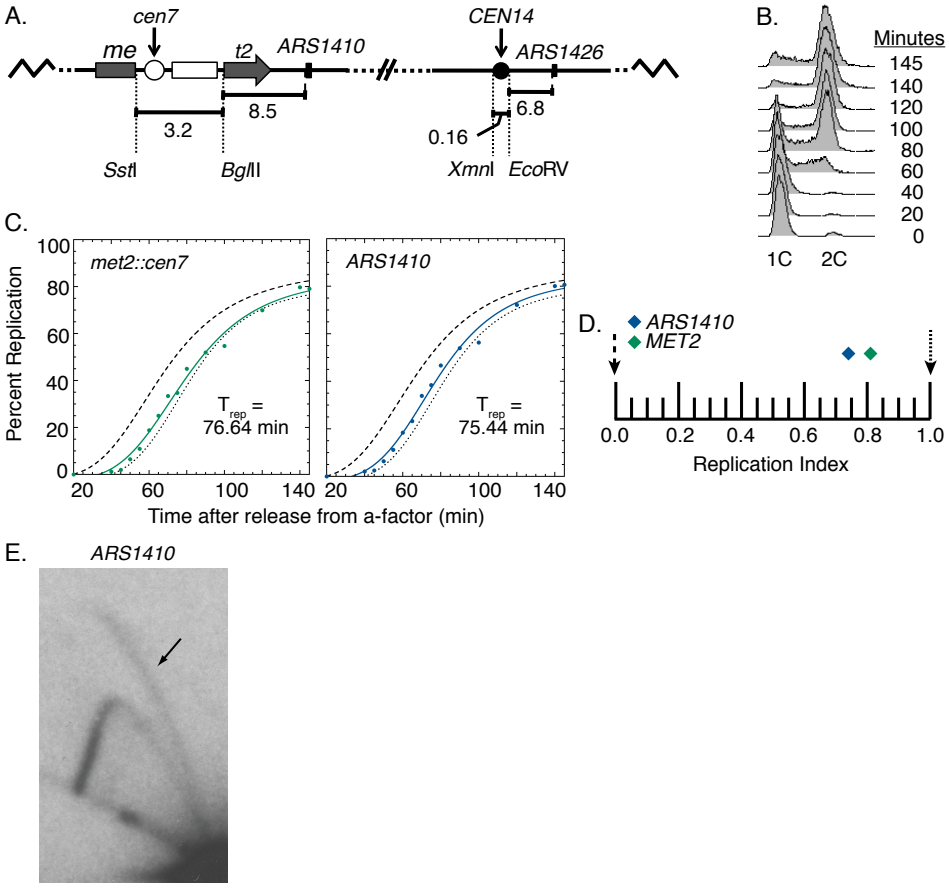

A.
cen7
me
t2
ARS1410
CEN14
ARS1426
3.2
8.5
0.16
6.8
SstI
BglII
XmnI
EcoRV
B.
Minutes
145
140
120
100
80
60
40
20
0
1C
2C
C.
met2::cen7
ARS1410
Percent Replication
Trep = 76.64 min
Trep = 75.44 min
Time after release from a-factor (min)
D.
ARS1410
MET2
0.0
0.2
0.4
0.6
0.8
1.0
Replication Index
E.
ARS1410

Figure 2.4: Replication time of point mutated centromere on chromosome XIV. (A) Cartoon depiction of chromosome XIV with a non-functional centromere (white circle) integrated at *MET2*. Chromosome XIV of WT cells was modified such that *MET2* was disrupted with the same sequence used to disrupt *MET2* in the rearranged strain (see Figure 2.1) except that the centromere was made inactive by mutating the essential CDEIII domain. This chromosome is maintained through its wild type centromere at the endogenous location (black circle). **(B)** Flow cytometry of cells with a non-functional centromere in the *MET2* locus. Similar to the WT strain (see Figure 2.2B), cells from this cell line enter S-phase by 40 minutes and achieved 2C DNA content by 140 minutes. **(C)** Replication kinetic curves for *met2*::*cen7* and *ARS1410*. As observed in the WT strain, the replication curves for *met2*::*cen7* (green) and *ARS1410* (blue) are positioned more closely to that of R11 (dotted line) than *ARS306* (dashed line) (compare to Figure 2.2C). **(D)** Replication indices for *met2*:*cen7* (green diamond) and *ARS1410* (blue diamond). RIs of *ARS306* and R11 are indicated by black dashed and dotted arrows, respectively. *met2*::*cen7* and *ARS1410* had RIs of 0.81 and 0.74, respectively. **(E)** 2D gel analysis of *ARS1410*. Presence of a bubble arc (black arrow) for *ARS1410* in the strain in which the non-functional centromere was integrated at *MET2* (compare with Figure 2.1B) indicates that *ARS1410* is a functional origin in this strain.

FIGURE 2.5

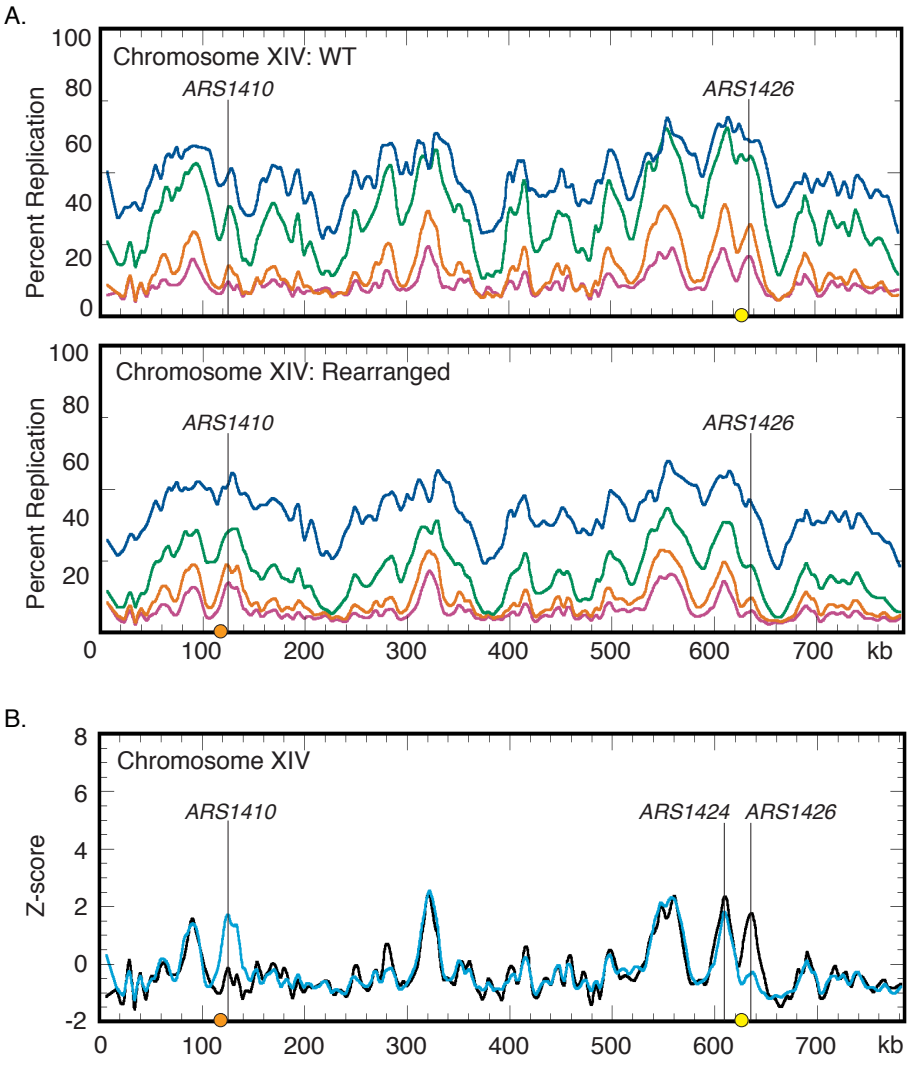

A.
Chromosome XIV: WT
ARS1410
ARS1426
Percent Replication
100
80
60
40
20
0
Chromosome XIV: Rearranged
ARS1410
ARS1426
Percent Replication
100
80
60
40
20
0
100
200
300
400
500
600
700
kb
B.
Chromosome XIV
ARS1410
ARS1424
ARS1426
Z-score
8
6
4
2
0
-2
0
100
200
300
400
500
600
700
kb

Figure 2.5: Replication dynamics for chromosome XIV in WT and rearranged strains. (A) Replication kinetic profiles of chromosome XIV in WT (top) and rearranged (bottom) strains. Percent replication was monitored across chromosome XIV at 40 (magenta), 45 (orange), 55 (green), and 65 (blue) minutes following release from alpha factor arrest. When the native centromere (yellow circle) is present near *ARS1426*, a prominent peak is seen in the 40 and 45 minute samples. In this strain, the peak at *ARS1410* is shallow in the 40 and 45 minute samples. When the centromere is repositioned (orange circle) near *ARS1410* in the rearranged strain, both the time of appearance and the prominence of the peaks at *ARS1410* and *ARS1426* are inverted with respect to the WT strain. See Figures S2.2 and S2.3 for all chromosomes. **(B)** Z-score plots of chromosome XIV in WT (black) and rearranged (blue) strains. Replication kinetic profiles from the 40 minute sample were normalized by converting percent replication values to Z-score values (see Materials and Methods). Genomic loci corresponding to *ARS1410* and *ARS1426* show significant differences in Z-scores. *ARS1424* is the next closest active origin to the endogenous centromere residing ~19 kb to the left. See Figures S2.4, S2.5, S2.6 and data sets S2.3 and S2.4 for all chromosomes and the 45- and 65- minute samples.

FIGURE 2.6

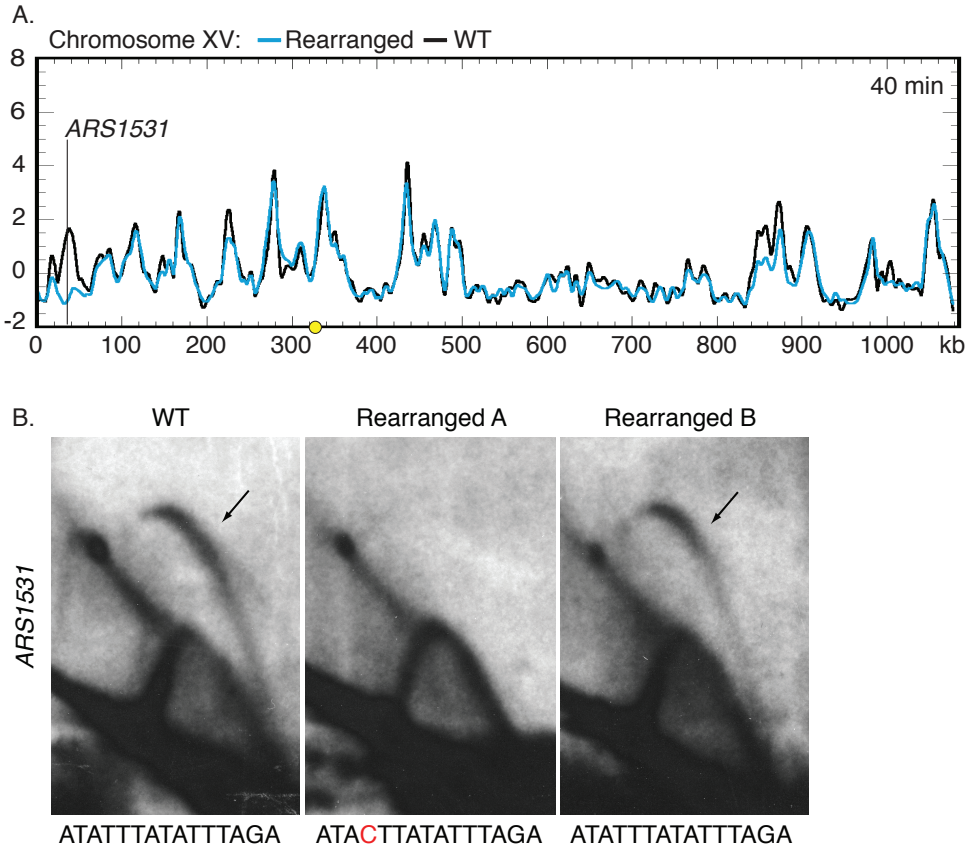

A.
Chromosome XV: Rearranged WT
40 min
ARS1531
8
6
4
2
0
-2
0
100
200
300
400
500
600
700
800
900
1000
kb
B.
WT
Rearranged A
Rearranged B
ARS1531
ATATTTATATTTAGA
ATACTTATATTTAGA
ATATTTATATTTAGA

Figure 2.6: Z-score and 2D gel analysis of *ARS1531*. (A) Z-score plot of chromosome XV in WT (black) and rearranged (blue) strains. *ARS1531* displayed a difference of Z-score values at least as large as that seen for *ARS1410* and *ARS1426* (see Figure 2.5B). See Figure S2.4 for all chromosomes. **(B)** 2D gel analysis of *ARS1531* in the WT (left), the rearranged strain used in microarray analysis (middle), and the rearranged strain used for prior slot blot analysis (right). DNA from all three strains was digested with NcoI and BglII to give a 3.18 kb fragment harboring *ARS1531* and then subjected to 2D gel analysis. The presence of a bubble arc in the WT (black arrow) indicates that *ARS1531* is a functional origin in this strain. The presence of a bubble arc in one of the two rearranged strains confirms that the absence of origin activity in rearranged A (used in microarray analysis) is not due to relocation of the centromere on chromosome XIV. Below each 2D gel image is the sequence for the WT or mutant (red) ACS.

FIGURE 2.7

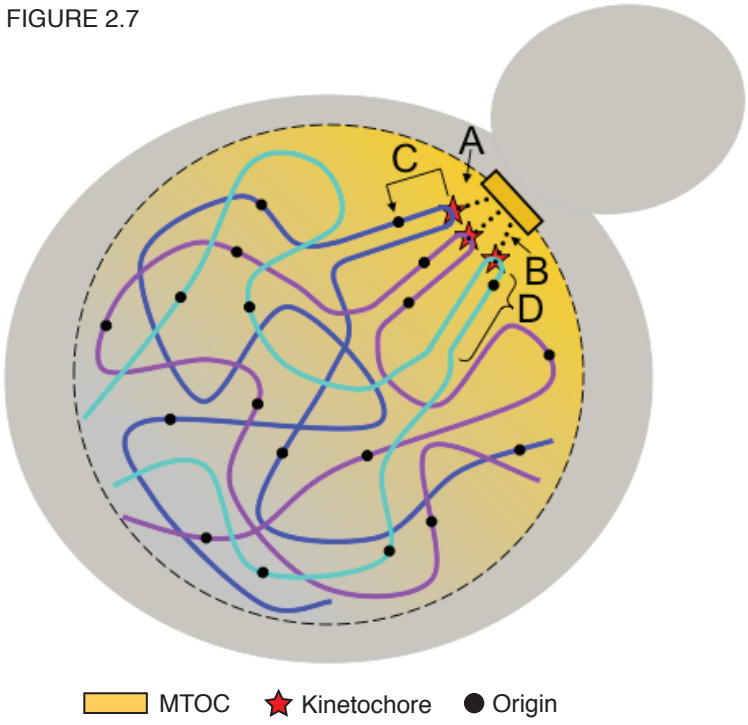


Figure 2.7: Models for centromere-mediated early origin activation.
(A) Kinetochore/microtubule interaction orients the centromere and pericentric DNA near the microtubule organizing center (MTOC) where there is an enrichment of replication initiation factors. **(B)** Tension exerted by the kinetochore/microtubule interaction induces an altered chromatin structure of pericentric DNA that provides accessibility of embedded origins to initiation factors. **(C)** Kinetochore proteins interact directly (or indirectly) with origin initiation factors recruiting them to nearby origins. **(D)** The organization of pericentric DNA into the C-loop orients origins within the C-loop to the periphery of the chromatin mass increasing their accessibility to initiation factors.

Chapter 3

Investigating the Significance of Early Centromere Replication in *Saccharomyces cerevisiae*

3.1: Introduction

Early centromere replication is a common feature of the replication landscape[55,106,129,130,139,140,187], suggesting that centromeres regulate the activation time of at least a subset of origins. In the previous chapter we confirmed this hypothesis by reporting that *S. cerevisiae* centromeres confer early activation on nearby origins[188]. Two recent findings in yeast add interesting perspectives on the centromere effect on origin activation. First, the lack of forkhead proteins, Fkh1 and Fkh2, results in delayed activation of a large number of normally early-firing origins, but centromere-proximal origins remain early-firing[189]. Second, while many origins show delayed activation in meiotic S-phase compared to mitotic S-phase, centromere-proximal origins do not show such a delay[190]. Both findings are consistent with the idea that some aspect of centromere function directly influences the activation of centromere-proximal origins through a mechanism that is independent of other, more global controls.

The discovery that budding yeast centromeres promote their own early replication in conjunction with the conservation of early centromere replication suggests that early replication of centromeres may be important for cell viability. Centromeres are the target site for the establishment of a large multi-protein complex known as the kinetochore (reviewed in[153]). Kinetochores not only act as the medium by which microtubules attach to the chromosome but also regulate chromosome segregation. In *S. cerevisiae*, kinetochores transiently disassemble upon centromere replication[191]. Presumably, a delay in centromere replication would provide

less time for subsequent kinetochore formation to occur prior to the onset of cell division. In light of the aforementioned data, we hypothesize that early centromere replication promotes the fidelity of chromosome segregation and decreases the potential for chromosome loss, aneuploidy, and other genomic rearrangements. We further speculate that there is selective pressure for cells to replicate their centromeres early. Here we examine these hypotheses through fluctuation analysis, competitive and long-term growth assays.

3.2: Results

Cells harboring a late replicating centromere are viable

The importance of early centromere replication is debatable, largely because the effects of having a late replicating centromere have not been specifically tested. The finding that centromeres mechanistically ensure their own early replication[188] implies that late replication of centromeres would be detrimental. To test whether *S. cerevisiae* can tolerate having a late replicating centromere, we replaced the origins that flank centromere V (*ARS510 and ARS511*) with selectable markers providing resistance to G418 (*KanMX*) and Nourseothricin (*clonNAT*^{R}), respectively. Two-dimensional (2D) gel electrophoresis[158] was used to confirm loss of origin activity upon replacement of *ARS510* and *ARS511* (**Figure 3.1A**). Cells lacking these origins appeared healthy and displayed no noticeable defects in colony morphology or their ability to transition through S-phase (**Figure 3.1B**) indicating that early replication of this centromere is not essential.

Alternatively, it is possible that in an attempt to maintain early replication of centromere V, either dormant or novel origins became active in the pericentric regions. To determine if new origins had become active, we examined the replication kinetics of centromere V using density

shift experiments coupled with microarray analysis (**Figure 3.2, S3.1, S3.2; also see Materials and Methods**). As expected, the pericentric region of chromosome V displayed strong early appearing peaks for *ARS510* and *ARS511* (**Figure 3.2A**). In cells lacking *ARS510* and *ARS511*, the genomic locations corresponding to these origins were replicated later than in WT cells despite the presence of nearby weak, residual replication peaks. To facilitate direct comparison of replication profiles, the percent replication values from the 15-minute samples were normalized by conversion to Z-scores (see Materials and Methods) and superimposed on the same axis (**Figure 3.2B**). From this comparison it is clear that centromere V was replicated drastically later in cells lacking *ARS510* and *ARS511* than in control cells. Thus we conclude that *S. cerevisiae* can tolerate a late replicating centromere. Furthermore, these data suggest that a majority of cells lacking *ARS510* and *ARS511* duplicate centromere V using replication forks that initiate from origins located over 40kb away.

As indicated above, low-level replication peaks were detected near the endogenous locations of *ARS510* and *ARS511*, the one near *ARS510* being the more prominent. We suspect that these are false peaks due to nonspecific hybridization during microarray analysis. However, it is possible that this signal is due to weak origin activity in these locations, as they have not been further analyzed.

Late centromere replication increases chromosome loss in the absence of the spindle checkpoint

The kinetochore is dissociated upon, and reestablished soon after, replication of the centromere[191]. Based on this mechanism it is logical that kinetochores will have less time to be properly reestablished on centromeres that are replicated late in S-phase. To test the hypothesis

that late centromere replication leads to an increased chromosome loss rate, we performed fluctuation analysis on a diploid that is heterozygous for centromere V replication time (**Figure 3.3**). In this diploid one copy of chromosome V is derived from the late centromere strain described in the previous section. The other copy is WT for *ARS510* and *ARS511*, and therefore replicates centromere V early in S phase (**Figure 3.3A**). On this chromosome, *CAN1* and *URA3* were replaced with the *can1-100* and *ura3-52* alleles. To create a control diploid with early replication of both copies of centromere V, we modified the *CAN1*, *URA3* chromosome by integrating the *KanMX* and *clonNAT*^{R} markers upstream of *ARS510* and downstream of *ARS511*, respectively. The control diploid has the same marker arrangement as the heterozygous diploid, but is homozygous for WT centromere V replication time (**Figure 3.3B**).

Cells were grown overnight at 23°C in medium that maintained selection (complete medium lacking uracil) for the chromosome harboring late centromere V. Fluctuation analysis was performed on cultures that were diluted into 96 well plates and grown to saturation under relaxed selection (complete medium; see Materials and Methods). The cultures in each well were then spotted to plates containing 5FOA and Canavanine. Each spot was then scored for the presence of colonies that were ≥0.5 mm in diameter (**Table 3.1**) from which a mutation (loss of *URA3* and *CAN1*) rate was calculated. Control cells (homozygous early replicating centromere V) had a mutation rate of 2.43×10^{-6} per cell per division (**Figure 3.3C, maroon bar**). Independent isolates (A and B) that were heterozygous for centromere V replication time displayed mutation rates similar to the control, 2.83×10^{-6} and 1.99×10^{-6} (**Figure 3.3C, red bars**). Because *URA3* and *CAN1* are substantially separated (82.7 kb), chromosome loss was thought to be the major event contributing to 5FOA + Canavanine resistance. To distinguish chromosome loss from other events such as break-induced replication (BIR), homologous recombination and

segregation, and point mutations, a single colony from each spotted culture was randomly selected and patched to 5FOA + Canavanine, then replica-plated to medium containing G418 or clonNAT (**Table 3.1**). Of the 126 control (homozygous early replicating centromere V) colonies patched to 5FOA + Canavanine, 114 were sensitive to G418 while 110 were sensitive to clonNAT indicating that a majority of 5FOA + Canavanine resistant colonies arose through loss of the entire chromosome. Similar results were obtained with cells that are heterozygous for centromere V replication time (**Table 3.1**); isolate A had 115 colonies that were sensitive to G418 and 111 that were sensitive to clonNAT whereas isolate B had 105 and 103 colonies that were respectively sensitive to G418 and clonNAT. All G418 resistant colonies were also resistant to clonNAT suggesting that these events arose through BIR that occurred to the left of *KanMX*. However, the rare possibility that both *URA3* and *CAN1* have suffered point mutations in these cells cannot be ruled out, as our assay cannot distinguish these events. Very few 5FOA + Canavanine resistant colonies were both sensitive to G418 and resistant to clonNAT (**Table 3.1**). These colonies could result from BIR, occurring to the left of $clonNAT^R$, and through point mutation of *KanMX*. These data indicate that having a late centromere does not increase chromosome instability.

Because the spindle checkpoint monitors deficient kinetochore-microtubule attachment, we reasoned that any ill effects that a late replicating centromere would impose on chromosome segregation might be masked by the presence of the checkpoint. Therefore, we performed fluctuation analysis on diploids that are identical to those used above except that they are deficient for *MAD2*, a vital component of the spindle checkpoint (**Figure 3.3C**). *mad2Δ* control cells had a slightly elevated mutation/loss rate (4.43×10^{-6}) compared to *MAD2* controls (2.43×10^{-6}). Interestingly, *mad2Δ* cells that had a late replicating centromere displayed significantly

higher mutation rates ($1.3x10^{-4}$). Together these data show that early replicating centromeres help improve chromosome stability in the absence of the spindle checkpoint.

Importance of early centromere replication on long-term growth

Because early centromere replication is conserved, we were curious to discern if having a late replicating centromere would impart a long-term growth disadvantage to the cell. To test this idea, we competed haploids that contained a late replicating centromere against a haploid WT strain that was marked with GFP in cultures that were inoculated with roughly a 50-50 mixture of each strain (**Figures 3.4**). The mixed cultures were kept in continuous log phase growth for up to 90 hours while the percentage of GFP cells within the population was determined via flow cytometry at intervals throughout the time course. If both GFP+ and GFP- cells grow at the same rate, then the population will remain at or near its initial ratio (roughly 50% GFP+ and 50% GFP-). However, the cell type that grows the fastest will eventually sweep the population.

Although the GFP+ cells displayed a growth advantage over the GFP- WT cells (**Figure 3.4A, maroon graph**), surprisingly no growth advantage was observed over cells that had a late replicating centromere V (**Figure 3.4A, red graphs**). This trend was reproducible with independent biological isolates of both the GFP+ and GFP- WT cells (**Figure S3.3A**). These data suggest that having a late replicating centromere confers a growth advantage to the cell. However, when we competed cells that had an early or late replicating centromere IX (obtained from Tomoyuki Tanaka and Conrad Nieduszynski[192]) against the same GFP+ control we detected no difference in growth rate (**Figure 3.4B, dark and light purple graphs,**

respectively). Together these data suggest that the replication time of centromere V is important for regulating cell growth, whereas the replication time of centromere IX is not.

Because a late replicating centromere coupled with a defective spindle checkpoint results in a higher mutation rate, we predicted that cells with these characteristics would be at a growth disadvantage relative to their early centromere counterparts. Similar to what we observed before, the *mad2Δ* GFP+ strain outgrew the GFP- version (**Figure 3.4C, dark blue graph**). However, the reciprocal occurred when *mad2Δ* was coupled with late replicating centromere V (**Figure 3.4C, light blue graphs**). The effect of late replicating centromere IX in combination with *mad2Δ* has not been tested in our hands. As expected, *MAD2* cells out competed *mad2Δ* cells (**Figure S3.3B**). From these data we conclude that the presence of a late replicating centromere V at least partially rescues the growth defect observed in the absence of the spindle checkpoint.

3.3: Discussion/Future Direction

This study is an ongoing attempt to determine the consequences of having a late replicating centromere. Here we have shown that although early replication of centromeres is not crucial for cell survival, this phenomenon helps to mediate proper chromosome segregation primarily in the absence of the spindle checkpoint. Consistent with our data, a study published earlier this year showed that cells with a late replicating centromere IX had an increased frequency of chromosome loss when present in a *mad2Δ* background[192]. However, the authors also reported a slight increase in chromosome IX loss in a *MAD2* background. This increase was absent in our analysis of chromosome V. The discrepancy between the two data sets could be technical or biological. Because Natsume et al. measured loss frequency as opposed to loss rate,

it is possible that their data are skewed by events that occurred either colonially during culture growth or after plating. Although these caveats exist, it is feasible that the differences between these two studies are specific to the chromosome being analyzed. Furthermore, the degree of replication delays of each centromere could also influence the results. It is plausible that centromere IX replicated later in S-phase than did our centromere V. This possibility is consistent with the idea that the later a centromere is duplicated the less time there is to establish a proper kinetochore.

Natsume et al. speculate that loss of one chromosome IX is followed by endo-reduplication of the other chromosome IX[192]. We are currently in the process of testing this hypothesis in our strains by determining the percentage of fluctuation analysis products (5FOA + Canavanine resistant, KanMX/Clonat sensitive) that have lost chromosome V. Chromosome copy number will be detected using pulse field gel electrophoresis followed by Southern blot analysis.

In addition to measuring chromosome loss, we were interested in determining if cells that have a late replicating centromere would be at a growth disadvantage relative to WT cells. We tested this possibility by competing WT cells against cells that contained either a late replicating centromere V or IX in continuously growing cultures for up to 90 hours (~30 generations). Additional competitions focusing on the significance of early replicating centromere V were conducted in the absence of the spindle checkpoint. Based on preliminary data we made the following two conclusions: (1) the replication time of centromere V is important for regulating cell growth whereas the replication time of centromere IX is not and (2) the presence of a late replicating centromere V at least partially rescues the growth defect observed in the absence of the spindle checkpoint.

It is possible that the replication times of different centromeres have variable effects on cell growth. This notion is exciting because it may have implications on how chromosomes are arranged within the nucleus. Alternatively, it is possible that changes in pericentric replication time have some bearing on the expression of genes within those regions. In this scenario, the time at which the centromere replicates does not affect growth rate but instead growth rate is influenced by the time at which the flanking DNA is replicated. However, it should be noted that no correlation between transcription and replication time has been found in *S. cerevisiae*[106].

Although the conclusions made from the growth competitions are enticing, there are caveats that need to be remedied and controls that should be performed before the conclusions can be made with high confidence. For example, it is possible that one or more of the strains have additional mutations that are affecting their growth rate. The strains in which centromere V was examined are isogenic with each other and of the A364A background. The strains used to analyze centromere IX are also isogenic with each other but of the W303 background. The GFP strain that was used as a reference was common to all competitions and of the A364A background. In this strain, GFP was obtained from S288C through an initial mating to A364A. Six subsequent backcrosses to A364A were performed. Either more backcrosses should be conducted or the GFP should be integrated to A364A and perhaps W303 through transformation to ensure isogenicity.

One key control that would add weight to this study would be to reintroduce an origin near late centromere V and test for rescue of the growth defects observed in figures 3.4A and 3.4C. If the growth differences presented in these figures are due to having a late replicating centromere then restoring that late centromere to its early status should revert these growth rates back to WT levels.

In addition to the experiments presented in this chapter, we are interested in determining if a late replicating centromere, in the presence and absence of the spindle checkpoint, will evolve to become early replicating over time. Currently, the details of the assay have not been completely established, as it will be challenging to determine when a timing change has occurred. Nevertheless, there are a number of ways in which the cell could accomplish such a task. First, the late centromere could be reverted to early replication status by activating dormant origins or through the creation of novel origins in the pericentric regions. Second, the rate at which replication forks travel could increase thus ensuring that the centromere is replicated in a timely manner. Third, the genome can be rearranged such that early activating origins are in close proximity to the centromere.

3.4: Materials and Methods

Yeast strains: Strains and their genotypes can be found in Table S3.1. Primers are listed in Table S3.2. All *S. cerevisiae* strains used in this study were derived from either A364A or W303. Strains in which chromosome V was studied are of the A364A background. To create a strain that replicates centromere V in late S-phase (YHLTP3 and YHLTP4), *ARS510* and *ARS511* were replaced with *KanMX* and *cloNAT*^{R}, respectively. Replacement of *ARS510* with *KanMX* was confirmed by three different PCR reactions using primer pairs 217:219, 217:221, and 219:220 (data not shown). Replacement of *ARS511* with *clonNAT*^{R} was confirmed using two PCR reactions amplifying the 5' and 3' insertion junctions, primer pairs 229:238, 230:237 (data not shown). Integration of *clonNAT*^{R} was also shown by Southern blot (data not shown). Loss of *ARS510* and *ARS511* activity was confirmed through 2D-gel analysis (**Figure 3.1A**).

The *MAD2* diploids (YTP40, YTP41, and YTP42) that were used in our fluctuation analysis were created by mating strains YHLTP3, YHLTP4, and YSL1_c with SL14-3α. *KanMX* and *clonNAT*^{R} were integrated upstream and downstream of *ARS510* and *ARS511*, respectively, and confirmed via Southern blot (data not shown). The *mad2*Δ diploids used in fluctuation analysis (YSL2, YSL3, and YSL4) were constructed by replacing *MAD2* with the hygromycin B resistance gene in YSL1_c (YSL1_c_*mad2Δ*), YHLTP3 (YHLTP3_*mad2Δ*), YHLTP4 (YHLTP4_*mad2Δ*), and SL14-3α (SL14-3α_*mad2Δ*). YSL1_c_*mad2Δ*, YHLTP3_*mad2Δ*, and YHLTP4_*mad2Δ* were then mated to SL14-3α_*mad2Δ*.

For the competition experiments, GFP was placed into YSL1_c and YSL1_c_*mad2Δ* by backcrossing our A364A strain to an S288C strain (obtained from Dunham lab) that has GFP integrated at the *HO* locus. GFP is constitutively expressed in this strain. The backcross was performed seven times. These strains were common to all competitions (YSL1_c_GFP for the *MAD2* competitions and YSL1_c_*mad2Δ*_GFP for the *mad2Δ* competitions) and used to track percent GFP.

The competition experiments focusing on the growth of early and late replicating chromosome V in a *MAD2* background (**Figure 3.4A**) were performed by growing each haploid strain (YSL1_c, YHLTP3, YHLTP4, YSL1_c_GFP) to log phase. Once in log phase, each GFP- strain was mixed at a 50-50 ratio with YSL1_c_GFP. The mixed cultures were kept in log phase for 90 hours. Periodically samples were taken and measured for percent GFP using a flow cytometer.

The competition experiments performed on the *mad2Δ* strains (**Figure 3.4C**) were conducted in the same manner at the *MAD2* competition except that strains YSL1_c_*mad2Δ*, YHLTP3_*mad2Δ*, YHLTP4_*mad2Δ*, YSL1_c_*mad2Δ*_GFP, were used.

The competition experiments focusing on chromosome IX (Figure 3.4B) were conducted as per the chromosome V competitions. However, the GFP- strains (T7107, 10650_a1, 10650_a2) are W303 background and were obtained from the Tanaka and Nieduszynski labs[192]. Strains 10650_a1 and 10650_a2 are *MATa* haploids derived from strain 10650[192].

Two-dimensional agarose gel electrophoresis: Origin activity was analyzed by standard 2D agarose gel electrophoresis techniques performed on total genomic DNA obtained from either asynchronous or synchronous S-phase cells[183-185]. Asynchronous samples were used for 2D gels conducted on *ARS1426*, *ARS1531*, and *met2::CEN7*. Synchronized samples were used for 2D gels conducted on *ARS1410* and *MET2*. For asynchronous samples, cells were collected in early log phase. For synchronized samples, cells were arrested in G1-phase with alpha factor (final concentration 3 μM) and released synchronously into S-phase by the addition of pronase (0.15 mg/mL). Cells were then collected every 2 minutes and pooled. Genomic DNA was harvested similar to asynchronous samples. In the first dimension, DNA was separated in 0.4% agarose for 18-20 hours at 1V/cm. The second dimension was run for 3-3.5 hours in 1.1% agarose containing Ethidium Bromide (0.3 μg/mL) at ~5-6 V/cm at 4°C.

Flow cytometry: Cells were harvested by mixing with 0.1% sodium azide in 0.2 M EDTA and then fixed with 70% ethanol. Flow cytometry was performed as previously described[186] upon staining cells with Sytox Green (Molecular Probes, Eugene, OR). The data were analyzed with CellQuest software (Becton-Dickinson, Franklin Lakes, NJ).

Density transfer experiments: Density transfer experiments were performed as described[130] with slight modifications in cell synchronization and sample preparation for microarray analysis[106]. Cells were grown overnight at 23°C in a 5 mL culture of dense medium containing ^{13}C-glucose at 0.1% (w/v) and ^{15}N-ammonium sulfate at 0.01% (w/v). Cells were

then diluted into a larger vessel in dense medium and allowed to reach an optical density of 0.3 (~$3.85x10^6$ cells/mL). Cells were arrested in G1 by incubating with 200 μM alpha factor until at least 95% of cells were unbudded based on microscopic analysis. Cells were then filtered and transferred to medium containing ^{12}C-glucose (2%), ^{14}N-ammonium sulfate (0.5%), and alpha factor. The cultures were then brought to 37°C (restrictive temperature for cdc7-1) and synchronously released from the α-factor arrest through the addition of pronase at a concentration of 0.05 mg/mL. At 37°C *cdc7-1* cells arrest at START. Following α-factor release the cells were monitored until ~95% of the population was unbudded. The temperature was then lowered to 23°C to release cells from the *cdc7-1* block. Samples were collected at 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 75, and 90 minutes. Cell samples were treated with a mixture of 0.1% sodium azide and 0.2 M EDTA then pelleted and frozen at -20°C. Genomic DNA was extracted from pelleted cells, digested with *Eco*RI, and the DNA fragments were separated based on density by ultracentrifugation in cesium chloride gradients. Gradients were drip fractionated and slot blotted to nylon membrane where they were hybridized with probes of interest. Unreplicated DNA is HH in density while replicated DNA is HL in density.

Microarray analysis: For microarray analysis, slot blots were hybridized with a genomic DNA probe to identify fractions containing either HH or HL DNA. Once identified, the HH and HL fractions were separately pooled and differentially labeled with cyanine (Cy3 or Cy5) conjugated dUTP (Perkin Elmer)[145]. Timed samples were hybridized to high-density Agilent yeast ChIP-to-chip 4x44K arrays according to the manufacturer's recommendations. The algorithms for analyzing microarray data are described in[147]. An 18 kb window was used for overall smoothing.

To facilitate direct comparison of replication profiles, percent replication values obtained from microarray analysis were converted to Z-score values. Z-score values were calculated using the following formula: $Z = (X-\mu)/\sigma$ where X = the percent replication value for a given probe, μ = the genomic average percent replication for a given sample, and σ = the standard deviation of the distribution of X. The 10 and 20 minute samples were not used because the genomic percent replication values were not well matched between the two strains.

Fluctuation analysis: All fluctuation analyses were conducted at 23°C. Overnight cultures of each strain were grown in C-Ura 2% glucose to maintain selection for the chromosome of interest. The cells were then diluted into complete medium at 0.005% glucose (relaxed selection for chromosome of interest) to a concentration of 0.03333 cells/mL. Cells from this dilution were used to inoculate 96 well plates at a density of 1000 cells/well. Each well held 30 ul cultures. In addition to inoculating the 96 well plates, the diluted cells were spotted to 5FOA + Canavanine plates (192 spots that contain the same number of cells that are present in each well). These plates were scored for colony formation after 96 hours. This plating allowed us to determine the number of false positive wells (those that have 5FOA + Canavanine resistant cells prior to non-selective growth). Furthermore, a subset of the diluted cells was also plated to YEPD plates for single colonies and incubated for 96 hours then scored for CFUs. This number was used to determine the actual number of cells in each well and on each spot prior to grow-out in non-selective medium. Cells were sonicated to disperse any cell clumps prior to inoculation of the 96 well plates and plating. The 96 well plates were sealed and cultures were growth to saturation (~72 hours). The seal was replaced roughly every 24 hours to allow gas exchange.

Following incubation of the 96 wells plates, 192 wells were spotted to 5FOA + Canavanine and 10 additional wells were counted using a flow cytometer, diluted, then plated to

YEPD plates for single colonies. These plates were incubated for 96 hours then scored. The YEPD plates were used to calculate the final number of cells in each well. After the number of false positive spots were removed from final 5FOA + Canavanine counts and the number of generations that occurred during non-selective growth were determined, a mutation rate (loss of *URA3* and *CAN1*) was calculated using the p_o method[193] first described by Luria and Delbruck[194]. This protocol, described in detail at http://fangman-brewer.genetics.washington.edu/fluctuation.htm, was adapted from one used by Dr. Greg Lang (http://glanglab.com/FluctuationAssay.pdf).

Due to the higher mutation rates, the fluctuation analysis that was performed on the *mad2Δ* strains differed slightly from what was just described. First, each well in the 96 well plates contained 8ul cultures with 300 cells. Second, the 96 well plates containing the mad2Δ late centromere V strains were incubated for 24 hours instead of 72 hours.

Characterization of 5FOA + Canavanine resistant colonies: To determine how many of the 5FOA + Canavanine resistant spots from fluctuation analysis were due to chromosome loss, one random colony was taken from each spot and patched to a 5FOA + Canavanine plate. This plate was then replica plated to plates containing either G418 or nourseothricin. Patches that were resistant to 5FOA + Canavanine and sensitive to G418 and nourseothricin were considered to have resulted from chromosome loss.

Competitive growth assay: Cells were grown to log phase growth in complete medium (2% glucose). Each culture was then diluted to an OD ~0.1 and mixed with the GFP labeled reference strain (YSL1_c_GFP or YSL1_c_*mad2Δ*_GFP). The mixture was incubated at 23°C for 90 hours. Throughout this time cells were allowed to reach an OD ≤ 2 where they were then

diluted to an OD ~0.1. The percentage of GFP-positive cells within the culture was calculated at the time of dilution using a flow cytometer.

3.5: Notes

The data presented in this chapter were obtained through the collaborative efforts of Thomas J. Pohl and two undergraduates under his mentorship, SeungBeen (Steven) Lee and Hung-Hsueh (Helen) Lai.

Supplemental information is located in Appendix B.

Acknowledgements

We thank members of the Brewer/ Raghuraman lab along with Meitreya Dunham and members of her lab for helpful discussions and suggestions. This research was supported by the National Institute of General Medical Sciences grant (18926) to BJB and MKR.

FIGURE 3.1

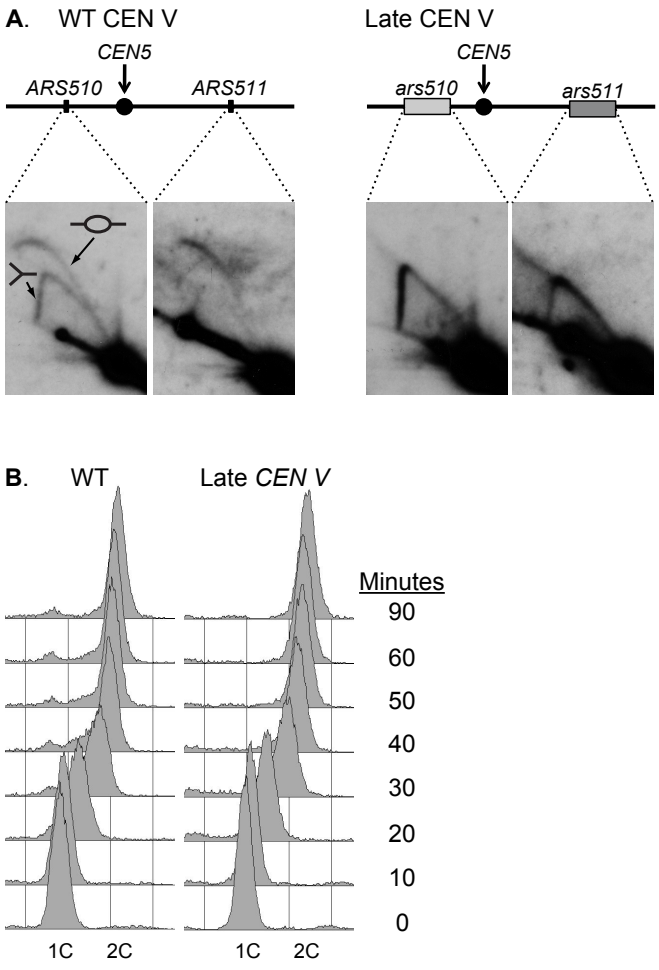

A. WT CEN V
CEN5
ARS510
ARS511
Late CEN V
CEN5
ars510
ars511
B. WT
Late CEN V
Minutes
90
60
50
40
30
20
10
0
1C
2C
1C
2C

Figure 3.1: Constructing a strain to replicate centromere V late in S-phase. (A) In WT cells, centromere V is flanked by two functional origins, *ARS510* and *ARS511*, as indicated by the presence of a bubble arc (depicted by the bubble fragment). A strain in which centromere V replicates late in S-phase was created by replacing *ARS510* and *ARS511* with *KanMX* (light grey box) and *clonNAT*^{R} (dark grey box), respectively. The lack of a bubble arc for either ARS confirms the replacement. **(B)** Flow cytometry of cells harboring either early (WT) and late replicating centromere V. Both strains progress through a synchronized S-phase with similar kinetics and achieve 2C DNA content by 50 minutes.

FIGURE 3.2

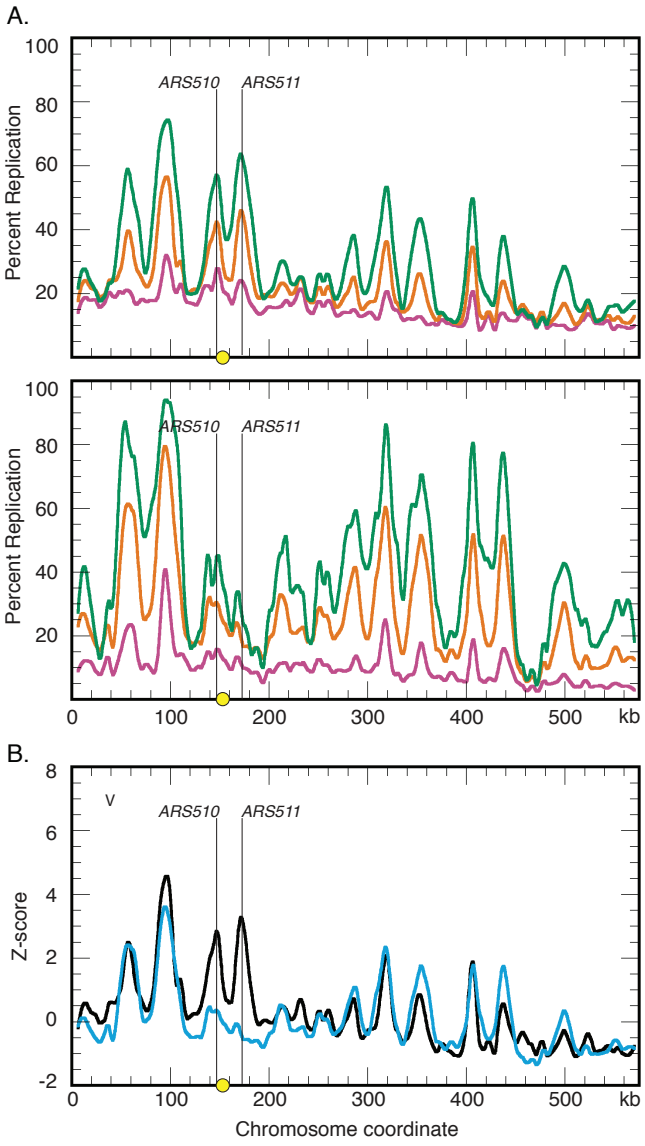
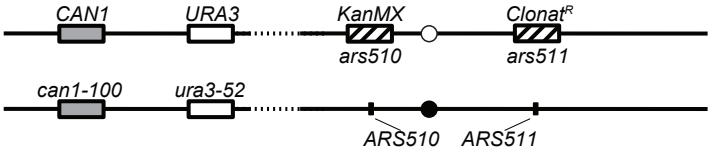

A.
Percent Replication
100
80
60
40
20
ARS510
ARS511
Percent Replication
100
80
60
40
20
ARS510
ARS511
0
100
200
300
400
500
kb
B.
Z-score
8
6
4
2
0
-2
V
ARS510
ARS511
0
100
200
300
400
500
kb
Chromosome coordinate

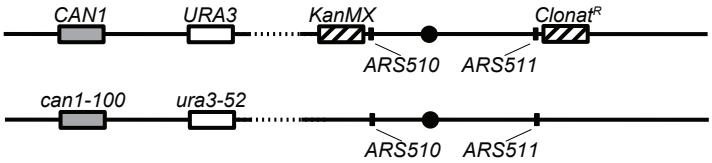
Figure 3.2: Replication dynamics for chromosome V in WT and *ars510 ars511* knockout strains. (A) Replication kinetic profiles of chromosome V in WT (top) and *ars510 ars511* knockout (bottom) strains. Percent replication was monitored across chromosome V at 10 (magenta), 15 (orange), and 20 (green) minutes following release into S-phase. In WT cells, all time points display prominent, early appearing peaks over the locations of *ARS510* and *ARS511*. When *ARS510* and *ARS511* are removed in the ARS knockout strain, replication of the centromere V region is drastically delayed relative to the rest of the chromosome. See Figures S3.1 and S3.2 for all chromosomes. **(B)** Z-score plots of chromosome V in WT (black) and *ars510 ars511* knockout (blue) strains. Replication kinetic profiles from the 15 minute sample were normalized by converting percent replication values to Z-score values (see Materials and Methods). Genomic loci corresponding to *ARS510*, *ARS511,* and the intermediate centromere V regions show significant differences in Z-scores. See Figure S3.3 for all chromosomes.

FIGURE 3.3

A. Heterozygous for *CEN V* replication time



B. Homozygous for *CEN V* replication time



C. Fluctuation Analysis

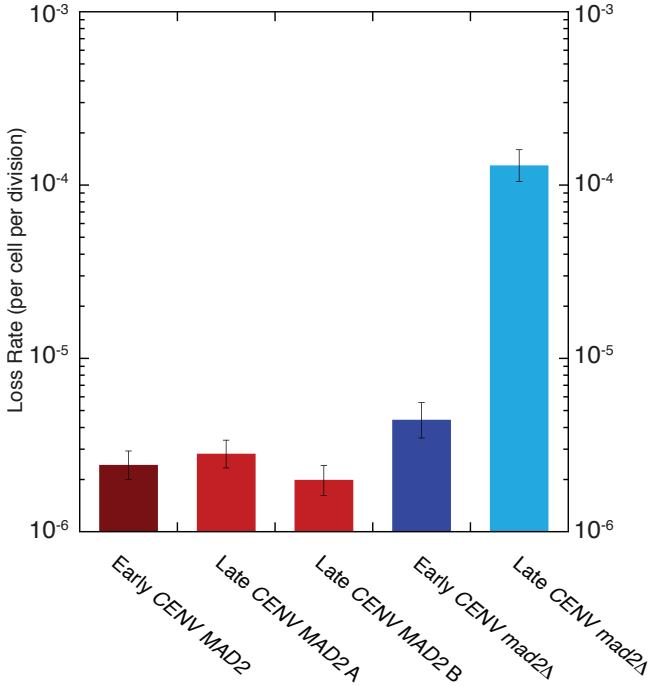


Figure 3.3: Fluctuation analysis of chromosomes with early or late replicating centromeres. (A) Schematics depicting both copies of chromosome V that are present in the diploid that is heterozygous for centromere V replication time. One copy of chromosome V (top) is WT for *CAN1* (grey box) and *URA3* (white box) at their endogenous locations. The centromere (open circle) of this chromosome has been made to replicate in late S-phase through replacement *ars510* and *ars511* with *KanMX* (hashed box left of centromere) and *clonNAT*^{R} (hashed box right of centromere). The other copy of chromosome V (bottom) is WT for *ARS510* and *ARS511* (black boxes); however *CAN1* and *URA3* have been replaced with can1-*100* (grey box) and *ura3-52* (white box). **(B)** Schematic depicting both copies of chromosome V in the diploid that is homozygous for early centromere V replication time. Both chromosomes are WT for *ARS510* and *ARS511* (black boxes). One chromosome (top) is WT for *CAN1* (grey box) and *URA3* (white box) at their endogenous locations. On this chromosome, *KanMX* is integrated down stream of *ARS510* and *clonNAT*^{R} is integrated upstream of *ARS511*. The second copy of chromosome V (bottom) has *can1-100* and *ura3-52* in place of *CAN1* and *URA3*. **(C)** Mutation rates (loss of *CAN1* and *URA3*) for *MAD2* or *mad2Δ* cells that have either an early or late replicating centromere. When centromere V is replicated at its normal time in S-phase *MAD2* cells display a mutation rate of 2.43×10^{-6} (maroon bar) while *mad2Δ* cells have a mutation rate of 4.43×10^{-6} (dark blue bar). When centromere V is replicated late in S-phase *MAD2* cells have a mutation rate of 2.83×10^{-6} and 1.99×10^{-6} (red bars). A muation rate of 1.3×10^{-4} (light blue bar) was observed for this same chromosome when in a *mad2Δ* background.

FIGURE 3.4

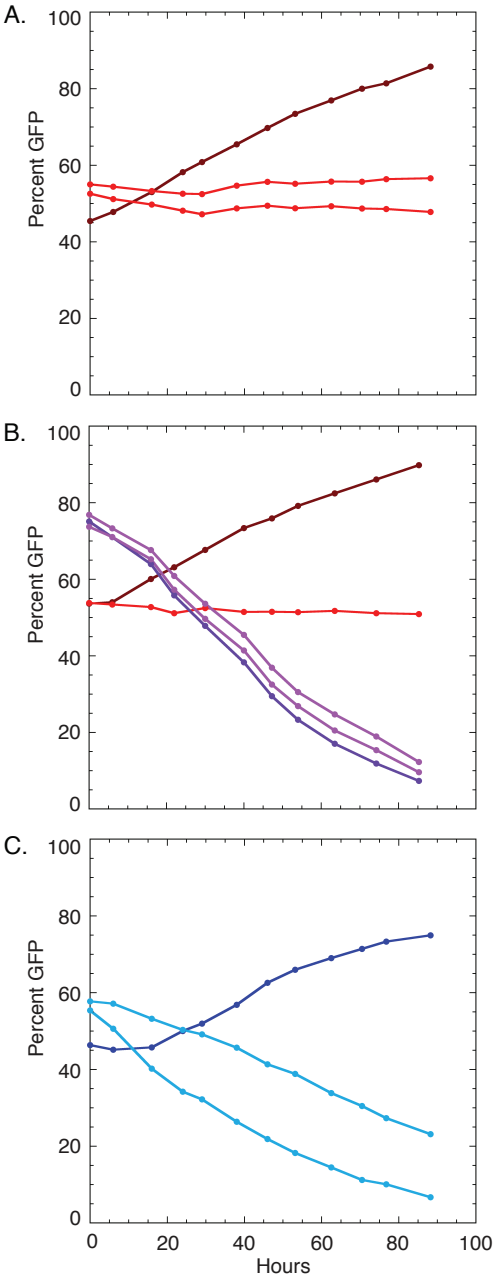


Figure 3.4: Competitive growth analysis of haploids that contain early or late replicating centromere V or centromere XI. All strains of interest were competed against a WT strain from A364A background that was labeled with GFP (GFP+), unless otherwise noted. Both GFP+ and GFP- cells were inoculated into a single culture that was maintained in log phase for 90 hours. Samples of the population were collected throughout a time course and measured for percent GFP. **(A)** GFP+ cells were competed against GFP- cells that replicate centromere V at normal times (maroon) or in late S-phase (red). All strains used to generate this graph are A364A background. **(B)** GFP+ cells were competed against GFP- W303 background cells that replicate centromere IX at normal times (dark purple) or in late S-phase (light purple). The maroon and red graphs are technical replicates of the competitions that generated the maroon and red graphs in part A of this figure. **(C)** GFP+ *mad2Δ* cells were competed against GFP- *mad2Δ* cells that replicate centromere V at normal times (dark blue) or in late S-phase (light blue). All strains used to generate this graph are A364A background.

Table 3.1: Fluctuation Analysis Colony Counts

Strain	Wells spotted to 5FOA + Canavanine	Spots with colonies ≥0.5mm	$G418^R$	$clonNAT^R$
Control (Early *CENV*)	192	126	11	15
Late *CENV* A	191	141	25	29
Late *CENV* B	191	110	5	7

Table 3.1: Fluctuation Analysis Colony Counts
The strains that were used in fluctuation analysis are listed in column 1. Late *CEN* A and B are independent isolates. Column 2 lists the number of wells for each strain that were spotted to 5FOA + Canavanine plates. Column 3 lists number of spots from column 2 that gave rise to colonies over 0.5mm in diameter. One colony from each of these spots was analyzed for the presence of KanMX (G418 resistance, column 4) and clonNAT^{R} (clonNAT resistance, column 5).

Chapter 4

A DNA Sequence Element that Advances Replication Origin Activation Time in *Saccharomyces cerevisiae*[1]

4.1: Introduction

[1]The replication of eukaryotic chromosomes is temporally regulated, with different sections along a chromosome replicating, on average, at different times during S-phase. In the baker's yeast, *Saccharomyces cerevisiae*, the times at which origins undergo activation, or fire, have been shown to contribute to this temporal regulation[106,107,146]. Two classes of origins have been described whose activation is delayed until late in S-phase; in both cases the origins fire late because of their location in the genome. One class of late origins includes those located near telomeres, with *ARS501* near the right end of chromosome V being the prototype of this class[154,195]. Late firing is not an intrinsic property of the origin: transfer to a circular plasmid results in activation in early S-phase. Moreover, when normally early-activated origins are inserted near a telomere they become late activated[111,196,197]. The replication timing of telomere proximal origins has been shown to be governed by the Ku complex (a heterodimer consisting of Yku70 and Yku80 subunits), the silencing chromatin protein Sir3[133], and the TG_{1-3} repeat binding protein, Rif1[132,134]. A second class of late origins is exemplified by the chromosome XIV origin, *ARS1412*[112]. Here, delayed activation of the origin is caused by unspecified non-telomere sequence elements in flanking DNA. These delay elements can influence the activation times of origins that are located up to 6 kilobases (kb) away.

---

[1]The data presented in this chapter were obtained in collaboration with Katherine Kolor, a previous graduate student in the Fangman/Brewer lab. The contribution of Thomas J. Pohl entails generating data presented in figures 5 and 6 in conjunction with writing of the manuscript.

These observations suggested that in the absence of external influences, origins are inherently early activated, and that later firing origins are delayed in their activation because of their particular surrounding sequences. However, it has recently been shown that centromeres function in *cis* to confer early activation on origins residing up to 20 kb away[188,192]. Additionally, some origins that are not close to centromeres owe their early activation time, at least in part, to the binding of transcription factors Fkh1 and Fkh2 up to 500 bp from the origins[189]. These data indicate that early origin activation is not merely a default state but that DNA elements such as centromeres and forkhead protein binding sites found in flanking sequences can promote early origin activation through the binding of kinetochores and forkhead proteins, respectively. In this manuscript, we provide evidence for a third DNA sequence element capable of advancing the time of activation of nearby origins.

The DNA sequence capable of imparting early origin activation was found through analysis of a plasmid, pN&S, that contains two copies of *ARS1* located 180° from each other and designated $ARS1^N$ and $ARS1^S$ (*ARS1 north* and *ARS1 south*, respectively[155] (**Figure 4.1A**). In most cell cycles, only one of the two ARSs on the molecule is activated, and although the two copies of *ARS1* on pN&S are identical in sequence, they are activated differentially such that initiation at $ARS1^S$ is favored approximately 4:1 over initiation at $ARS1^N$. The element causing this origin bias on plasmid pN&S has been localized by deletion analysis to the 3' end of the *URA3* gene, which is near $ARS1^S$ (**Figure 4.1A**;[155]). We have concentrated on two hypotheses to explain the bias in origin use, origin competence and origin timing.

The origin competence hypothesis is that the inherent biochemical ability of one of the *ARS1* copies to fire is affected because either its licensing or some subsequent initiation step is influenced by the 3' *URA3* sequence. Consequently, either the ability of $ARS1^N$ to fire is reduced

or that of *ARS1*^{S} is increased. If this hypothesis were true, then the proposed difference in competence might be manifested as a difference in the ability of cells to maintain plasmids containing only the *ARS1*^{N} region or the *ARS1*^{S} region in the same context relative to the 3' *URA3* sequences as plasmid pN&S. However, the observed relative rates of plasmid loss for such constructs do not reflect the initiation bias. A plasmid bearing *ARS1*^{S}, which is predominantly utilized on pN&S, exhibits a slightly higher rate of loss than does a plasmid with *ARS1*$^{N\ 155}$.

An alternative hypothesis is that the 3' *URA3* sequence causes *ARS1*^{S} to fire earlier on average than *ARS1*^{N}, resulting in passive replication of *ARS1*^{N} before it has the opportunity to fire. A difference in the average time of activation of as little as one minute would be adequate to account for the passive replication of *ARS1*^{N} in a majority of cells given that the distance between the two origins on plasmid pN&S is only ~3.75 kb and the average replication fork rate at 23°C is about 3.5 kb/min[198]. This hypothesis more readily explains the absence of molecules in which both origins are active and is also consistent with the similar maintenance observed for plasmids containing either just the *ARS1*^{N} region or just the *ARS1*^{S} region.

In this study, we have dissected the bias element by deletion analysis and by scanning mutagenesis, and show that this element does indeed cause earlier activation in S-phase of the nearby copy of *ARS1* relative to the second, more distant copy of *ARS1*. Our results provide the identification of a potential third mechanism by which DNA sequences can advance origin activation time.

4.2: Results and Discussion

Identifying the target of the bias determinant

We have previously shown that in the plasmid pN&S (**Figure 4.1A**), *ARS1*^{S} is utilized four times more frequently than is *ARS1*^{N} and that the preferred utilization of *ARS1*^{S} is dependent on sequences residing in the 3' end of the *URA3* gene[155]. Furthermore, this biased utilization of *ARS1*^{S} is not dependent on *URA3* transcription.

Initiation bias in this two-*ARS1* plasmid was determined by examining the direction in which replication forks move through a restriction fragment that lies between the two copies of *ARS1* (**Figure 4.1A**). If *ARS1*^{S} is active, forks move northward; if *ARS1*^{N} is active, forks move southward. Replication intermediates of the *Bgl*II fragment from the west half of the plasmid generate an arc of Y molecules on standard 2D gels with replication forks from *ARS1*^{N} producing the same arc of Ys as forks from *ARS1*^{S} [155]. To distinguish the two sets of replication intermediates we included an in-gel restriction digestion with *Pst*I prior to the second dimension of electrophoresis and probed for the larger *Pst*I-*Bgl*II fragment. Ys created by forks emanating from *ARS1*^{S} create an arc that rises from the 1N spot, while Ys from *ARS1*^{N} create an arc that rises from a position displaced from the 1N spot (**Figure 4.1B**;[199]). The amount of signal in each arc of Ys is quantified (**Figure 4.1B**) to yield the percentage of molecules replicated by forks from *ARS1*^{S}—values above 50% indicating biased initiation from *ARS1*^{S}. The original two-ARS plasmid, pN&S, (**Figure 4.1C**) shows *ARS1*^{S} initiation 75.2% of the time based on the values obtained from five different cultures (*ARS1*^{S} initiation values of 71, 74, 74, 77, and 80%). These values provide a baseline for comparison with other constructs. Because the labor involved in each gel experiment limits the sample sizes that can reasonably be obtained, in testing other constructs (see below), we have chosen to be conservative in deciding whether biased initiation is significantly different from that of pN&S. Thus, we conservatively conclude that *ARS1*^{S}

initiation percentages falling outside the range of 68.1% to 82.3% (75.2+/-7.1%; Students t-test 99% confidence interval) are significantly different from pN&S at the 1% level.

We first determined the target of the bias determinant residing at the 3' end of the *URA3* gene. This initiation bias could be due to an effect on the activation of $ARS1^N$ or on $ARS1^S$. To determine which copy of *ARS1* is being affected by the 3' *URA3* sequences we made two derivatives of pN&S, each having one of the two copies of *ARS1* inverted compared to its original orientation. ARS sequences are inherently asymmetrical as they are composed of discrete A and B elements: the three B elements in *ARS1* lie on the 3'-side of the T-rich strand of the 17 bp AT-rich ARS consensus sequence (A element) (**Figure 4.1A**; [200]). Because of this asymmetry, we reasoned that the orientation of one of the copies of *ARS1* with respect to the bias determinant might be important for the initiation bias.

We found that inversion of $ARS1^N$ in pN&S fails to reduce the bias (74% vs. 83%; compare **Figures 4.1C and 4.1D**), whereas inversion of $ARS1^S$ reduces it substantially (74% vs. 56%; compare **Figures 4.1C and 4.1E**). Since the endpoints of the $ARS1^S$ inversion do not include the sequences required for the initiation bias (described below), the reduction of the initiation bias can be attributed to the failure of the bias determinant to act on the ARS when the ARS is in inverted orientation and not to the disruption of the sequence causing the bias. These data indicate that the primary target of the bias determinant is $ARS1^S$ rather than $ARS1^N$.

Location and structure of the bias determinant

Determination of the initiation bias in the two-ARS plasmid provides a relatively simple assay for a mutational analysis of this element. Previous observations showed that deletion of the *Nco*I-*Nru*I fragment, which included the 5' half of *URA3* (plasmid pΔ34), does not eliminate

the preferential use of *ARS1*^{S}, but that deletion of the *Nsi*I-*Nco*I fragment from the 3' end of *URA3* (plasmid pΔ43) results in equivalent use of both copies of *ARS1* (**Figure 4.2**; [155]). To further localize the early determinant, four deletions (plasmids pΔ5, pΔ6, pΔ7, and pΔ8) that remove sections of the *Nsi*I-*Nco*I sequence were constructed using convenient restriction sites, and tested for origin bias. Deletion of the 169 bp *Apa*I-*Nco*I fragment (pΔ5), the 302 bp *Alw*NI-*Apa*I fragment (pΔ6), or the 113 bp *Hae*III-*Alw*NI fragment (pΔ7) had no significant reduction of the bias. However, the 34 bp *Nsi*I-*Hae*III fragment contributes substantially to the bias since its deletion (pΔ8) significantly reduced the bias (**Figure 4.2**).

To examine a possible role of sequences immediately to the left of the *Nsi*I-*Hae*III fragment, a 56 bp *Sma*I-*Nsi*I deletion was made (plasmid pΔ9). This deletion greatly reduced the bias (**Figure 4.2**) showing that these deleted sequences also play a major role. Further deletion to the left, removing the 228 bp *Hind*III-*Sma*I fragment including a portion of the *TRP1* sequences (plasmid pΔ10), had little effect on the initiation bias (**Figure 4.2**). Based on these results, plasmid sequences that contribute substantially to initiation bias appear to lie largely within the 90 bp *Sma*I to *Hae*III sequence in the *URA3* gene (**Figure 4.2, bracket 1**). However, despite finding little effect of deletions Δ5, Δ6, and Δ7, which together span the 584 bp *Hae*III-*Nco*I fragment to the right of the *Sma*I-*Hae*III section, deletion of the entire *Hae*III-*Nco*I fragment (plasmid pΔ11) eliminated the origin bias (**Figure 4.2**). Because no single internal deletion in this region (pΔ5, pΔ6, and pΔ7) dramatically reduced the initiation bias, there appear to be functionally redundant contributions to origin bias by the *Hae*III-*Nco*I region. In summary, sequences within the *Hae*III-*Nco*I fragment (**Figure 4.2, bracket 2**), along with those in the *Sma*I-*Hae*III fragment, contribute to biased origin use. Due to the apparent complexity of

sequences within the *Hae*III-*Nco*I region, we focused our further studies on the unique 90 bp fragment (**Figure 4.2, bracket 1**).

To further determine which sequences within the 90 bp *Sma*I-*Hae*III fragment (**Figure 4.2, bracket 1**) contribute to bias determinant activity, this fragment was examined by scanning site-directed mutagenesis. Typically, 10 bp changes were made (one mutation, m7, was a 9 bp change) (**Figure 4.3A**). Because an *Xho*I restriction site was also included for screening purposes two base pairs out of the 90 bp examined were not changed by the mutagenesis (**Figure 4.3A, overlines**). Analysis of $ARS1^S$ initiation percentage for the ten mutants (**Figure 4.3B**) revealed that two of them, m6 and m7 (**Figure 4.3C**), exhibited dramatic reductions in the initiation bias while a third mutant, m5, showed a slightly reduced initiation bias. The other seven mutations gave initiation biases that were not significantly different from that of the original two-ARS plasmid (**Figure 4.3B, vertical bracket**). Therefore, through scanning mutagenesis we identified a 29 bp sequence (**Figure 4.3A, red**) within the *Sma*I-*Hae*III fragment that contains bases important for origin bias.

To further characterize bases that are essential for the origin bias, three plasmids were constructed to have a 6 bp mutation on either the 5' or the 3' ends of the 29 bp sequence or containing a central 4 bp deletion. These plasmids were also analyzed via fork direction 2D gel analysis (**Figure 4.3D and 4.3E**). While the 6 bases on the 3' end of the fragment are essential for origin bias, the 6 bases on the 5' end and the 4 bases that were centrally deleted are not. These data result in the identification of a 19 bp bipartite sequence that is essential for origin bias (**Figure 4.3D, red**). The essential 19 bases are consistent with the results obtained for plasmids pΔ8 and pΔ9 in our deletion analysis (**see Figure 4.2**). Interestingly, our data suggest that

although the identified 19 bases are essential for the bias they are not sufficient to cause the bias, requiring the presence of sequences within the broader 584 bp *Hae*III-*Nco*I fragment.

The bias determinant advances *ARS1*^{S} origin firing time relative to *ARS1*^{N}

The observation that the bias determinant at the 3' end of *URA3* is influencing *ARS1*^{S} (**Figure 4.1**) without preferentially affecting its efficiency relative to that of *ARS1*^{N} [155] led us to the hypothesis that the determinant causes the bias by advancing *ARS1*^{S} activation time. Since *ARS1*^{N} is used about 20% of the time it suggests that, even if the *ARS1*^{S} is usually activated earlier than is *ARS1*^{N}, there is some overlap in the distribution of their activation times. It seems likely, therefore, that the difference between their mean times of activation could differ by a small increment, perhaps by as little as one minute. We tested this origin timing hypothesis using three different experimental approaches.

Comparison of 3'URA3 sequence and a centromere. If the bias determinant were to impart early activation to *ARS1*^{S} we would expect that other DNA elements, such as centromeres, that are known to advance origin timing would produce a similar bias when placed in the context of the two-ARS plasmid. We would also predict that a centromere placed near the opposite copy of *ARS1* would diminish or neutralize any bias that would be imposed by the 3' *URA3* bias element.

We first asked if a centromeric sequence could produce a similar bias as the 3' *URA3* sequence in the context of pN&S. To do so, we utilized a derivative of pN&S that lacks the sequence necessary for 3' *URA3* dependent bias and therefore shows no ARS bias (pΔ9, **see Figure 4.2**). Integration of a centromere close to *ARS1*^{N} in this plasmid (pΔ9C) resulted in a marked bias toward *ARS1*^{N} utilization, indicating that sequences that are known to advance

origin activation time in chromosomal DNA can in fact impose an origin utilization bias in pN&S (**Figure 4.4A, left panel**). Although a centromere's ability to advance origin activation time is bi-directional, the centromere also has a greater timing effect on closer origins than more distant ones[188]. Therefore, it is conceivable that in the context of the two-ARS plasmid the centromere is advancing the activation time of both *ARS1* copies, with $ARS1^N$ being advanced to a greater extent. Insertion of the centromere near $ARS1^N$ can also neutralize the bias imparted by the 3' *URA3* sequence: $ARS1^S$ and $ARS1^N$ were equally used on a plasmid harboring both a centromere nearest to $ARS1^N$ and the intact 3'*URA3* sequences (**Figure 4.4A, right panel**). The equal usage of the two *ARS1* copies can be interpreted as $ARS1^S$ undergoing activation at a time similar to $ARS1^N$, which is replicated early in this situation due to the nearby centromere. These data suggest that the 3' *URA3* sequences are capable of advancing origin activation time.

The bias determinant promotes early $ARS1^S$ activation. To further test whether the bias determinant promotes the earlier activation of $ARS1^S$, replication times were determined for plasmids that contain only $ARS1^S$, either with or without a functional bias determinant. If the bias determinant were causing a delay in the activation of $ARS1^N$, then deletion of the bias determinant should have no effect on the time of activation of $ARS1^S$ and the two plasmids should replicate at the same time. If, instead, the bias determinant were advancing the activation time of $ARS1^S$ then the plasmid with the bias determinant should replicate earlier than the one without it. Plasmid p5-6C is identical to the two-*ARS1* plasmid, pN&S, with two exceptions: (1) the 1.45 kb *EcoRI* fragment containing $ARS1^N$ has been removed, and (2) a 1.1 kb *CEN5* fragment has been cloned into the *BamHI* site. A second plasmid, p5-6m7C, is identical to p5-6C except that it contains a 9 bp substitution mutation, m7, which eliminates the initiation bias in the two-ARS plasmid (**Figure 4.3C**). Because *CEN5* is present at the same location in both p5-

6C and p5-6m7C, any timing effect that *CEN5* would impart on *ARS1* would be equal in both plasmids. Integration of *CEN5* does not interfere with detecting the influence of the 3' *URA3* sequence on $ARS1^S$ (**Figure 4.4A**) and was necessary to achieve the efficient plasmid maintenance required for density transfer experiments.

Plasmids p5-6C and p5-6m7C were transformed separately into *S. cerevisiae* and their times of replication were determined in parallel cultures using a modified version of the Meselson-Stahl density shift experiment with synchronized cultures[130]. Cultures of cells with these two different plasmids have the same population doubling times. The percent replication of p5-6C and p5-6m7C was calculated for each sample and plotted with respect to time (**Figure 4.4B**). The time of replication (T_{rep}) for each locus was calculated as the time at which that locus reached half maximal replication. *ARS305*, one of the earliest known origins, and R11, a late replicating fragment on chromosome V, were used as timing standards. To facilitate comparison between cultures, these T_{rep} values were converted to replication indices[112] by assigning *ARS305* a replication index (RI) of 0 and R11 a RI of 1.0 (**Figure 4.4B, lower panel**). Most other genomic loci have a RI between 0 and 1.0. The T_{rep} values for the two plasmids and a sequence immediately flanking chromosomal *ARS1* were then converted to RIs corresponding to the fraction of the *ARS305*-R11 interval elapsed when the T_{rep} for each locus was obtained. p5-6C (RI = -0.14) replicated 3-4 minutes earlier than did a chromosomal fragment containing the early replicating origin, *ARS305*, whereas p5-6m7C (RI = 0.1) replicated at the same time as *ARS305* (**Figure 4.4B**). As expected, the replication indices for chromosomal *ARS1* fragments were similar between the two strains (RI = 0.35 and 0.36). These results indicate that the bias determinant acts positively on the $ARS1^S$ origin to advance its time of activation.

Consistent with these results, the 3-minute timing advantage that the bias determinant imparts on *ARS1*^{S} was recapitulated using a site-directed recombinase to split plasmid pN&S into a south half and a north half *in vivo*, allowing the firing time of *ARS1*^{N} and *ARS1*^{S} to be directly compared within the same cell (**Supplemental Figure 1**). Together, the results of these three assays indicate that *URA3* sequences are influencing replication initiation time on the two-ARS plasmid by advancing *ARS1*^{S} activation time, and that this difference in initiation time can explain the initiation bias. The 3-4 minute advancement in activation time caused by the 3'*URA3* sequence explains how two equally efficient ARSs can exhibit large differences in levels of activation when placed in competition on a single plasmid molecule.

The bias determinant can regulate origin activation time over a distance of at least 1 kb

Other ***cis***-acting timing determinants that have been identified can affect replication timing over great distances. Centromeres advance origin activation time over a distance of 20 kb[188,192] while telomeres delay the activation time of origins over about 30 kb[154]. The activation times of two origins on chromosome XIV, *ARS1412* and *ARS1413*, are delayed by cryptic sequences that reside up to 6 kb away[112]. The forkhead transcription factors, Fkh1 and Fkh2, are capable of advancing origin activation time over a distance of at least 500 bp[189]. We therefore sought to determine the distance over which the *URA3* timing determinant can act. An 8 bp *Xho*I linker was inserted at the *Sma*I/*Eco*RV junction between *ARS1*^{S} and the bias determinant to provide a cloning site for several subsequent insertions. This small insertion had no effect on the bias (**Figure 4.5A**). Insertion of either one or two copies of a 75 bp *Xho*I PCR lambda fragment at the *XhoI* site (**Figure 4.5B, C**) or of a 604 bp lambda *Stu*I fragment at the *Sma*I/*Eco*RV junction (**Figure 4.5D**) also had no effect on the bias. Since the distance from the *Sma*I/*Eco*RV

junction to the ACS of *ARS1*^{S} is 465 bp, these data indicate that the bias determinant can function over a distance of at least 1069 bp.

An alternative, though unlikely, interpretation of this experiment is that the bias determinant is not able to function over this distance and that instead, the lambda sequences are promoting an initiation bias of their own. To eliminate this possibility, the bias determinant was compromised in the pN&S+604 bp construct by an *Nco*I-*Nsi*I deletion (see pΔ43 in **Figure 4.2**). This construct exhibits no bias (data not shown), indicating that the lambda sequences are not capable of conferring a bias on their own.

At its native location *URA3* resides at a distance from its neighboring origins that is ideal to test the outer limits over which the bias element can function. In the genome, *URA3* is flanked by two efficient early activating origins, *ARS508* and *ARS510,* that are located at respective distances of 21.9 and 28.8 kb from *URA3*[65]. However, only *ARS510* is in the correct orientation to have its activation time regulated by the bias determinant. Therefore, we expect that if the bias determinant regulates origin activation at its native location its effect would be on *ARS510*. To determine if the bias determinant influences replication at its native genomic location we performed a density shift experiment on cells in which the chromosomal copy of the 3' *URA3* bias element had been removed (**Figure 4.5E**). Cells lacking the bias determinant replicated both the *URA3* gene and *ARS510* similar to control cells, suggesting that the bias determinant does not detectably regulate the activation time of flanking origins at its native location. These results could also be obtained if the bias determinant was effective only on plasmids and incapable of advancing origin activation time in the context of a linear chromosome. However, we can rule out this possibility because the bias effect persists when pN&S is integrated into the chromosomal *ARS1* locus[155].

Taken together, these experiments indicate that the *URA3* timing determinant can act over a distance of at least 1 kb but not over a distance of tens of kb. As more sequences that are capable of regulating origin activation time are identified it will be interesting to determine more precise limits or boundaries for each given regulator.

DNA binding motif search of the bias determinant

The effects of deletions in the 3' *URA3* region and of the inversion of *ARS1* on bias activity indicate that the *ARS1* element and the bias determinant are separate entities. Does the bias determinant communicate with the ARS through the binding of specific proteins to the bias determinant? Based on our observation that the 3' end of *URA3* consists of a complex, multipartite timing determinant, we were aware that identifying any trans-acting factor(s) mediating the effect of this determinant might be challenging: for example, that redundancy introduced by multiple trans-acting factors interacting with the determinant might complicate the analysis. Nevertheless, finding that at least one component of the determinant could be narrowed to a small (19-23 bp) element encouraged us to embark on a search for proteins such as transcription factors that might mediate the timing effect by binding to that sequence. To this end, we first considered proteins known to influence origin activity. *ARS1* contains a binding site for the transcription factor Abf1 in element B3 that contributes to ARS activity[67], and it has been demonstrated for *ARS121* that a second Abf1 binding site, outside of the B region, contributes to maintenance of an *ARS121* plasmid under some culture conditions[201]. In addition, a binding site for a different transcription factor (Rap1) was shown to substitute for the Abf1 site at *ARS1*[67]. Although the mechanism by which transcription factors enhance ARS function is not known, these observations suggest that transcription factor binding might contribute to the bias

observed in pN&S. There is a potential binding site for Abf1 and one for Rap1 in the 90 bp *Sma*I-*Hae*III fragment of the bias element region. However, neither of these potential sites overlaps the 23 bp essential core element identified by scanning mutagenesis (**Figure 4.3**), indicating that these transcription factors are not primarily responsible for the origin initiation bias. Given the complexity of the bias determinant, however, a possible redundant role for Abf1 and Rap1 cannot be excluded.

The forkhead transcription factors, Fkh1 and Fkh2, recently have been shown to influence the initiation time of some origins[189], raising the possibility that these transcription factors might have some role in the observed bias determinant. Furthermore, there are potential binding sites for Fkh1 and Fkh2 close to or partially overlapping the core bias element (**Figure 4.3A**). However, these binding sites do not overlap with all sequences that were shown to be crucial for the bias (see m11, **Figure 4.3D**) suggesting that Fkh1 and Fkh2 are likely also not involved in the observed origin initiation bias. Consistent with this logic, preliminary chromatin immunoprecipitation studies performed by the Fox lab indicate that Fkh1 does not bind this sequence at its native chromosomal location (Tim Hoggard and Catherine Fox, personal communication).

Since the obvious candidates are likely not responsible for the bias, we embarked on a more systematic search for *trans*-acting factors that might bind to the bias determinant. Both the essential 23 bp sequence and the 19 bp sequence (with the four central bases deleted) were run through the motif comparison tool TOMTOM[202]. In both instances, binding motifs for multiple proteins were identified, many of which are specific to transcription factors (**Figure 4.6A**).

The top hit in both searches was the binding motif for the transcription factor Cst6 (**Figure 4.6**). Of particular interest was the finding that while the Cts6 binding motif overlaps

the *Nsi*I site, the motif would be restored in the four-base (TGCA) deletion (**Figure 4.6B**). The next candidate on the list was the transcription factor, Yap3. Like all other proteins on the list, the Yap3 consensus binding motif only overlapped with part of the sequence that was found to be essential for the bias and like many of the other proteins Yap3 has an imperfect target-site match (**Figure 4.6B**). To determine if either of the top two hits (Cst6 or Yap3) was responsible for the bias, plasmid pN&S was transformed into cells lacking either *CST6* or *YAP3* and analyzed for the presence or absence of the bias (**Figure 4.6B**). As in WT cells, an initiation bias for $ARS1^S$ was observed for pN&S indicating that neither Cst6 nor Yap3 binding is responsible for the bias.

As we have excluded the "top" two candidates from our motif search, a systematic approach will need to be taken to analyze the remaining proteins that are potentially involved. However, from these studies, we have concluded that given the large number of potential binding candidates, the degeneracy of transcription factor binding sites, the possibility of redundancy at the bipartite site, and the labor involved in testing each candidate, an approach other than gene knockouts will be required to identify the putative trans-acting factors that regulate this initiation timing determinant. Furthermore, it is possible that none of the candidate proteins identified in our motif search are uniquely involved in creation of the bias. This possibility is supported by the observation that none of the motifs for the remaining candidates span the entire 23 bp motif that is essential to the bias. It is worth noting that the percentages of $ARS1^S$ usage in both *CST6* and *YAP3* knockout strains are either very close to or slightly below the bottom of the confidence interval. Though these data were not deemed significantly different from pN&S, it would be interesting to determine how a double knockout of *CST6* and *YAP3* would affect $ARS1^S$ bias.

The observation that double knockouts of transcription factors can present adverse phenotypes such as pseudohyphal growth[189] adds to the complexity of such studies.

Finally, we considered the possibility that bias determinant acts directly via some intrinsic property of its own, not necessarily mediated by the binding of *trans*-acting factors. The B2 element of *ARS1* is part of a DNA unwinding element (DUE), which is believed to enhance ARS function through its inherent helical instability[77,203]. Computer analysis of the *Sma*I-*Hae*III fragment of pN&S shows that it is part of a DUE[77,203]. Does this DUE contribute, autonomously or in conjunction with the binding of a protein, to earlier activation of $ARS1^S$? We made several site-directed mutations of the 19 bp core sequence, scrambling the sequence without altering the predicted helical stability (Materials and Methods). These mutations did substantially reduce the origin bias (**Supplemental Figure 2**). Therefore, the 19 bp sequence appears to contribute to origin bias in a way that depends on its specific sequence and not simply on its propensity to unwind. While a role for this particular DUE at the 3' end of the *URA3* gene in origin bias cannot be excluded, a DUE alone is not sufficient to advance the time of origin firing.

Other than our studies examining the bias determinant at its native location, the ability of this early timing determinant to influence other ARS elements has not yet been tested. The possible response of other ARS elements can be approached first, and most simply, by using the origin bias assay. However, it may be that response to the early determinant requires specific sequences within *ARS1*. If so, the sequence elements in *ARS1* that respond to the determinant could be identified using mutagenesis of $ARS1^S$ in the context of the two-ARS plasmid assay. For any mutation that reduces the origin bias, it would be important to alter both ARSs in the plasmid identically to distinguish between mutations that reduce origin function from those that

are specific to the origin's ability to respond to the bias determinant. In addition, the bias assay with the two-ARS plasmid has the potential for providing a sensitive tool for characterizing other *cis*-acting factors involved in replication initiation timing control. For example, sequences on chromosome XIV have been shown to delay the time of activation of origins that they flank, and they have been shown to act on *ARS1* transplanted from chromosome IV[112]. These late determinants can delay origin initiation by 10 minutes or more, and might cause an even greater bias in the two-ARS plasmid than that conferred by the 3' *URA3* early determinant.

4.3: Materials and Methods

Strains and plasmids: Strain 3xRZeoDIR, which is RM14-3a (background A364a, *MATa*, *cdc7-1*, *bar1*, *his6*, *leu2-3,112*, *trp1-289*, *ura3-52*; [130] with three copies of the 6891 bp *Cla*I fragment of pRINT[204] tandemly integrated at the *LEU2* locus and a zeocin ARS-less cassette on the right arm of chromosome V, was used for the comparative hybridization experiments[204]. The ARS-less cassette was created by integrating the 3083 bp *Pvu*II fragment of pZeodir (described below) by homologous recombination. pN&S deletion derivatives were analyzed in RM14-3a or MR14-3a. Since the *trp1-1* allele of RM14-3a is somewhat revertible, MR14-3a, which is RM14-3a with a *TRP1* knockout allele, was used for all deletion constructs except pΔ43 and pΔ5. MR14-3a was used for all of the oligonucleotide mutagenesis constructs.

Plasmids were constructed in *E. coli* strain DH5α. All bias determinant plasmid constructs are based on plasmid pN&S[155]. Due to the duplication of the *TRP1ARS1* sequences in the two-*ARS1* plasmid, it was often difficult to find appropriate unique restriction sites for plasmid modifications. To circumvent this problem, several plasmids were constructed by a multi-step process beginning with plasmids that contain only the south half of the plasmid.

These plasmids are pBB3, pBB6, pBB6mut and pBB6TA. pBB3 has the 967 bp *Nde*I-*Sma*I *URA3* fragment from YIp5 cloned into the 245 bp *Nde*I-*Sma*I fragment of pUC18. pBB6 has the 1068 bp *Eco*RI-*Eco*RV fragment of *TRP1ARS1* fragment cloned into the 3386 bp *Eco*RI-*Sma*I fragment of pBB3. pBB6TA has a 1453 bp *Eco*RI *TRP1ARS1* fragment cloned into the *Eco*RI site of pBB3 in the same orientation as pBB6. pBB6mut is pBB6TA with a 293 bp *Nde*I-*Nco*I deletion. After the desired alteration had been made, the two-ARS plasmid was re-constructed. pZeodir was constructed by ligating the 1181 bp blunt *Kpn*I-*Hin*dIII fragment of pUT322 (CAYLA) into the 4258 bp blunt *Hin*dIII fragment of pKAN-DIR[204].

Mutagenesis and consensus sequence alignments: The revertible selectable marker method was used for mutagenesis[205]. Oligonucleotides were purchased from Gibco BRL. The sequences of the mutations m1-10 are given in **Figure 4.3A**. The sequences of the mutations m11, m12 and ΔNsiI are given in **Figure 4.3D**. The sequences of the DUE mutations corresponding to the m5-m7 positions were m5DUE (ATCAAGTACT; data not shown), m6DUE (ATAGACGTCA; **Supplemental Figure 2**), and m7DUE (ATTAATAGT; **Supplemental Figure 2**). Correct transformants were identified by restriction digestion screening for incorporation of a restriction site (*Xho*I for m1-10; *Sca*I for m5DUE; *Aat*II for M6DUE; *Ase*I for m7DUE), and the mutations were confirmed by sequencing. The success rate of the mutagenesis was generally around 60%.

Consensus matches for the binding sites of Abf1p and Rap1p were identified using the Intelligenetics program Seq. The parameters used were as follows: GapPenalty=1, GapSizePenalty=0.3333, MinAlign=4, MisMatchPenalty=0, OverLap=33.3%, SegmentSize=30, Expect=10, LoopOut=3, MinMatch=5. Sequences were entered using the Intelligenetics program GENED. Computer analysis of DNA unwinding elements (DUE) was performed using

VooDUE (M. K. Raghuraman, unpublished), a Macintosh version of the Thermodyne program[77,203]. It calculates the energy of unwinding of the DNA duplex in the same manner as the Thermodyne program.

Two-dimensional gel electrophoresis: Samples for two-dimensional (2-D) gels were collected at 1.5-minute intervals for 33 minutes following release of the cells into S phase, then pooled. Cell synchrony and release were performed as described[184]. DNA was prepared as described[184,185]. Determining the direction of replication fork movement by 2-D gels involved an in-gel digestion as described[199]. For every construct, the west half of the plasmid was analyzed. The blots were hybridized to a radioactively labeled 1185 bp *Bgl*I-*Nde*I fragment of pBR322 and analyzed on an InstantImager (Packard) for quantitation.

Determination of replication time: Density shift experiments on the strains analyzed in **Figures 4.4B and 4.5E** were respectively performed in strains RM14-3a and KK14-3a (a URA+ version of RM14-3a) as previously described[130]. The data were analyzed using an InstantImager or phosphorimager (Packard or BioRad, respectively). The comparative hybridization experiments were performed essentially as described[204]. Strain 3xRZeoDIR containing pN&Sdir or pN&SDdir (described in Results) was incubated for 1.25 generations in α-factor. Upon arrest in the G1 phase, an uninduced sample was taken and galactose was added to 2%. Excision of the ARS-less cassette was allowed to proceed for four hours. Glucose was then added to 2% and the cells were released from α-factor arrest at 37˚C, the restrictive temperature for *cdc7*^{ts}. Upon *cdc7* arrest, a sample was taken, and the cells were returned to 23˚C to allow entry into S phase. Samples were taken at 4-min intervals throughout S (60 min). The DNA was prepared by the "Smash and Grab" procedure[206] and digested with *Bam*HI and *Eag*I. The samples were run on a 0.7% agarose gel, blotted onto Hybond (Amersham) nylon membrane and hybridized to a

*Bam*HI-*Eag*I fragment of pBR322 labeled with ^{32}P. Quantification of the blots was performed on a PhosphorImager (Molecular Dynamics) using ImageQuant software.

4.4: Notes

This chapter has been published: Pohl, T.J.[1], Kolor, K.[1], Fangman, W.L., Brewer, B.J., and Raghuraman, M.K. (2013). *G3:Genes, Genomes, and Genetics*. doi:10.1534/g3.113.008250. PMID:24022751

[1]Co-first authors

Supplemental information is located in Appendix C.

Acknowledgements

We thank past and present members of the Brewer/Raghuraman laboratory along with members of the Dunham laboratory for thoughtful discussions and comments on this manuscript. This research was supported by a National Institute of General Medical Sciences grant (18926) to WLF, BJB and MKR.

FIGURE 4.1

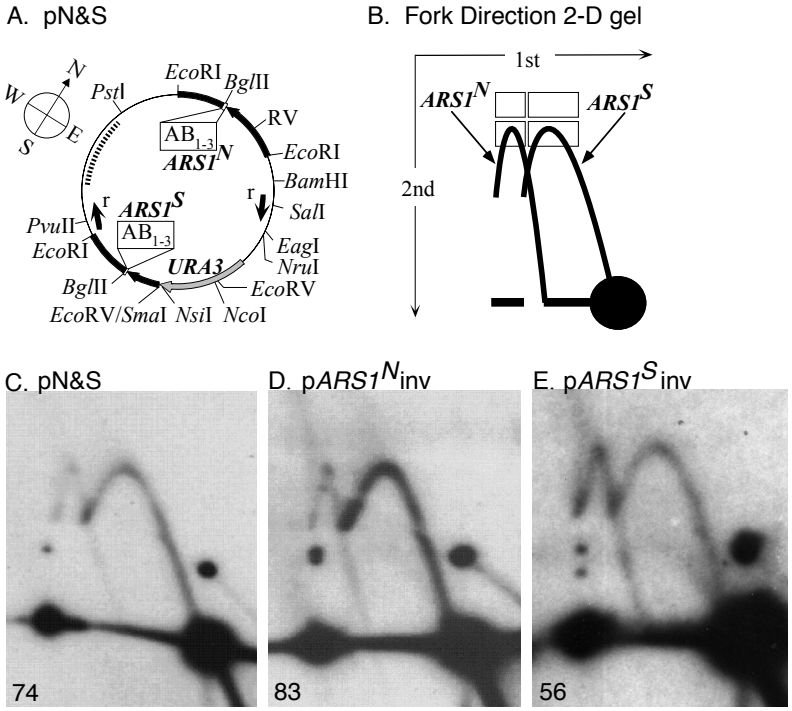

A. pN&S
B. Fork Direction 2-D gel
N
W
E
S
PstI
EcoRI
BglII
RV
AB1-3
ARS1N
EcoRI
BamHI
SalI
ARS1S
AB1-3
PvuII
EcoRI
EagI
NruI
URA3
BglII
EcoRV
EcoRV/SmaI NsiI NcoI
1st
2nd
ARS1N
ARS1S
C. pN&S
D. pARS1Ninv
E. pARS1Sinv
74
83
56

Figure 4.1: Fork direction analysis of the two-*ARS1* plasmid by 2-D gel electrophoresis. **(A)** Map of pN&S, a ~7.5 kb plasmid described previously[155]. Thick lines indicate sequences derived from yeast while thin lines indicate prokaryotic vector sequences. The "compass" bearings (see text) are indicated by N, S, E and W. A *Bgl*II site occurs immediately 3' to the T-rich strand of the 11 bp ARS consensus sequence of each *ARS1*. The orientation of each copy of *ARS1* is indicated by the boxes containing letters representing the A and B regions (B1-3) defined for *ARS1*[67] . The *URA3* gene is drawn as a gray arrow. The black arrows represent *TRP1* sequences. The *TRP1* gene fragment nearest to *ARS1s* lacks its promoter and 5' sequences. The small arrows (labeled "r") located in the plasmid's interior represent the positions of the recombinase sites (cloned at the *Sal*I and *Pvu*II sites indicated) to make pN&Sdir (see Supplemental Fig. 1). Not all sites for each enzyme are shown. The dotted line represents the region and the probe used to analyze fork direction 2-D gels. **(B)** Cartoon depicting 2-D gel separation of forks moving in opposite directions through the southwestern *Bgl*II-*Pst*I fragment of pN&S. *Bgl*II digestion was carried out before electrophoresis in the first dimension and an in-gel digestion with *Pst*I before the second dimension. Forks emanating from $ARS1^N$ (left arc) are distinguished from the forks emanating from $ARS1^S$ (right arc). The efficiency of each ARS is expressed as a percentage of the total hybridization signal from an arc (two lower rectangles) after subtracting background hybridization (two upper rectangles). **(C-E)** Fork direction gels for **(C)** pN&S, **(D)** $pARS1^N$inv (pN&S with the 1.45 kb *Eco*RI fragment containing $ARS1^N$ reversed), and **(E)** $pARS1^S$inv (pN&S with the 1.05 kb *Eco*RI-*Eco*RV fragment containing $ARS1^S$ reversed). The numbers in the lower left of each panel refer to the percent initiation at $ARS1^S$ (see text for details).

FIGURE 4.2

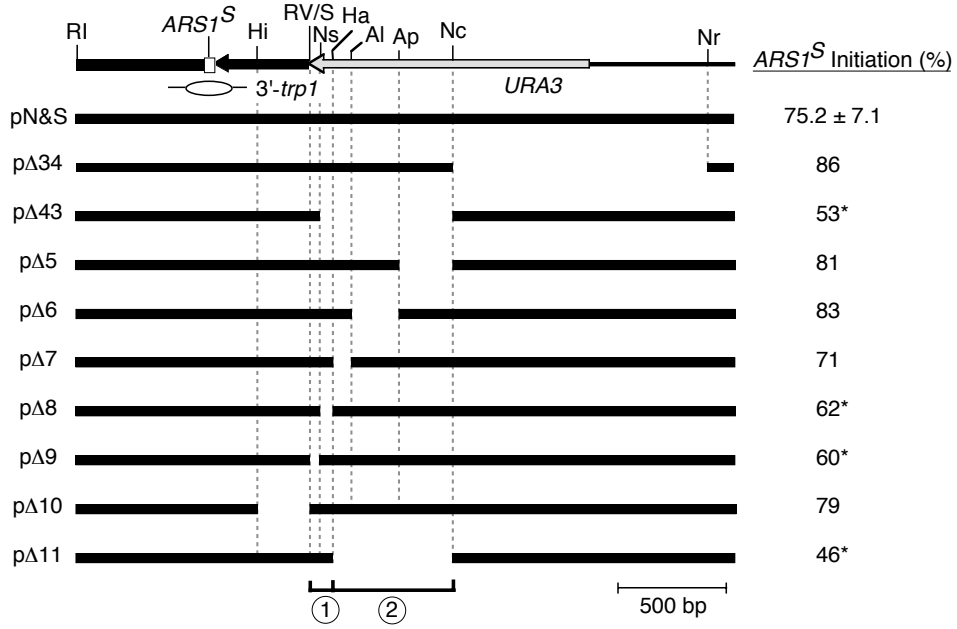

RI
ARS1S
Hi
RV/S
Ns
Ha
AI
Ap
Nc
Nr
3'-trp1
URA3
ARS1S Initiation (%)
pN&S
75.2 ± 7.1
pΔ34
86
pΔ43
53*
pΔ5
81
pΔ6
83
pΔ7
71
pΔ8
62*
pΔ9
60*
pΔ10
79
pΔ11
46*
1
2
500 bp

Figure 4.2: Deletion analysis of 3' *URA3* sequences important for the initiation bias. A map of the southern region of pN&S including $ARS1^S$ and the *URA3* gene is shown at *top*. A bubble indicates the position of $ARS1^S$. Values for percent initiation at $ARS1^S$ are given for pN&S and nine deletion derivatives. Values that are significantly reduced as compared to the value obtained for pN&S on five separate experiments are indicated by (*). Deletion constructs are as follows: (pΔ34) Δ*Nco*I-*Nru*I (Brewer and Fangman 1994); (pΔ43) Δ*Nsi*I-*Nco*I (Brewer and Fangman 1994); (pΔ5) Δ*Apa*I-*Nco*I; (pΔ6) Δ*Alw*NI-*Apa*I; (pΔ7) Δ*Hae*III-*Alw*NI; (pΔ8) Δ*Nsi*I-*Hae*III; (pΔ9) Δ*Sma*I-*Nsi*I; (pΔ10) Δ*Hin*dIII-*Eco*RV; (pΔ11) Δ*Hae*III-*Nco*I. Restriction sites are as follows: (Al) *Alw*NI; (Ap) *Apa*I; (RI) *Eco*RI; (RV) *Eco*RV; (Ha) *Hae*III; (Hi) *Hin*dIII; (Nc) *Nco*I; (Nr) *Nru*I; (Ns) *Nsi*I; (S) *Sma*I.

FIGURE 4.3

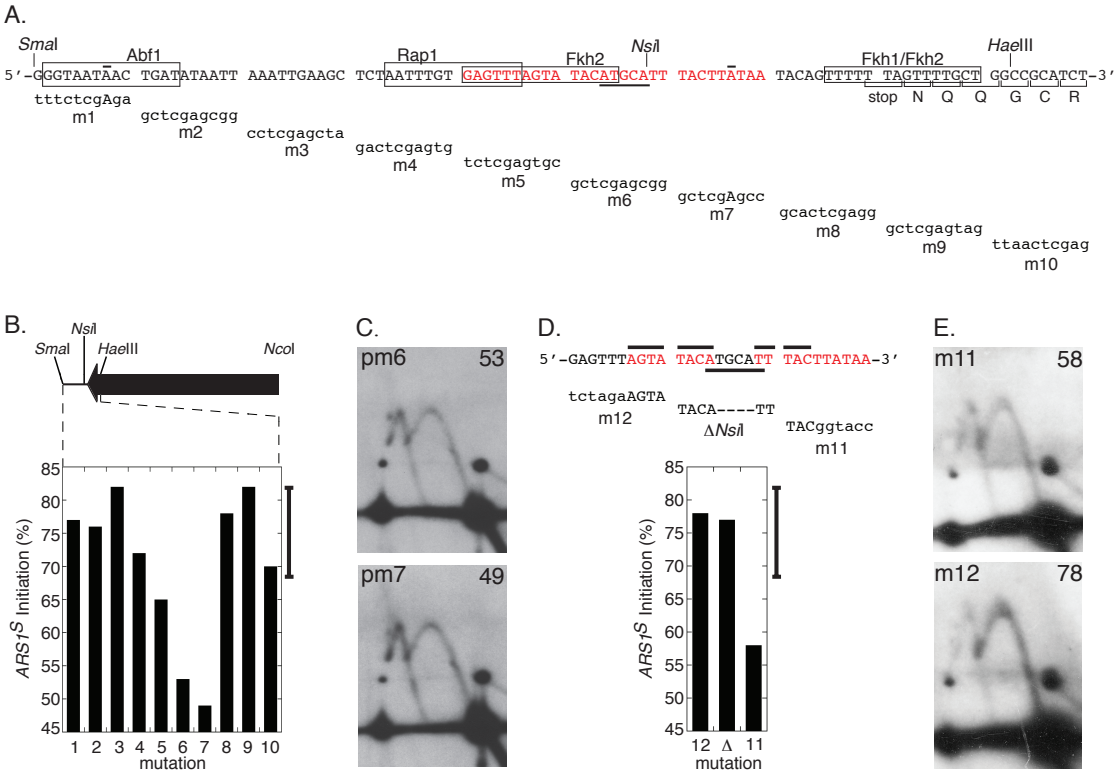
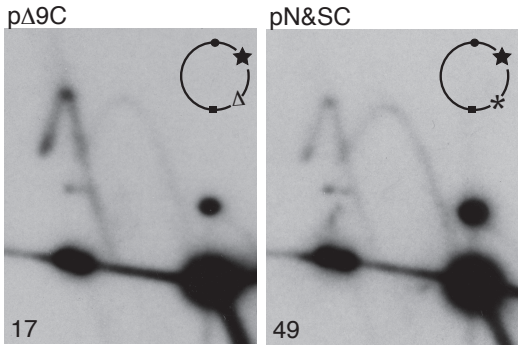

A.
SmaI
Abf1
Rap1
Fkh2
NsiI
Fkh1/Fkh2
HaeIII
5′-GGGTAATAAC TGATATAATT AAATTGAAGC TCTAATTTGT GAGTTTAGTA TACATGCATT TACTTATAA TACAGTTTTT TAGTTTTGCT GGCCGCATCT-3′
stop N Q Q G C R
tttctcgAga m1
gctcgagcgg m2
cctcgagcta m3
gactcgagtg m4
tctcgagtgc m5
gctcgagcgg m6
gctcgAgcc m7
gcactcgagg m8
gctcgagtag m9
ttaactcgag m10
B.
SmaI
NsiI
HaeIII
NcoI
ARS1S Initiation (%)
mutation
C.
pm6 53
pm7 49
D.
5′-GAGTTTAGTA TACATGCATT TACTTATAA-3′
tctagaAGTA m12
TACA----TT ΔNsiI
TACggtacc m11
ARS1S Initiation (%)
12 Δ 11
mutation
E.
m11 58
m12 78

Figure 4.3: Scanning mutagenesis analysis for sequences involved in initiation bias. (A) The wild-type sequence of the *Sma*I-*Hae*III fragment of *URA3* is shown in capital letters. Sequences of the mutations are given in lowercase letters beneath the wild-type sequences they replace. The two overlined nucleotides were not changed by the mutagenesis. The boxes identify 11/13 bp matches to Rap1 or Abf1 consensus binding sites as well as matches to Fkh1 and Fkh2. The 3' end of the *URA3* coding sequence is indicated by the single letter amino acid code. The locations of the *Sma*I, *Nsi*I, and *Hae*III restriction enzyme sites are indicated. Sequences that when mutated significantly reduce the $ARS1^S$ bias are indicated in red. **(B)** Values for percent initiation at $ARS1^S$ are plotted for the 10 mutations that span the *Sma*I-*Hae*III fragment. The bar on the right of the graph indicates a 99% confidence interval extending on both sides of the mean percent initiation at $ARS1^S$ for pN&S based on five separate experiments. Values that fall outside of this range are considered significantly different from pN&S at the 1% level. The approximate locations of the mutations relative to the 3' half of the *URA3* gene (arrow) are shown. **(C)** Fork-direction gel analysis of mutants m6 and m7 as examples of the mutations that eliminate the origin bias. **(D)** The wild-type sequence corresponding to the sequences of the *URA3* fragment that were mutated in m5, m6, and m7 are shown in capital letters. Sequences of three different mutations (m12, Δ*Nsi*I, m11) are given in lowercase letters beneath the wild-type sequences they replace. Deleted bases are depicted by a dash. The overlined nucleotides were not changed by the mutagenesis while the underlined nucleotides mark the *Nsi*I site. Sequences important for the bias are indicated in red letters. Percent initiation at $ARS1^S$ is indicated as bar graphs for the three mutations, m11, m12, and Δ*Nsi*I. **(E)** Fork direction 2-D gels of mutants m11 and m12 (see **Figure 4.3D** for sequences). Percent initiation at $ARS1^S$ is indicated in the upper right corners.

FIGURE 4.4

A. Comparison of CEN and *URA3* bias determinants



B. Density Transfer

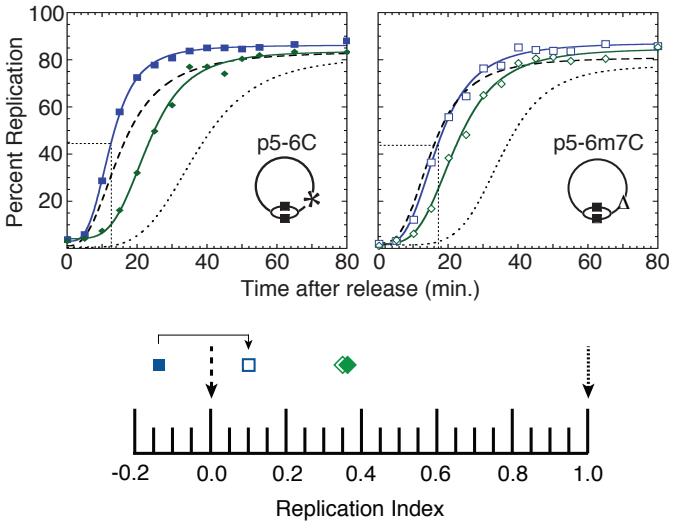


Figure 4.4: Replication kinetics of *ARS1* in the presence and absence of *URA3* sequences that cause biased origin activation. **(A)** Fork direction 2-D gel analysis was performed on plasmids in which the bias determinant is present (pN&SC, cartoon: *) or absent (pΔ9C, cartoon: Δ). Both plasmids contained a centromere near $ARS1^N$ (cartoon: star). Percent initiation at $ARS1^S$ is indicated in the bottom left corners. **(B)** Replication kinetics of stable plasmids containing only $ARS1^S$ from dense isotope transfer experiments. The Trep value for each plasmid is shown as a dotted line intersecting the X-axis. Upper left panel: replication kinetics for a plasmid (p5-6C) containing the wild-type 3' *URA3* sequence that results in preferential use of $ARS1^S$ in the original pN&S plasmid (solid blue squares). Upper right panel: replication kinetics for a plasmid (p5-6m7C) with a 9 bp substitution mutation in the 3' *URA3* sequence that results in equivalent origin use in the original pN&S plasmid (hollow blue squares). The kinetics of plasmid replication are compared with a sequence that replicates early in S phase, *ARS305* (dashed); in mid-S phase, chromosomal *ARS1* (solid or empty green diamonds); and late in S phase, R11 (dotted). The presence (*) or absence (Δ) of the wild type 3' *URA3* sequence is shown on the cartoons of the plasmids. Replication indices for the plasmids and *ARS1* in each strain are shown in the lower panel.

FIGURE 4.5

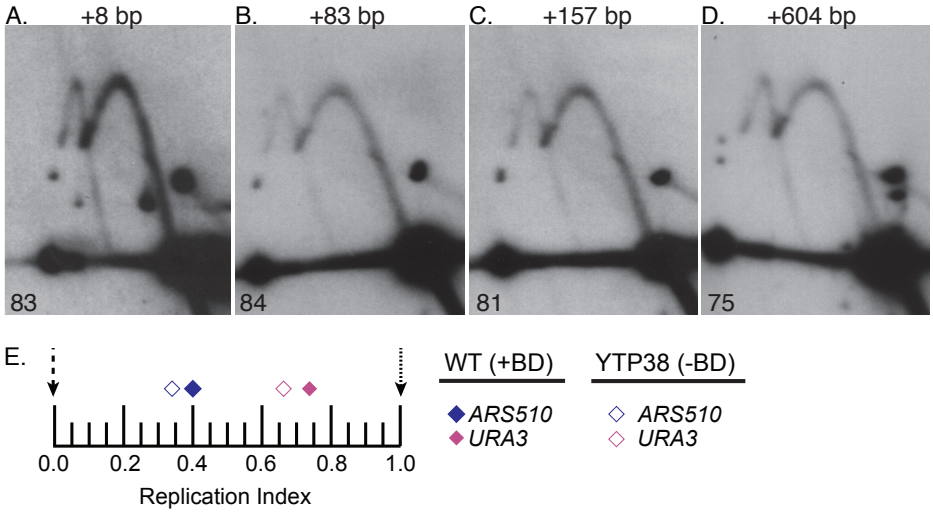

A. +8 bp
B. +83 bp
C. +157 bp
D. +604 bp
83
84
81
75
E.
0.0
0.2
0.4
0.6
0.8
1.0
Replication Index
WT (+BD)
YTP38 (-BD)
ARS510
URA3
ARS510
URA3

Figure 4.5: Determining distances over which the bias determinant can function. (A-D) Fork direction 2-D gel analysis on pN&S plasmids that have various lengths of sequence inserted within the *Nco*I-*Sma*I fragment, between the bias determinant and $ARS1^S$. The amount of sequence that was inserted for each construct is indicated above each image. Percent initiation at $ARS1^S$ is indicated in the lower left corners. **(E)** Replication indices for chromosomal *URA3* (pink) and *ARS510* (blue) in WT cells harboring the bias determinant at its native location (solid diamonds) and in cells lacking the bias determinant (empty diamonds) are plotted. *ARS607* (black dashed arrow) and R11 (black dotted arrow) were used as early and late timing standards, respectively. In the WT strain, *URA3* and *ARS510* had replication indices of 0.73 and 0.4, respectively. In the strain lacking the bias determinant (YTP38), *URA3* had a replication index of 0.66 while the replication index for *ARS510* was 0.34.

FIGURE 4.6

A.

Protein	19 bp	23 bp	Both
Cst6			X
Yap3			X
Pho2			X
Ino4		X	
Sfp1		X	
MATalpha2			X
Yox1	X		
Ino2		X	
Yap7			X
TBP			X
Yap6			X
Fhl1		X	
Spt2			X
Fkh1	X		
Fkh2	X		
Cad1			X
Gzf3			X
Yhp1			X
Arr1	X		

B.

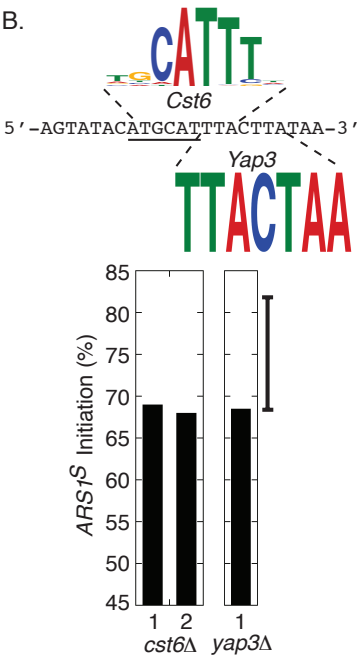


Figure 4.6: Motif search of sequences important for the bias determinant. (A) Both the 23 bp (+TGCA) and the bipartite 19 bp (-TGCA) sequences were run through the motif identification tool TOMTOM[202]. Candidate proteins that were identified for each sequence are listed in order of significance (top to bottom) with proteins common to both searches in the last column. Transcription factors are indicated by red text. **(B)** The WT sequence corresponding to the 23 bps (+TGCA) that are required for the bias is indicated by small black capital letters. The *Nsi*I site is underlined. The logos for Cst6 and Yap3 binding motifs are indicated above or below the corresponding sequence. Percent initiation at $ARS1^S$ on pN&S in cells lacking either *CST6* or *YAP3* are indicated as bar graphs.

Chapter 5

Limiting Factors and the Control of Replication Timing

Over the past two years many important advances, including the work presented in this dissertation, have been made toward deciphering the mechanisms that regulate origin activation time. In 2011, Mantiero et al., discovered that Sld2, Sld3, Dpb11, and Dbf4 (SSDD) are limiting for initiation of DNA replication in *S. cerevisiae*[207]. Sld2 and Sld3 are essential CDK substrates that interact with Dpb11 to recruit the GINS complex and DNA polymerase ε to the pre-RC[208-213]. Dbf4 is the regulatory subunit of DDK, required for the phosphorylation of the MCM helicase and other proteins of the replisome. The authors of this study showed that simultaneous over-expression of all four limiting factors causes late origins to activate earlier in S-phase. The authors suggest that these factors might be limiting in both abundance and in the rate at which they are recycled following origin initiation. Interestingly, just this year, Collart et al., discovered that the replication factors RecQ4, Treslin, Cut5, and Drf1 (the equivalent proteins to SSDD in yeast) become limiting for the initiation of DNA replication at the mid-blastula transition in *Xenopus* embryos[214] suggesting that the role these proteins play in regulating origin initiation time are conserved in higher eukaryotes.

The association of Sld3 with the pre-RC depends on the presence of Cdc45[18,210]. Because Cdc45 and the Sld3 interacting protein, Sld7, preferentially target early activating origins, Mantiero and colleagues investigated whether these proteins could also be limiting for replication initiation[207]. No detectable differences in origin activation time were observed between over expression of SSDD and over expression of SSDD plus Cdc45 and Sld7 (SSDDCS). Presumably this lack of differences between the two strains was due to the still

limiting amounts of Sld3, though this presumption has not been tested. However, the authors did observe an increase in origin efficiency. Data supporting the observation that over expression of SSDDCS leads to increased origin efficiency was obtained by an independent group in a more recent study[215].

The discovery that these limiting factors can regulate origin activation time potentially provides a common mechanistic theme for replication timing control. These data suggest that early replication is determined by mechanisms that can recruit or make a subset of origins available to at least a subset of these liming factors. Conversely, it is logical that mechanisms promoting the exclusion of these limiting factors from origins could impart late activation time on those origins.

Data consistent with this hypothesis was presented earlier this year. Natsume et al., expanded upon the work that is presented in chapter 2 of this dissertation by providing mechanistic evidence into how centromeres are able to regulate origin activation time[192]. The authors of this study show that the kinetochore component and cohesin loading factor, Ctf19, promotes early activation of pericentric origins through the recruitment of DDK and subsequently Sld3/Sld7. Furthermore, the recruitment of these replication factors occurs between telophase and G1[192]. The timing of recruitment is consistent with when origin activation time is thought to be established[108,110]. Currently, it is unclear how recruitment of DDK and Sld3/Sld7 to kinetochores regulates the activation time of origins that can be located up to 20 kb away. It is possible that regulation of origin activation over this distance is achieved through one or more of the mechanisms that are proposed in chapter 2 of this dissertation[188]. Because Dbf4, Dpb11, Sld3 and Sld7 are limiting factors, it is enticing to postulate that DDK recruitment to kinetochores increases the local abundance of these factors near centromeric

regions. Therefore, pericentric origins are more readily available to be acted upon by these limiting initiation factors. It is not known whether other kinetochore proteins are involved in this process; however, it seems plausible that other proteins are involved as Ctf19 is a member of a multi-protein complex that links inner and outer kinetochore components.

Although it has not been tested in great detail, a recent study by Knott et al., indicates that the recruitment of limiting factors may also be important for establishing early activation of non-pericentric origins as well[189]. The authors of this study show that initiation of many of the earliest activating non-pericentric origins is governed by the forkhead transcription factors Fkh1 and Fkh2. Fkh1- and Fkh2-mediated early origin activation is regulated at least in part through the recruitment of Cdc45. Future work is needed to determine how this recruitment occurs. The authors propose that the Fkh1/2 proteins promote the clustering of early origins to regions of the nucleus that are enriched for these replication factors. The mechanisms by which clustering occurs and the extent to which clustering affects origin activation remain a mystery.

All in all, it is likely that the replication limiting factors will be found to play a central role in regulating origin activation time. It is clear that these factors are essential for establishing early timing at pericentric origins in *S. cerevisiae* thus ensuring that centromeres are replicated in early S-phase. Future work will need to be conducted to determine if other kinetochore proteins regulate origin activation time and how recruitment of limiting factors to centromeres regulates distantly positioned pericentric origins. Furthermore, it will be of interest to determine if recruitment of these factors to kinetochores occurs in species that have regional centromeres. As described in chapter 3, future work is needed to determine the importance of early replicating centromeres.

As suggested by the Fkh1/2 study, the recruitment or exclusion of limiting factors may also be important for establishing timing at non-pericentric origins. The observation that the timing of not all origins are influenced by Fkh1/2 in conjunction with the data presented in chapter 4 suggests that additional factors that regulate origin activation time exist. Although the bias determinant that we characterize in chapter 4 has not been directly tested for dependency on Fkh1/2, this sequence may be important for identifying other regulators of origins timing. Future work will need to be conducted to determine what protein(s) bind this sequence and whether or not those proteins can recruit any of the limiting factors. The sequence we identify in chapter 4 contains many imperfect matches to transcription factor binding sites suggesting that there are other transcription factors that regulate origin activation time. As alluded to in the introduction, the B3 element, which stimulates origin activity, can be substituted for many transcription factor binding sites. In light of the forkhead data and our findings detailed in chapter 4, it is possible that transcription factors that function as B3 elements also regulate origin timing. Presumably these transcription factors would interact with at least a subset of the limiting factors.

In summary, the work presented in this thesis lays the groundwork for important advances in the replication field. The data presented in chapters 2 and 4 are the first to conclusively show that there are at least two independent mechanisms that can actively advance origin activation time. Furthermore, the data presented in chapter 3 suggest that the establishment of temporal DNA replication is biologically relevant for cellular health. In the future, it will be interesting to see what other DNA sequences and their associated factors that contribute to the regulation of origin activation time will be discovered. My finding that early centromere replication helps protect against chromosome loss will likely spur research interest in determining why other regions of the genome replicate at specific times. For example, what is

the importance of late telomere and rDNA replication? What is the significance of early and late replicating regions that are internal to the chromosome?

References

1 Bachrati, C. Z. & Hickson, I. D. RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma* **117**, 219-233, doi:10.1007/s00412-007-0142-4 (2008).

2 Burhans, W. C. & Weinberger, M. DNA replication stress, genome instability and aging. *Nucleic Acids Res* **35**, 7545-7556, doi:gkm1059 [pii]10.1093/nar/gkm1059 (2007).

3 Lavin, M. F., Gueven, N. & Grattan-Smith, P. Defective responses to DNA single- and double-strand breaks in spinocerebellar ataxia. *DNA Repair (Amst)* **7**, 1061-1076, doi:S1568-7864(08)00117-1 [pii]10.1016/j.dnarep.2008.03.008 (2008).

4 Ruzankina, Y., Asare, A. & Brown, E. J. Replicative stress, stem cells and aging. *Mech Ageing Dev* **129**, 460-466, doi:S0047-6374(08)00075-4 [pii]10.1016/j.mad.2008.03.009 (2008).

5 Sidorova, J. M., Kehrli, K., Mao, F. & Monnat, R., Jr. Distinct functions of human RECQ helicases WRN and BLM in replication fork recovery and progression after hydroxyurea-induced stalling. *DNA Repair (Amst)* **12**, 128-139, doi:S1568-7864(12)00284-4 [pii]10.1016/j.dnarep.2012.11.005 (2013).

6 Stetson, D. B., Ko, J. S., Heidmann, T. & Medzhitov, R. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* **134**, 587-598, doi:S0092-8674(08)00819-2 [pii]10.1016/j.cell.2008.06.032 (2008).

7 Yang, Y. G., Lindahl, T. & Barnes, D. E. Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease. *Cell* **131**, 873-886, doi:S0092-8674(07)01288-3 [pii]10.1016/j.cell.2007.10.017 (2007).

8 Hand, R. & German, J. A retarded rate of DNA chain growth in Bloom's syndrome. *Proc Natl Acad Sci U S A* **72**, 758-762 (1975).

9 Bicknell, L. S. *et al.* Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat Genet* **43**, 356-359, doi:ng.775 [pii]10.1038/ng.775 (2011).

10 Kerzendorfer, C., Colnaghi, R., Abramowicz, I., Carpenter, G. & O'Driscoll, M. Meier-Gorlin syndrome and Wolf-Hirschhorn syndrome: two developmental disorders highlighting the importance of efficient DNA replication for normal development and neurogenesis. *DNA Repair (Amst)* **12**, 637-644, doi:S1568-7864(13)00097-9 [pii]10.1016/j.dnarep.2013.04.016 (2013).

11 Kerzendorfer, C. *et al.* Characterizing the functional consequences of haploinsufficiency of NELF-A (WHSC2) and SLBP identifies novel cellular phenotypes in Wolf-Hirschhorn syndrome. *Hum Mol Genet* **21**, 2181-2193, doi:dds033 [pii]10.1093/hmg/dds033 (2012).

12 Bell, S. P. & Dutta, A. DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**, 333-374, doi:10.1146/annurev.biochem.71.110601.135425110601.135425 [pii] (2002).

13 Sclafani, R. A. & Holzen, T. M. Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**, 237-280, doi:10.1146/annurev.genet.41.110306.130308 (2007).

14 Takeda, D. Y. & Dutta, A. DNA replication and progression through S phase. *Oncogene* **24**, 2827-2843, doi:1208616 [pii]10.1038/sj.onc.1208616 (2005).

15 Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128-134, doi:10.1038/357128a0 (1992).

16 Houchens, C. R. *et al.* Multiple mechanisms contribute to Schizosaccharomyces pombe origin recognition complex-DNA interactions. *J Biol Chem* **283**, 30216-30224, doi:M802649200 [pii]10.1074/jbc.M802649200 (2008).
17 Vashee, S. *et al.* Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev* **17**, 1894-1908, doi:10.1101/gad.108420317/15/1894 [pii] (2003).
18 Labib, K. & Diffley, J. F. Is the MCM2-7 complex the eukaryotic DNA replication fork helicase? *Curr Opin Genet Dev* **11**, 64-70, doi:S0959-437X(00)00158-1 [pii] (2001).
19 Randell, J. C., Bowers, J. L., Rodriguez, H. K. & Bell, S. P. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol Cell* **21**, 29-39, doi:S1097-2765(05)01811-3 [pii]10.1016/j.molcel.2005.11.023 (2006).
20 Tanaka, S. & Diffley, J. F. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nat Cell Biol* **4**, 198-207, doi:10.1038/ncb757ncb757 [pii] (2002).
21 Whittaker, A. J., Royzman, I. & Orr-Weaver, T. L. Drosophila double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes Dev* **14**, 1765-1776 (2000).
22 Austin, R. J., Orr-Weaver, T. L. & Bell, S. P. Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev* **13**, 2639-2649 (1999).
23 Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type Drosophila origin recognition complex. *Proc Natl Acad Sci U S A* **98**, 11997-12002, doi:10.1073/pnas.211342798211342798 [pii] (2001).
24 Kong, D. & DePamphilis, M. L. Site-specific DNA binding of the Schizosaccharomyces pombe origin recognition complex is determined by the Orc4 subunit. *Mol Cell Biol* **21**, 8095-8103, doi:10.1128/MCB.21.23.8095-8103.2001 (2001).
25 Lee, J. K., Moon, K. Y., Jiang, Y. & Hurwitz, J. The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. *Proc Natl Acad Sci U S A* **98**, 13589-13594, doi:10.1073/pnas.25153039898/24/13589 [pii] (2001).
26 Lee, D. G. & Bell, S. P. Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol Cell Biol* **17**, 7159-7168 (1997).
27 Chuang, R. Y., Chretien, L., Dai, J. & Kelly, T. J. Purification and characterization of the Schizosaccharomyces pombe origin recognition complex: interaction with origin DNA and Cdc18 protein. *J Biol Chem* **277**, 16920-16927, doi:10.1074/jbc.M107710200M107710200 [pii] (2002).
28 Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc Natl Acad Sci U S A* **96**, 2656-2661 (1999).
29 Moon, K. Y., Kong, D., Lee, J. K., Raychaudhuri, S. & Hurwitz, J. Identification and reconstitution of the origin recognition complex from Schizosaccharomyces pombe. *Proc Natl Acad Sci U S A* **96**, 12367-12372 (1999).
30 Tuduri, S., Tourriere, H. & Pasero, P. Defining replication origin efficiency using DNA fiber assays. *Chromosome Res* **18**, 91-102, doi:10.1007/s10577-009-9098-y (2010).
31 Santocanale, C. & Diffley, J. F. ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. *EMBO J* **15**, 6671-6679 (1996).

32 Moyer, S. E., Lewis, P. W. & Botchan, M. R. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A* **103**, 10236-10241, doi:0602400103 [pii]10.1073/pnas.0602400103 (2006).
33 Kubota, Y. *et al.* A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev* **17**, 1141-1152, doi:10.1101/gad.107000317/9/1141 [pii] (2003).
34 Takayama, Y. *et al.* GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* **17**, 1153-1165, doi:10.1101/gad.106590317/9/1153 [pii] (2003).
35 Archambault, V., Ikui, A. E., Drapkin, B. J. & Cross, F. R. Disruption of mechanisms that prevent rereplication triggers a DNA damage response. *Mol Cell Biol* **25**, 6707-6721, doi:25/15/6707 [pii]10.1128/MCB.25.15.6707-6721.2005 (2005).
36 Green, B. M. & Li, J. J. Loss of rereplication control in Saccharomyces cerevisiae results in extensive DNA damage. *Mol Biol Cell* **16**, 421-432, doi:E04-09-0833 [pii]10.1091/mbc.E04-09-0833 (2005).
37 Green, B. M., Finn, K. J. & Li, J. J. Loss of DNA replication control is a potent inducer of gene amplification. *Science* **329**, 943-946, doi:329/5994/943 [pii]10.1126/science.1190966 (2010).
38 Drury, L. S., Perkins, G. & Diffley, J. F. The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J* **16**, 5966-5976, doi:10.1093/emboj/16.19.5966 (1997).
39 Elsasser, S., Chi, Y., Yang, P. & Campbell, J. L. Phosphorylation controls timing of Cdc6p destruction: A biochemical analysis. *Mol Biol Cell* **10**, 3263-3277 (1999).
40 Gopalakrishnan, V. *et al.* Redundant control of rereplication in fission yeast. *Proc Natl Acad Sci U S A* **98**, 13114-13119, doi:10.1073/pnas.221467598221467598 [pii] (2001).
41 Jallepalli, P. V., Brown, G. W., Muzi-Falconi, M., Tien, D. & Kelly, T. J. Regulation of the replication initiator protein p65cdc18 by CDK phosphorylation. *Genes Dev* **11**, 2767-2779 (1997).
42 Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068-1073, doi:10.1038/3508260035082600 [pii] (2001).
43 Liku, M. E., Nguyen, V. Q., Rosales, A. W., Irie, K. & Li, J. J. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol Biol Cell* **16**, 5026-5039, doi:E05-05-0412 [pii]10.1091/mbc.E05-05-0412 (2005).
44 Mendez, J. *et al.* Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol Cell* **9**, 481-491, doi:S1097276502004677 [pii] (2002).
45 Sugimoto, N. *et al.* Cdt1 phosphorylation by cyclin A-dependent kinases negatively regulates its function without affecting geminin binding. *J Biol Chem* **279**, 19691-19697, doi:10.1074/jbc.M313175200M313175200 [pii] (2004).
46 Coue, M., Kearsey, S. E. & Mechali, M. Chromotin binding, nuclear localization and phosphorylation of Xenopus cdc21 are cell-cycle dependent and associated with the control of initiation of DNA replication. *EMBO J* **15**, 1085-1097 (1996).

47 Fujita, M. *et al.* Cell cycle- and chromatin binding state-dependent phosphorylation of human MCM heterohexameric complexes. A role for cdc2 kinase. *J Biol Chem* **273**, 17095-17101 (1998).
48 Delmolino, L. M., Saha, P. & Dutta, A. Multiple mechanisms regulate subcellular localization of human CDC6. *J Biol Chem* **276**, 26947-26954, doi:10.1074/jbc.M101870200M101870200 [pii] (2001).
49 Saha, P. *et al.* Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase. *Mol Cell Biol* **18**, 2758-2767 (1998).
50 Cook, J. G., Chasse, D. A. & Nevins, J. R. The regulated association of Cdt1 with minichromosome maintenance proteins and Cdc6 in mammalian cells. *J Biol Chem* **279**, 9625-9633, doi:10.1074/jbc.M311933200M311933200 [pii] (2004).
51 McGarry, T. J. & Kirschner, M. W. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**, 1043-1053, doi:S0092-8674(00)81209-X [pii] (1998).
52 Wohlschlegel, J. A. *et al.* Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290**, 2309-2312, doi:10.1126/science.290.5500.2309290/5500/2309 [pii] (2000).
53 Stinchcomb, D. T., Struhl, K. & Davis, R. W. Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39-43 (1979).
54 Broach, J. R. *et al.* Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harb Symp Quant Biol* **47 Pt 2**, 1165-1173 (1983).
55 Di Rienzi, S. C. *et al.* Maintaining replication origins in the face of genomic change. *Genome Res* **22**, 1940-1952, doi:gr.138248.112 [pii]10.1101/gr.138248.112 (2012).
56 Eaton, M. L., Galani, K., Kang, S., Bell, S. P. & MacAlpine, D. M. Conserved nucleosome positioning defines replication origins. *Genes Dev* **24**, 748-753, doi:gad.1913210 [pii]10.1101/gad.1913210 (2010).
57 Fabiani, L., Frontali, L. & Newlon, C. S. Identification of an essential core element and stimulatory sequences in a Kluyveromyces lactis ARS element, KARS101. *Mol Microbiol* **19**, 756-766 (1996).
58 Liachko, I. *et al.* A comprehensive genome-wide map of autonomously replicating sequences in a naive genome. *PLoS Genet* **6**, e1000946, doi:10.1371/journal.pgen.1000946 (2010).
59 Marahrens, Y. & Stillman, B. Replicator dominance in a eukaryotic chromosome. *EMBO J* **13**, 3395-3400 (1994).
60 Theis, J. F. & Newlon, C. S. The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. *Proc Natl Acad Sci U S A* **94**, 10786-10791 (1997).
61 Van Houten, J. V. & Newlon, C. S. Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol Cell Biol* **10**, 3917-3925 (1990).
62 Liachko, I. *et al.* Novel features of ARS selection in budding yeast Lachancea kluyveri. *BMC Genomics* **12**, 633, doi:1471-2164-12-633 [pii]10.1186/1471-2164-12-633 (2011).
63 Feng, W. *et al.* Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. *Nat Cell Biol* **8**, 148-155, doi:ncb1358 [pii]10.1038/ncb1358 (2006).

64 MacAlpine, D. M. & Bell, S. P. A genomic view of eukaryotic DNA replication. *Chromosome Res* **13**, 309-326, doi:10.1007/s10577-005-1508-1 (2005).
65 Nieduszynski, C. A., Hiraga, S., Ak, P., Benham, C. J. & Donaldson, A. D. OriDB: a DNA replication origin database. *Nucleic Acids Res* **35**, D40-46, doi:gkl758 [pii]10.1093/nar/gkl758 (2007).
66 Nieduszynski, C. A., Knox, Y. & Donaldson, A. D. Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev* **20**, 1874-1879, doi:20/14/1874 [pii]10.1101/gad.385306 (2006).
67 Marahrens, Y. & Stillman, B. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**, 817-823 (1992).
68 Rao, H., Marahrens, Y. & Stillman, B. Functional conservation of multiple elements in yeast chromosomal replicators. *Mol Cell Biol* **14**, 7643-7651 (1994).
69 Theis, J. F. & Newlon, C. S. Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. *Mol Cell Biol* **14**, 7652-7659 (1994).
70 Chang, F. *et al.* High-resolution analysis of four efficient yeast replication origins reveals new insights into the ORC and putative MCM binding elements. *Nucleic Acids Res* **39**, 6523-6535, doi:gkr301 [pii]10.1093/nar/gkr301 (2011).
71 Rao, H. & Stillman, B. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc Natl Acad Sci U S A* **92**, 2224-2228 (1995).
72 Rowley, A., Cocker, J. H., Harwood, J. & Diffley, J. F. Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J* **14**, 2631-2641 (1995).
73 DePamphilis, M. L. Replication origins in metazoan chromosomes: fact or fiction? *Bioessays* **21**, 5-16, doi:10.1002/(SICI)1521-1878(199901)21:1<5::AID-BIES2>3.0.CO;2-6 [pii]10.1002/(SICI)1521-1878(199901)21:1<5::AID-BIES2>3.0.CO;2-6 (1999).
74 Huang, R. Y. & Kowalski, D. A DNA unwinding element and an ARS consensus comprise a replication origin within a yeast chromosome. *EMBO J* **12**, 4521-4531 (1993).
75 Huang, R. Y. & Kowalski, D. Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome. *Nucleic Acids Res* **24**, 816-823, doi:5e0780 [pii] (1996).
76 Kowalski, D. & Eddy, M. J. The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. *EMBO J* **8**, 4335-4344 (1989).
77 Natale, D. A., Umek, R. M. & Kowalski, D. Ease of DNA unwinding is a conserved property of yeast replication origins. *Nucleic Acids Res* **21**, 555-560 (1993).
78 Liachko, I., Youngblood, R. A., Keich, U. & Dunham, M. J. High-resolution mapping, characterization, and optimization of autonomously replicating sequences in yeast. *Genome Res* **23**, 698-704, doi:gr.144659.112 [pii]10.1101/gr.144659.112 (2013).
79 Wilmes, G. M. & Bell, S. P. The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. *Proc Natl Acad Sci U S A* **99**, 101-106, doi:10.1073/pnas.012578499012578499 [pii] (2002).

80 Diffley, J. F. & Stillman, B. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc Natl Acad Sci U S A* **85**, 2120-2124 (1988).
81 Eisenberg, S., Civalier, C. & Tye, B. K. Specific interaction between a Saccharomyces cerevisiae protein and a DNA element associated with certain autonomously replicating sequences. *Proc Natl Acad Sci U S A* **85**, 743-746 (1988).
82 Chang, V. K., Donato, J. J., Chan, C. S. & Tye, B. K. Mcm1 promotes replication initiation by binding specific elements at replication origins. *Mol Cell Biol* **24**, 6514-6524, doi:10.1128/MCB.24.14.6514-6524.200424/14/6514 [pii] (2004).
83 Irlbacher, H., Franke, J., Manke, T., Vingron, M. & Ehrenhofer-Murray, A. E. Control of replication initiation and heterochromatin formation in Saccharomyces cerevisiae by a regulator of meiotic gene expression. *Genes Dev* **19**, 1811-1822, doi:19/15/1811 [pii]10.1101/gad.334805 (2005).
84 Sharma, K., Weinberger, M. & Huberman, J. A. Roles for internal and flanking sequences in regulating the activity of mating-type-silencer-associated replication origins in Saccharomyces cerevisiae. *Genetics* **159**, 35-45 (2001).
85 Walker, S. S., Francesconi, S. C. & Eisenberg, S. A DNA replication enhancer in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **87**, 4665-4669 (1990).
86 Chang, F. *et al.* Analysis of chromosome III replicators reveals an unusual structure for the ARS318 silencer origin and a conserved WTW sequence within the origin recognition complex binding site. *Mol Cell Biol* **28**, 5071-5081, doi:MCB.00206-08 [pii]10.1128/MCB.00206-08 (2008).
87 Chang, V. K. *et al.* Mcm1 binds replication origins. *J Biol Chem* **278**, 6093-6100, doi:10.1074/jbc.M209827200M209827200 [pii] (2003).
88 Segurado, M., de Luis, A. & Antequera, F. Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe. *EMBO Rep* **4**, 1048-1053, doi:10.1038/sj.embor.embor7400008embor7400008 [pii] (2003).
89 Clyne, R. K. & Kelly, T. J. Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe. *EMBO J* **14**, 6348-6357 (1995).
90 Dai, J., Chuang, R. Y. & Kelly, T. J. DNA replication origins in the Schizosaccharomyces pombe genome. *Proc Natl Acad Sci U S A* **102**, 337-342, doi:0408811102 [pii]10.1073/pnas.0408811102 (2005).
91 Dubey, D. D., Kim, S. M., Todorov, I. T. & Huberman, J. A. Large, complex modular structure of a fission yeast DNA replication origin. *Curr Biol* **6**, 467-473, doi:S0960-9822(02)00514-6 [pii] (1996).
92 Heichinger, C., Penkett, C. J., Bahler, J. & Nurse, P. Genome-wide characterization of fission yeast DNA replication origins. *EMBO J* **25**, 5171-5179, doi:7601390 [pii]10.1038/sj.emboj.7601390 (2006).
93 Patel, P. K., Arcangioli, B., Baker, S. P., Bensimon, A. & Rhind, N. DNA replication origins fire stochastically in fission yeast. *Mol Biol Cell* **17**, 308-316, doi:E05-07-0657 [pii]10.1091/mbc.E05-07-0657 (2006).
94 Okuno, Y., Satoh, H., Sekiguchi, M. & Masukata, H. Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. *Mol Cell Biol* **19**, 6699-6709 (1999).
95 Kitsberg, D., Selig, S., Keshet, I. & Cedar, H. Replication structure of the human beta-globin gene domain. *Nature* **366**, 588-590, doi:10.1038/366588a0 (1993).

96 Wang, L. *et al.* The human beta-globin replication initiation region consists of two modular independent replicators. *Mol Cell Biol* **24**, 3373-3386 (2004).
97 Paixao, S. *et al.* Modular structure of the human lamin B2 replicator. *Mol Cell Biol* **24**, 2958-2967 (2004).
98 Zhang, H. & Tower, J. Sequence requirements for function of the Drosophila chorion gene locus ACE3 replicator and ori-beta origin elements. *Development* **131**, 2089-2099, doi:10.1242/dev.01064131/9/2089 [pii] (2004).
99 Dijkwel, P. A., Wang, S. & Hamlin, J. L. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. *Mol Cell Biol* **22**, 3053-3065 (2002).
100 Dijkwel, P. A., Vaughn, J. P. & Hamlin, J. L. Replication initiation sites are distributed widely in the amplified CHO dihydrofolate reductase domain. *Nucleic Acids Res* **22**, 4989-4996 (1994).
101 Vaughn, J. P., Dijkwel, P. A. & Hamlin, J. L. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. *Cell* **61**, 1075-1087, doi:0092-8674(90)90071-L [pii] (1990).
102 Burhans, W. C., Vassilev, L. T., Caddle, M. S., Heintz, N. H. & DePamphilis, M. L. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. *Cell* **62**, 955-965, doi:0092-8674(90)90270-O [pii] (1990).
103 Taylor, J. H. Asynchronous duplication of chromosomes in cultured cells of Chinese hamster. *J Biophys Biochem Cytol* **7**, 455-464 (1960).
104 Cayrou, C. *et al.* Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Res* **21**, 1438-1449, doi:gr.121830.111 [pii]10.1101/gr.121830.111 (2011).
105 Hiratani, I. *et al.* Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* **6**, e245, doi:08-PLBI-RA-1977 [pii]10.1371/journal.pbio.0060245 (2008).
106 Raghuraman, M. K. *et al.* Replication dynamics of the yeast genome. *Science* **294**, 115-121, doi:10.1126/science.294.5540.115294/5540/115 [pii] (2001).
107 Yabuki, N., Terashima, H. & Kitada, K. Mapping of early firing origins on a replication profile of budding yeast. *Genes Cells* **7**, 781-789, doi:559 [pii] (2002).
108 Dimitrova, D. S. & Gilbert, D. M. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol Cell* **4**, 983-993, doi:S1097-2765(00)80227-0 [pii] (1999).
109 Li, F. *et al.* The replication timing program of the Chinese hamster beta-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *J Cell Biol* **154**, 283-292 (2001).
110 Raghuraman, M. K., Brewer, B. J. & Fangman, W. L. Cell cycle-dependent establishment of a late replication program. *Science* **276**, 806-809 (1997).
111 Ferguson, B. M. & Fangman, W. L. A position effect on the time of replication origin activation in yeast. *Cell* **68**, 333-339, doi:0092-8674(92)90474-Q [pii] (1992).
112 Friedman, K. L. *et al.* Multiple determinants controlling activation of yeast replication origins late in S phase. *Genes Dev* **10**, 1595-1607 (1996).
113 Cadoret, J. C. *et al.* Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proc Natl Acad Sci U S A* **105**, 15837-15842, doi:0805208105 [pii]10.1073/pnas.0805208105 (2008).

114 Farkash-Amar, S. *et al.* Global organization of replication time zones of the mouse genome. *Genome Res* **18**, 1562-1570, doi:gr.079566.108 [pii]10.1101/gr.079566.108 (2008).
115 MacAlpine, D. M., Rodriguez, H. K. & Bell, S. P. Coordination of replication and transcription along a Drosophila chromosome. *Genes Dev* **18**, 3094-3105, doi:18/24/3094 [pii]10.1101/gad.1246404 (2004).
116 Schubeler, D. *et al.* Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. *Nat Genet* **32**, 438-442, doi:10.1038/ng1005ng1005 [pii] (2002).
117 Hiratani, I. & Gilbert, D. M. Replication timing as an epigenetic mark. *Epigenetics* **4**, 93-97, doi:7772 [pii] (2009).
118 Woodfine, K. *et al.* Replication timing of the human genome. *Hum Mol Genet* **13**, 191-202, doi:10.1093/hmg/ddh016ddh016 [pii] (2004).
119 Gilbert, D. M. Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* **14**, 377-383, doi:S0955067402003265 [pii] (2002).
120 Aparicio, J. G., Viggiani, C. J., Gibson, D. G. & Aparicio, O. M. The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae. *Mol Cell Biol* **24**, 4769-4780, doi:10.1128/MCB.24.11.4769-4780.200424/11/4769 [pii] (2004).
121 Knott, S. R., Viggiani, C. J., Tavare, S. & Aparicio, O. M. Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae. *Genes Dev* **23**, 1077-1090, doi:23/9/1077 [pii]10.1101/gad.1784309 (2009).
122 Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J. & Grunstein, M. Histone acetylation regulates the time of replication origin firing. *Mol Cell* **10**, 1223-1233, doi:S1097276502007025 [pii] (2002).
123 Bickmore, W. A. & Carothers, A. D. Factors affecting the timing and imprinting of replication on a mammalian chromosome. *J Cell Sci* **108 (Pt 8)**, 2801-2809 (1995).
124 Weinreich, M., Palacios DeBeer, M. A. & Fox, C. A. The activities of eukaryotic replication origins in chromatin. *Biochim Biophys Acta* **1677**, 142-157, doi:10.1016/j.bbaexp.2003.11.015S0167478103002860 [pii] (2004).
125 Venditti, P., Costanzo, G., Negri, R. & Camilloni, G. ABFI contributes to the chromatin organization of Saccharomyces cerevisiae ARS1 B-domain. *Biochim Biophys Acta* **1219**, 677-689, doi:0167-4781(94)90227-5 [pii] (1994).
126 Simpson, R. T. Nucleosome positioning can affect the function of a cis-acting DNA element in vivo. *Nature* **343**, 387-389, doi:10.1038/343387a0 (1990).
127 Aggarwal, B. D. & Calvi, B. R. Chromatin regulates origin activity in Drosophila follicle cells. *Nature* **430**, 372-376, doi:10.1038/nature02694nature02694 [pii] (2004).
128 Miotto, B. & Struhl, K. HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. *Mol Cell* **37**, 57-66, doi:S1097-2765(09)00916-2 [pii]10.1016/j.molcel.2009.12.012 (2010).
129 Koren, A. *et al.* Epigenetically-inherited centromere and neocentromere DNA replicates earliest in S-phase. *PLoS Genet* **6**, e1001068, doi:10.1371/journal.pgen.1001068 (2010).

130 McCarroll, R. M. & Fangman, W. L. Time of replication of yeast centromeres and telomeres. *Cell* **54**, 505-513 (1988).
131 Bianchi, A. & Shore, D. Early replication of short telomeres in budding yeast. *Cell* **128**, 1051-1062, doi:S0092-8674(07)00242-5 [pii]10.1016/j.cell.2007.01.041 (2007).
132 Cosgrove, A. J., Nieduszynski, C. A. & Donaldson, A. D. Ku complex controls the replication time of DNA in telomere regions. *Gene Dev* **16**, 2485-2490, doi:Doi 10.1101/Gad.231602 (2002).
133 Stevenson, J. B. & Gottschling, D. E. Telomeric chromatin modulates replication timing near chromosome ends. *Genes Dev* **13**, 146-151 (1999).
134 Lian, H. Y. *et al.* The effect of Ku on telomere replication time is mediated by telomere length but is independent of histone tail acetylation. *Mol Biol Cell* **22**, 1753-1765, doi:mbc.E10-06-0549 [pii]10.1091/mbc.E10-06-0549 (2011).
135 Hayano, M. *et al.* Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes Dev* **26**, 137-150, doi:26/2/137 [pii]10.1101/gad.178491.111 (2012).
136 Cornacchia, D. *et al.* Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J* **31**, 3678-3690, doi:emboj2012214 [pii]10.1038/emboj.2012.214 (2012).
137 Yamazaki, S. *et al.* Rif1 regulates the replication timing domains on the human genome. *EMBO J* **31**, 3667-3677, doi:emboj2012180 [pii]10.1038/emboj.2012.180 (2012).
138 Yompakdee, C. & Huberman, J. A. Enforcement of late replication origin firing by clusters of short G-rich DNA sequences. *J Biol Chem* **279**, 42337-42344, doi:10.1074/jbc.M407552200M407552200 [pii] (2004).
139 Ahmad, K. & Henikoff, S. Centromeres are specialized replication domains in heterochromatin. *J Cell Biol* **153**, 101-110 (2001).
140 Kim, S. M., Dubey, D. D. & Huberman, J. A. Early-replicating heterochromatin. *Genes Dev* **17**, 330-335, doi:10.1101/gad.1046203 (2003).
141 Weidtkamp-Peters, S., Rahn, H. P., Cardoso, M. C. & Hemmerich, P. Replication of centromeric heterochromatin in mouse fibroblasts takes place in early, middle, and late S phase. *Histochem Cell Biol* **125**, 91-102, doi:10.1007/s00418-005-0063-3 (2006).
142 Hegemann, J. H. & Fleig, U. N. The centromere of budding yeast. *Bioessays* **15**, 451-460, doi:10.1002/bies.950150704 (1993).
143 Jallepalli, P. V. & Lengauer, C. Chromosome segregation and cancer: cutting through the mystery. *Nat Rev Cancer* **1**, 109-117, doi:10.1038/35101065 (2001).
144 McGrew, J., Diehl, B. & Fitzgerald-Hayes, M. Single base-pair mutations in centromere element III cause aberrant chromosome segregation in Saccharomyces cerevisiae. *Mol Cell Biol* **6**, 530-538 (1986).
145 Alvino, G. M. *et al.* Replication in hydroxyurea: it's a matter of time. *Mol Cell Biol* **27**, 6396-6406, doi:MCB.00719-07 [pii]10.1128/MCB.00719-07 (2007).
146 McCune, H. J. *et al.* The temporal program of chromosome replication: genomewide replication in clb5Δ Saccharomyces cerevisiae. *Genetics* **180**, 1833-1847, doi:genetics.108.094359 [pii]10.1534/genetics.108.094359 (2008).

147 Feng, W., Bachant, J., Collingwood, D., Raghuraman, M. K. & Brewer, B. J. Centromere replication timing determines different forms of genomic instability in Saccharomyces cerevisiae checkpoint mutants during replication stress. *Genetics* **183**, 1249-1260, doi:genetics.109.107508 [pii]10.1534/genetics.109.107508 (2009).
148 Furuyama, S. & Biggins, S. Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc Natl Acad Sci U S A* **104**, 14706-14711, doi:0706985104 [pii]10.1073/pnas.0706985104 (2007).
149 Tanaka, K. *et al.* Molecular mechanisms of kinetochore capture by spindle microtubules. *Nature* **434**, 987-994, doi:nature03483 [pii]10.1038/nature03483 (2005).
150 Hayashi, M. *et al.* Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. *EMBO J* **26**, 1327-1339, doi:7601585 [pii]10.1038/sj.emboj.7601585 (2007).
151 Clarke, L. Centromeres of budding and fission yeasts. *Trends Genet* **6**, 150-154 (1990).
152 Fitzgerald-Hayes, M., Clarke, L. & Carbon, J. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* **29**, 235-244, doi:0092-8674(82)90108-8 [pii] (1982).
153 Santaguida, S. & Musacchio, A. The life and miracles of kinetochores. *EMBO J* **28**, 2511-2531, doi:emboj2009173 [pii]10.1038/emboj.2009.173 (2009).
154 Ferguson, B. M., Brewer, B. J., Reynolds, A. E. & Fangman, W. L. A yeast origin of replication is activated late in S phase. *Cell* **65**, 507-515, doi:0092-8674(91)90468-E [pii] (1991).
155 Brewer, B. J. & Fangman, W. L. Initiation preference at a yeast origin of replication. *Proc Natl Acad Sci U S A* **91**, 3418-3422 (1994).
156 Kolor, K. Advancement of the timing of an origin activation by a cis-acting DNA element. *PhD Dissertation, University of Washington* (1997).
157 Spell, R. M. & Holm, C. Nature and distribution of chromosomal intertwinings in Saccharomyces cerevisiae. *Mol Cell Biol* **14**, 1465-1476 (1994).
158 Brewer, B. J. & Fangman, W. L. The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell* **51**, 463-471, doi:0092-8674(87)90642-8 [pii] (1987).
159 Donaldson, A. D. *et al.* CLB5-dependent activation of late replication origins in S. cerevisiae. *Mol Cell* **2**, 173-182, doi:S1097-2765(00)80127-6 [pii] (1998).
160 Lechner, J. & Ortiz, J. The Saccharomyces cerevisiae kinetochore. *FEBS Lett* **389**, 70-74, doi:0014-5793(96)00563-7 [pii] (1996).
161 Duan, Z. *et al.* A three-dimensional model of the yeast genome. *Nature* **465**, 363-367, doi:nature08973 [pii]10.1038/nature08973 (2010).
162 Kitamura, E., Blow, J. J. & Tanaka, T. U. Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. *Cell* **125**, 1297-1308, doi:S0092-8674(06)00652-0 [pii]10.1016/j.cell.2006.04.041 (2006).
163 Jin, Q., Trelles-Sticken, E., Scherthan, H. & Loidl, J. Yeast nuclei display prominent centromere clustering that is reduced in nondividing cells and in meiotic prophase. *J Cell Biol* **141**, 21-29 (1998).

164 Heath, C. V. *et al.* Nuclear pore complex clustering and nuclear accumulation of poly(A)+ RNA associated with mutation of the Saccharomyces cerevisiae RAT2/NUP120 gene. *J Cell Biol* **131**, 1677-1697 (1995).
165 Winey, M., Yarar, D., Giddings, T. H., Jr. & Mastronarde, D. N. Nuclear pore complex number and distribution throughout the Saccharomyces cerevisiae cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. *Mol Biol Cell* **8**, 2119-2132 (1997).
166 Heun, P., Laroche, T., Raghuraman, M. K. & Gasser, S. M. The positioning and dynamics of origins of replication in the budding yeast nucleus. *J Cell Biol* **152**, 385-400 (2001).
167 Ebrahimi, H. *et al.* Early initiation of a replication origin tethered at the nuclear periphery. *J Cell Sci* **123**, 1015-1019, doi:jcs.060392 [pii]10.1242/jcs.060392 (2010).
168 Hiraga, S., Robertson, E. D. & Donaldson, A. D. The Ctf18 RFC-like complex positions yeast telomeres but does not specify their replication time. *EMBO J* **25**, 1505-1514, doi:7601038 [pii]10.1038/sj.emboj.7601038 (2006).
169 Bouck, D. C., Joglekar, A. P. & Bloom, K. S. Design features of a mitotic spindle: balancing tension and compression at a single microtubule kinetochore interface in budding yeast. *Annu Rev Genet* **42**, 335-359, doi:10.1146/annurev.genet.42.110807.091620 (2008).
170 Akiyoshi, B. *et al.* Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature* **468**, 576-579, doi:nature09594 [pii]10.1038/nature09594 (2010).
171 Fernius, J. & Marston, A. L. Establishment of cohesion at the pericentromere by the Ctf19 kinetochore subcomplex and the replication fork-associated factor, Csm3. *PLoS Genet* **5**, e1000629, doi:10.1371/journal.pgen.1000629 (2009).
172 Ranjitkar, P. *et al.* An E3 ubiquitin ligase prevents ectopic localization of the centromeric histone H3 variant via the centromere targeting domain. *Mol Cell* **40**, 455-464, doi:S1097-2765(10)00753-7 [pii]10.1016/j.molcel.2010.09.025 (2010).
173 Pietrasanta, L. I. *et al.* Probing the Saccharomyces cerevisiae centromeric DNA (CEN DNA)-binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy. *Proc Natl Acad Sci U S A* **96**, 3757-3762 (1999).
174 Yeh, E. *et al.* Pericentric chromatin is organized into an intramolecular loop in mitosis. *Curr Biol* **18**, 81-90, doi:S0960-9822(07)02415-3 [pii]10.1016/j.cub.2007.12.019 (2008).
175 Anderson, M., Haase, J., Yeh, E. & Bloom, K. Function and assembly of DNA looping, clustering, and microtubule attachment complexes within a eukaryotic kinetochore. *Mol Biol Cell* **20**, 4131-4139, doi:E09-05-0359 [pii]10.1091/mbc.E09-05-0359 (2009).
176 Megee, P. C., Mistrot, C., Guacci, V. & Koshland, D. The centromeric sister chromatid cohesion site directs Mcd1p binding to adjacent sequences. *Mol Cell* **4**, 445-450, doi:S1097-2765(00)80347-0 [pii] (1999).
177 Tanaka, T., Cosma, M. P., Wirth, K. & Nasmyth, K. Identification of cohesin association sites at centromeres and along chromosome arms. *Cell* **98**, 847-858, doi:S0092-8674(00)81518-4 [pii] (1999).

178 Weber, S. A. *et al.* The kinetochore is an enhancer of pericentric cohesin binding. *PLoS Biol* **2**, E260, doi:10.1371/journal.pbio.0020260 (2004).
179 Tanaka, Y. *et al.* Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. *EMBO J* **20**, 6612-6618, doi:10.1093/emboj/20.23.6612 (2001).
180 Laible, M. & Boonrod, K. Homemade site directed mutagenesis of whole plasmids. *J Vis Exp* **27**, 1135, doi:1135 [pii]10.3791/1135 (2009).
181 Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355-360, doi:10.1002/yea.320110408 (1995).
182 Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**, 339-346 (1989).
183 Brewer, B. J. & Fangman, W. L. Mapping replication origins in yeast chromosomes. *Bioessays* **13**, 317-322, doi:10.1002/bies.950130702 (1991).
184 Brewer, B. J., Lockshon, D. & Fangman, W. L. The arrest of replication forks in the rDNA of yeast occurs independently of transcription. *Cell* **71**, 267-276, doi:0092-8674(92)90355-G [pii] (1992).
185 Huberman, J. A., Spotila, L. D., Nawotka, K. A., el-Assouli, S. M. & Davis, L. R. The in vivo replication origin of the yeast 2 μm plasmid. *Cell* **51**, 473-481, doi:0092-8674(87)90643-X [pii] (1987).
186 Hutter, K. J. & Eipel, H. E. Microbial determinations by flow cytometry. *J Gen Microbiol* **113**, 369-375 (1979).
187 Muller, C. A. & Nieduszynski, C. A. Conservation of replication timing reveals global and local regulation of replication origin activity. *Genome Res* **22**, 1953-1962, doi:gr.139477.112 [pii]10.1101/gr.139477.112 (2012).
188 Pohl, T. J., Brewer, B. J. & Raghuraman, M. K. Functional centromeres determine the activation time of pericentric origins of DNA replication in Saccharomyces cerevisiae. *PLoS Genet* **8**, e1002677, doi:10.1371/journal.pgen.1002677PGENETICS-D-11-02710 [pii] (2012).
189 Knott, S. R. *et al.* Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. *Cell* **148**, 99-111, doi:S0092-8674(11)01515-7 [pii] 10.1016/j.cell.2011.12.012 (2012).
190 Blitzblau, H. G., Chan, C. S., Hochwagen, A. & Bell, S. P. Separation of DNA replication from the assembly of break-competent meiotic chromosomes. *PLoS Genet* **8**, e1002643, doi:10.1371/journal.pgen.1002643PGENETICS-D-11-02760 [pii] (2012).
191 Kitamura, E., Tanaka, K., Kitamura, Y. & Tanaka, T. U. Kinetochore microtubule interaction during S phase in Saccharomyces cerevisiae. *Genes Dev* **21**, 3319-3330, doi:21/24/3319 [pii]10.1101/gad.449407 (2007).
192 Natsume, T. *et al.* Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol Cell* **50**, 661-674, doi:S1097-2765(13)00374-2 [pii]10.1016/j.molcel.2013.05.011 (2013).
193 Rosche, W. A. & Foster, P. L. Determining mutation rates in bacterial populations. *Methods* **20**, 4-17, doi:10.1006/meth.1999.0901S1046-2023(99)90901-5 [pii] (2000).
194 Luria, S. E. & Delbruck, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **28**, 491-511 (1943).

195 Fangman, W. L. & Brewer, B. J. A question of time: replication origins of eukaryotic chromosomes. *Cell* **71**, 363-366, doi:0092-8674(92)90505-7 [pii] (1992).
196 Wellinger, R. J., Wolf, A. J. & Zakian, V. A. Origin activation and formation of single-strand TG1-3 tails occur sequentially in late S phase on a yeast linear plasmid. *Mol Cell Biol* **13**, 4057-4065 (1993).
197 Wellinger, R. J., Wolf, A. J. & Zakian, V. A. Structural and temporal analysis of telomere replication in yeast. *Cold Spring Harb Symp Quant Biol* **58**, 725-732 (1993).
198 Rivin, C. J. & Fangman, W. L. Replication fork rate and origin activation during the S phase of Saccharomyces cerevisiae. *J Cell Biol* **85**, 108-115 (1980).
199 Friedman, K. L. & Brewer, B. J. Analysis of replication intermediates by two-dimensional agarose gel electrophoresis. *Methods Enzymol* **262**, 613-627, doi:0076-6879(95)62048-6 [pii] (1995).
200 Newlon, C. S. *DNA replication in yeast*. 873-914 (Cold Spring Harbor Press, 1996).
201 Walker, S. S., Francesconi, S. C., Tye, B. K. & Eisenberg, S. The OBF1 protein and its DNA-binding site are important for the function of an autonomously replicating sequence in Saccharomyces cerevisiae. *Mol Cell Biol* **9**, 2914-2921 (1989).
202 Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Genome Biol* **8**, R24, doi:gb-2007-8-2-r24 [pii]10.1186/gb-2007-8-2-r24 (2007).
203 Natale, D. A., Schubert, A. E. & Kowalski, D. DNA helical stability accounts for mutational defects in a yeast replication origin. *Proc Natl Acad Sci U S A* **89**, 2654-2658 (1992).
204 Friedman, K. L., Raghuraman, M. K., Fangman, W. L. & Brewer, B. J. Analysis of the temporal program of replication initiation in yeast chromosomes. *J Cell Sci Suppl* **19**, 51-58 (1995).
205 Deng, W. P. & Nickoloff, J. A. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem* **200**, 8 (1992).
206 Hoffman, C. S. & Winston, F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. *Gene* **57**, 267-272 (1987).
207 Mantiero, D., Mackenzie, A., Donaldson, A. & Zegerman, P. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J* **30**, 4805-4814, doi:emboj2011404 [pii]10.1038/emboj.2011.404 (2011).
208 Kamimura, M. *et al.* Improvements in artery occlusion by low-density lipoprotein apheresis in a patient with peripheral arterial disease. *Ther Apher* **6**, 467-470, doi:tap470 [pii] (2002).
209 Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y. & Araki, H. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes Dev* **24**, 602-612, doi:24/6/602 [pii]10.1101/gad.1883410 (2010).
210 Kamimura, Y., Tak, Y. S., Sugino, A. & Araki, H. Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. *EMBO J* **20**, 2097-2107, doi:10.1093/emboj/20.8.2097 (2001).
211 Kanemaki, M. & Labib, K. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J* **25**, 1753-1763, doi:7601063 [pii]10.1038/sj.emboj.7601063 (2006).

212 Tanaka, S. *et al.* CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* **445**, 328-332, doi:nature05465 [pii]10.1038/nature05465 (2007).
213 Zegerman, P. & Diffley, J. F. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* **445**, 281-285, doi:nature05432 [pii]10.1038/nature05432 (2007).
214 Collart, C., Allen, G. E., Bradshaw, C. R., Smith, J. C. & Zegerman, P. Titration of four replication factors is essential for the Xenopus laevis midblastula transition. *Science* **341**, 893-896, doi:science.1241530 [pii]10.1126/science.1241530 (2013).
215 McGuffee, S. R., Smith, D. J. & Whitehouse, I. Quantitative, genome-wide analysis of eukaryotic replication initiation and termination. *Mol Cell* **50**, 123-135, doi:S1097-2765(13)00207-4 [pii]10.1016/j.molcel.2013.03.004 (2013).

Appendix A: Supplement for Chapter 2

SUPPLEMANTAL FIGURE 2.1

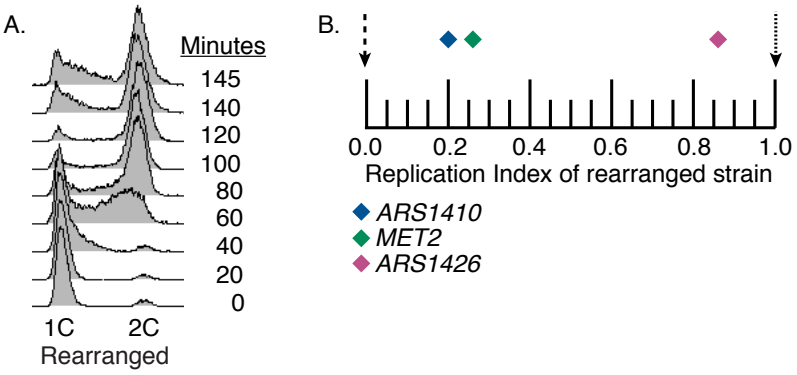


Figure S2.1: Replication kinetic data for an independent segregant of the rearranged strain. (A) Flow cytometry of independent segregant of the rearranged strain. The shift from 1C to 2C DNA content shows that cells entered S-phase at around 40 min and DNA synthesis was complete by 140 minutes (compare to **Figure 2.2B**). **(B)** Replication indices for met2 (green diamond), *ARS1410* (blue diamond), and *ARS1426* (magenta diamond) in the independent segregant of the rearranged strain were 0.26, 0.20, and 0.86, respectively. Timing standards, *ARS306* and R11 are plotted as a black dashed arrow and a black dotted arrow, respectively (compare to **Figure 2.2D**).

SUPPLEMANTAL FIGURE 2.2

Wild Type Replication Profiles

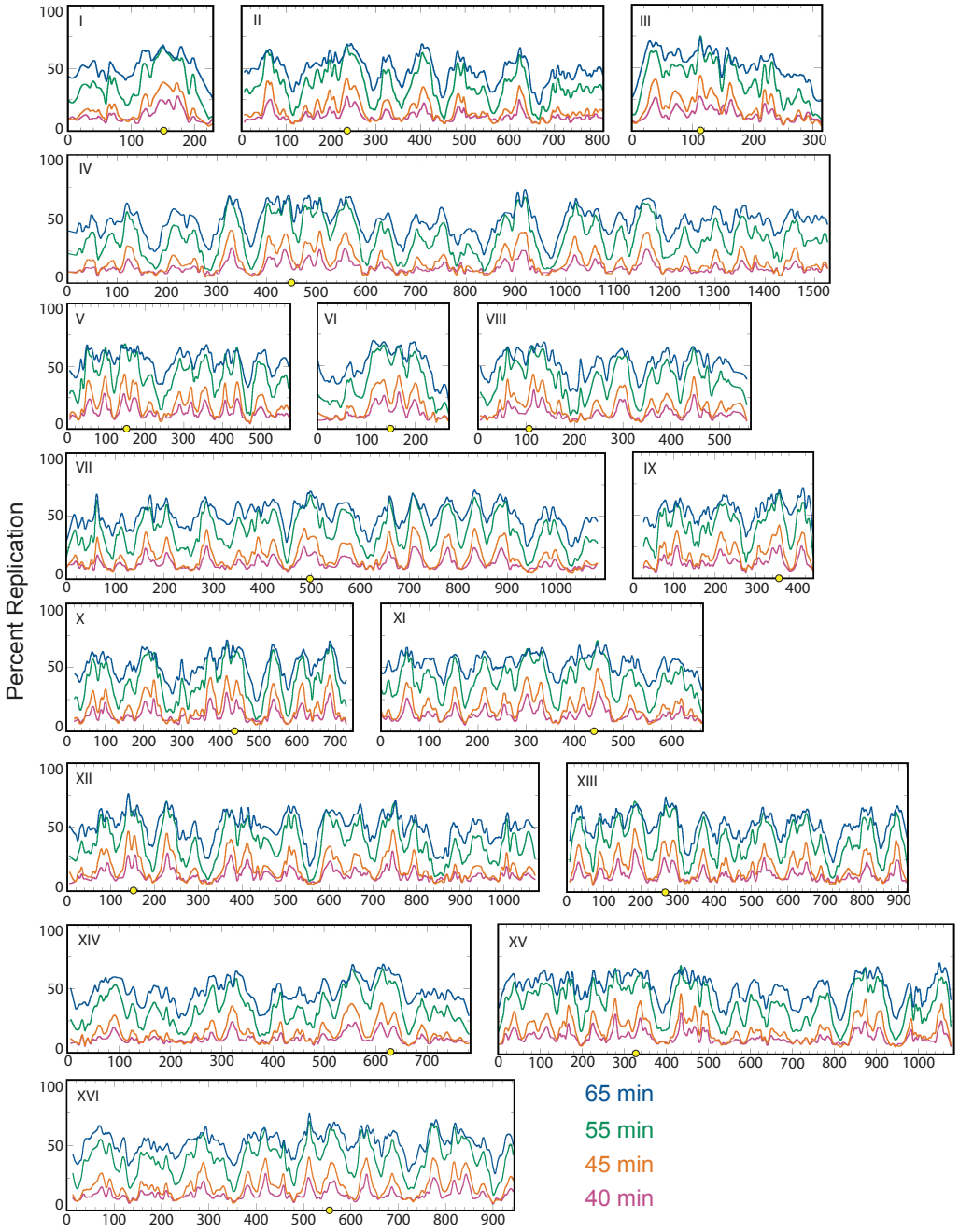


Figure S2.2: Replication kinetic profiles for the WT strain. Microarray analysis was conducted on the 40 (magenta), 45 (orange), 55 (green), and 65 (blue) minute samples. Smoothed data are plotted for each of the 16 *S. cerevisiae* chromosomes.

SUPPLEMANTAL FIGURE 2.3

Rearranged Replication Profiles

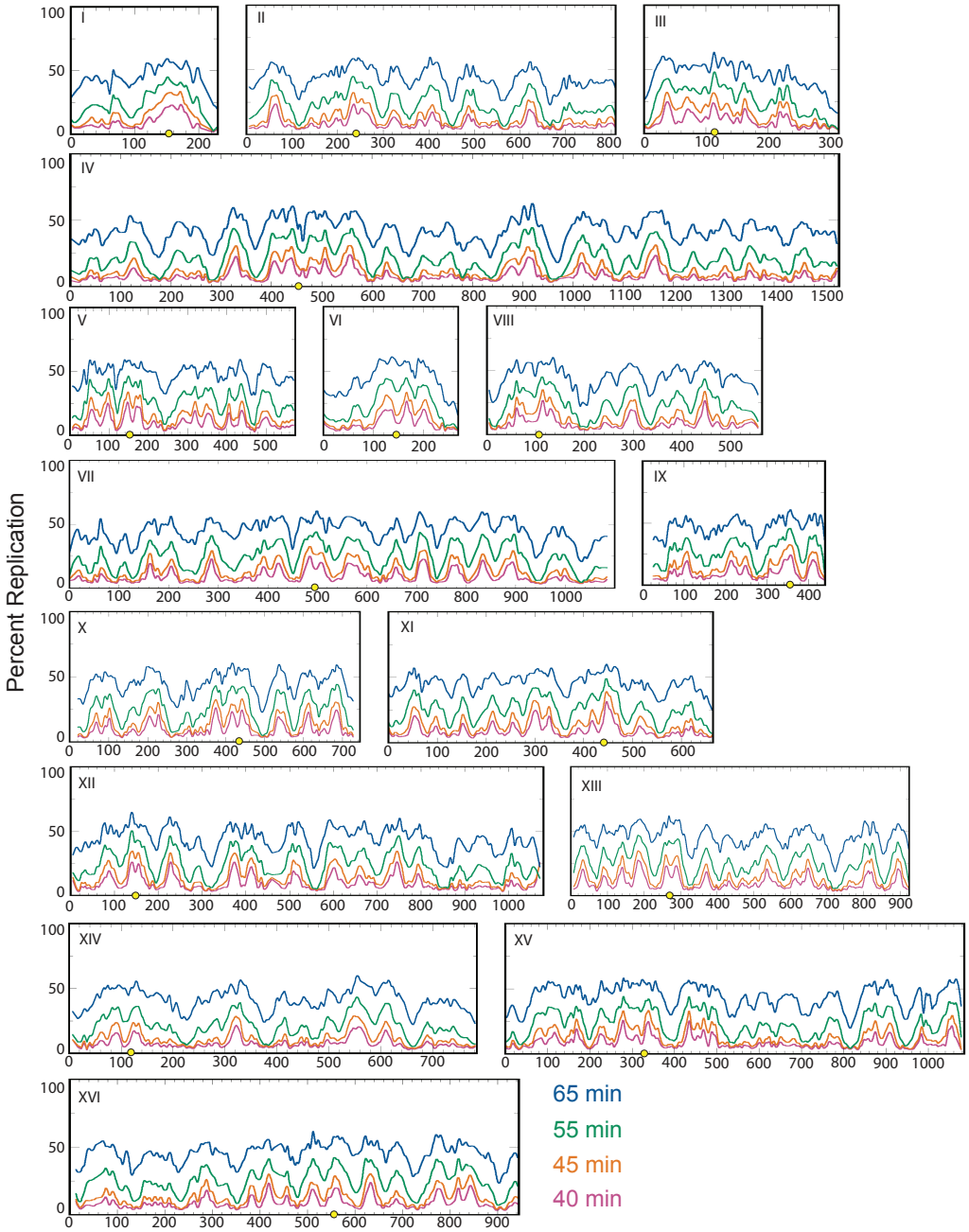


Figure S2.3: Replication kinetic profiles for rearranged strain. Microarray analysis was conducted on the 40 (magenta), 45 (orange), 55 (green), and 65 (blue) minute samples. Smoothed data are plotted for each of the 16 *S. cerevisiae* chromosomes.

SUPPLEMANTAL FIGURE 2.4

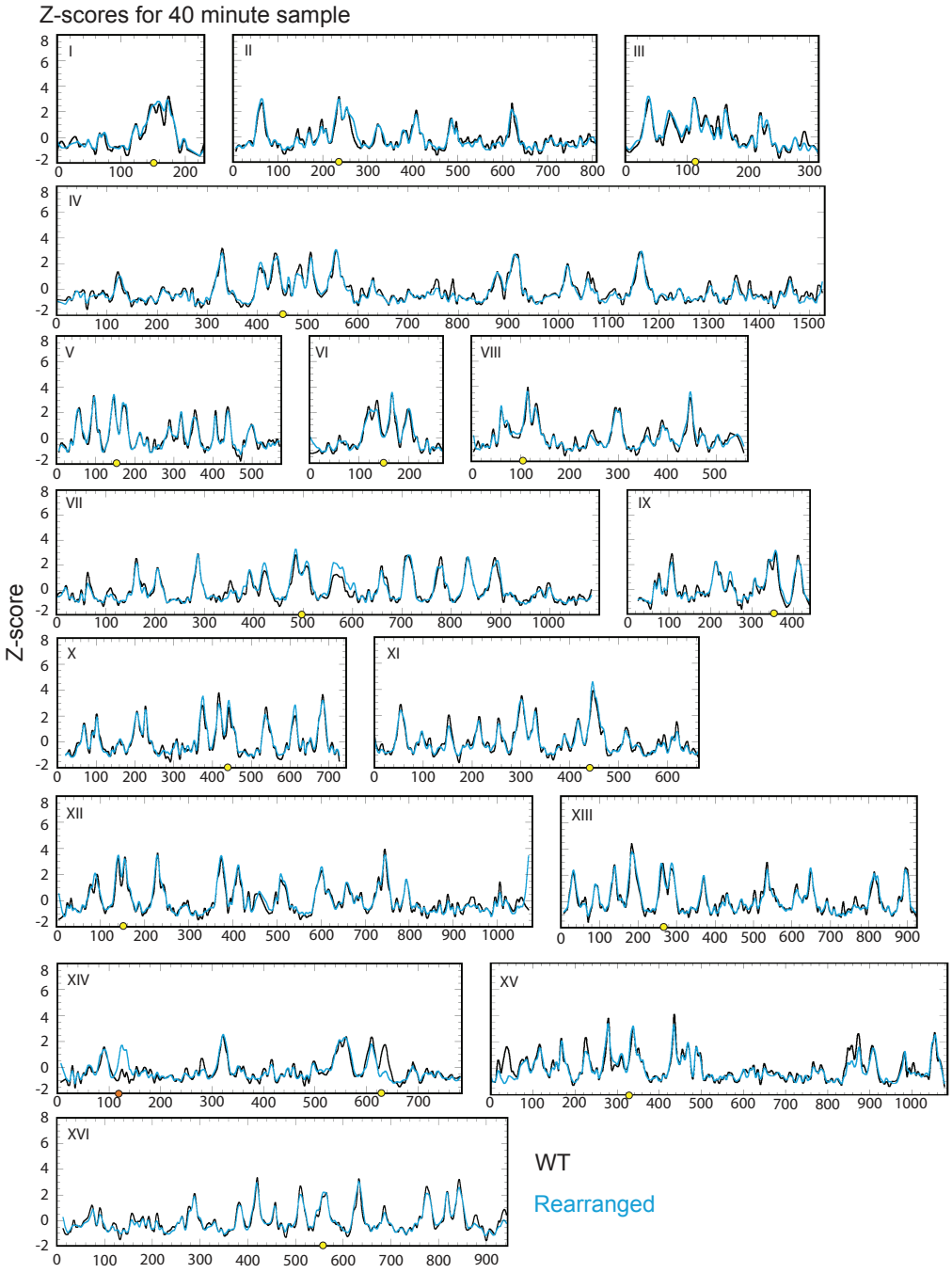


Figure S2.4: Comparison of WT and rearranged strain Z-score data for 40-minute samples. Replication kinetic data for the 40-minute samples in WT and rearranged cells were converted to Z scores and overlaid over the 16 *S. cerevisiae* chromosomes. WT data are plotted in black and rearranged data are plotted in blue. Endogenous and ectopic centromeres are depicted as yellow and orange circles, respectively.

SUPPLEMANTAL FIGURE 2.5

Z-scores for 45 minute sample

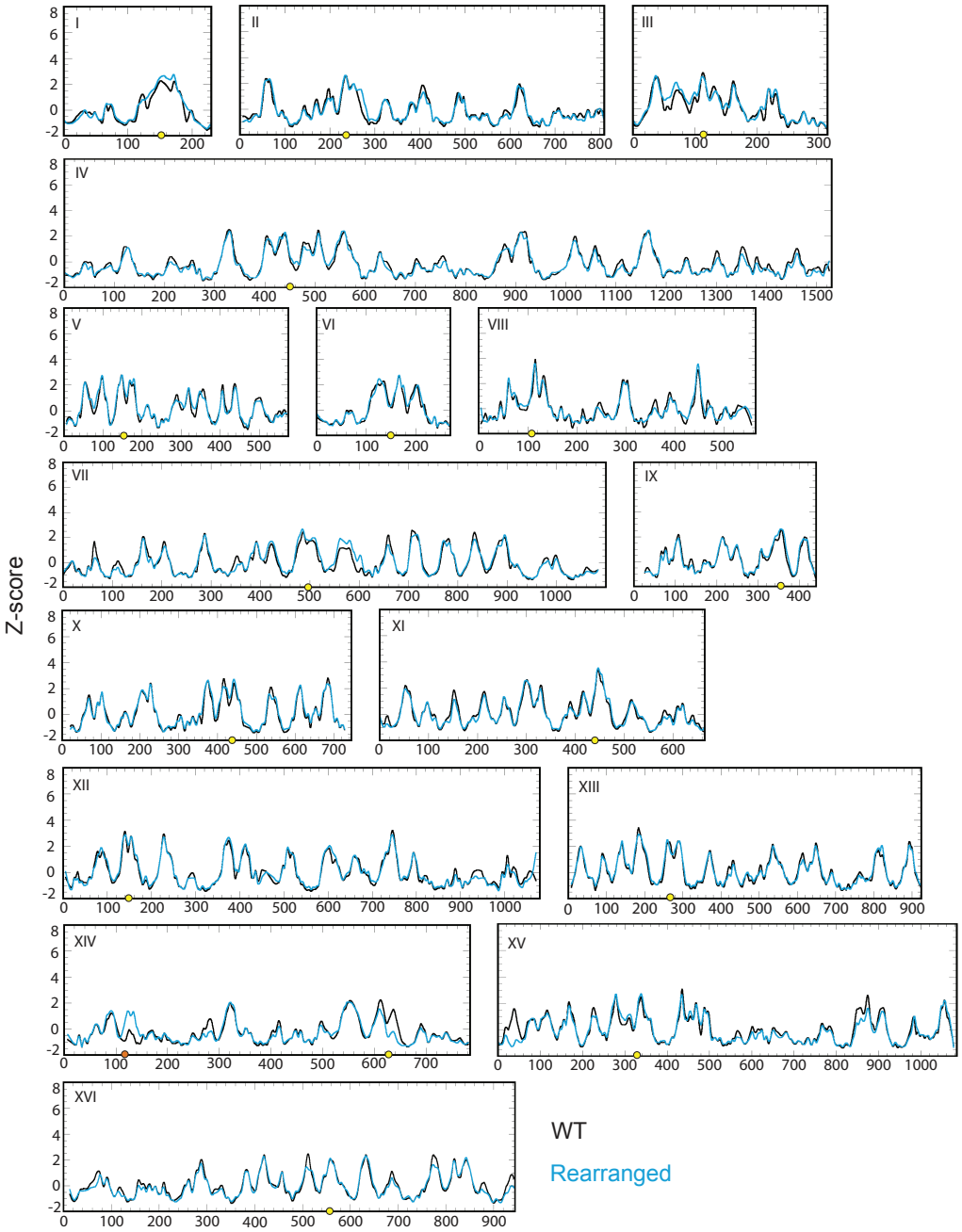


Figure S2.5: Comparison of WT and rearranged stain Z-score data for 45-minute samples. Replication kinetic data for the 45-minute samples in WT and rearranged cells were converted to Z-scores and overlaid over the 16 *S. cerevisiae* chromosomes. WT data are plotted in black and rearranged data are plotted in blue. Endogenous and ectopic centromeres are depicted as yellow and orange circles, respectively.

SUPPLEMANTAL FIGURE 2.6

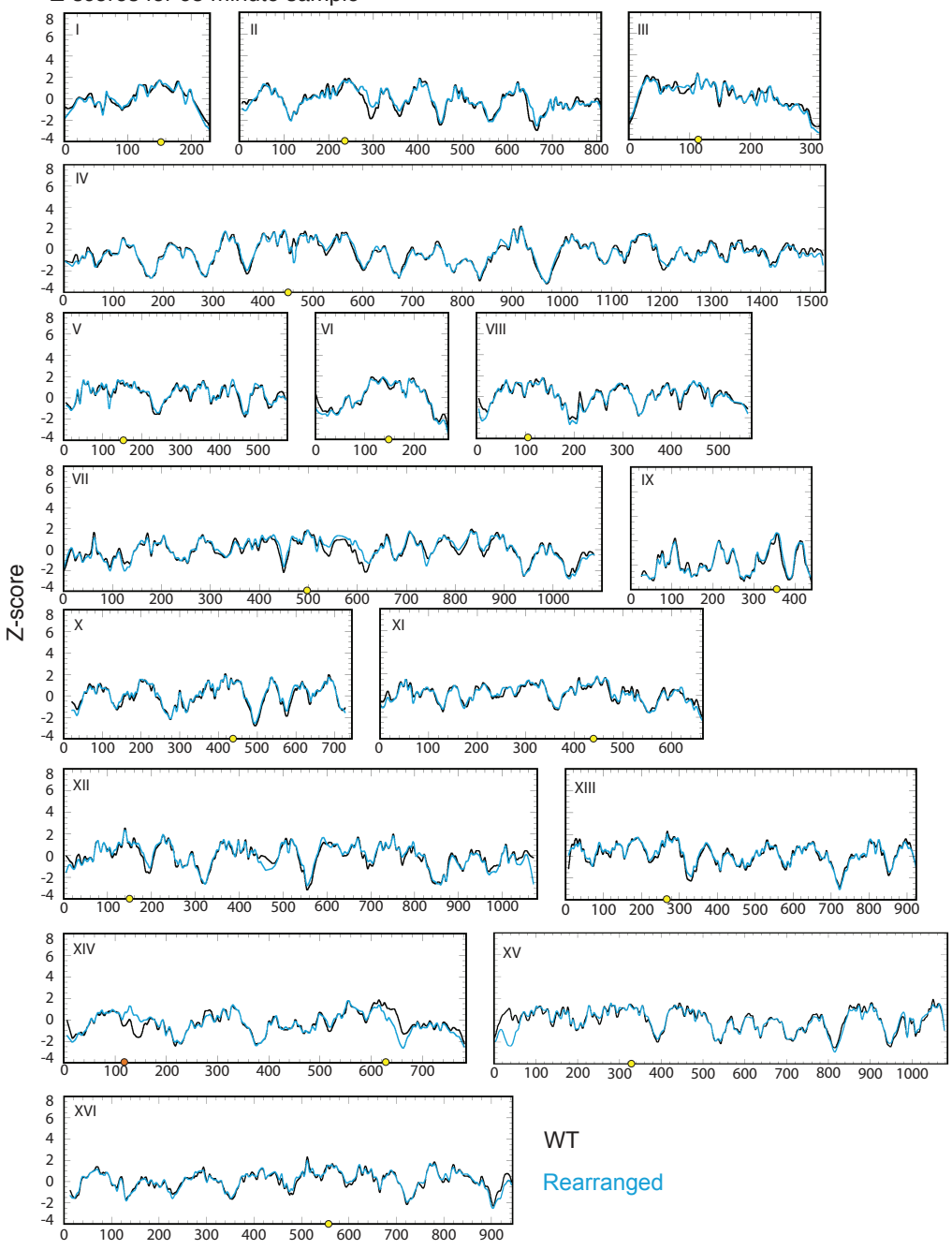


Figure S2.6: Comparison of WT and rearranged strain Z-score data for 65-minute samples. Replication kinetic data for the 65-minute samples in WT and rearranged cells were converted to Z-scores and overlaid over the 16 *S. cerevisiae* chromosomes. WT data are plotted in black and rearranged data are plotted in blue. Endogenous and ectopic centromeres are depicted as yellow and orange circles, respectively.

Supplemental Table S2.1. Yeast strains and plasmids used in this study

Name	Genotype	Source
CH1870	*MATa*, *ura3-52*, *lys2-801am*, *leu2-3,112 met2*::*CEN7.LEU2*, *cen14*::*URA3* in an S288C background	(45)
S288C	*MATa*, *ura3-52*, *trp1*Δ, *leu2-3,112*, *ade2-1*, *met-* (not *met2*)	(Byers lab)
YTP12	*MATa, ura3-52*, *trp1*Δ, *leu2-3,112*, *lys2-801am*, *met2*::*CEN7.LEU2 cen14*::*URA3*	This Study
YTP13	*MATa*, *ura3-52*, *trp*Δ *leu2-3,112*, *lys2-801am*	This Study
YTP15	*MATa*, *trp*Δ, *leu2-3,112*, *lys2-801am*	This Study
YTP16	*MATa*, *ura3-52*, *trp1Δ*, *leu2-3,112*, *lys2-801am*, *met2*::*CEN7.LEU2 cen14*::*URA3*	This Study
YTP19	*MATa*, *trp*Δ *leu2-3,112*, *lys2-801am*, *met2::cen7(cdeIII:XbaI).LEU2*	This Study
Plasmids	Genotype	Source
pUC18-KanMX-ARS228	*Amp, KanMX, ScARS228*	Brewer lab
pTP18	*Amp, ScARS228, met2::CEN7.LEU2*	This Study
pTP19	*Amp, ScARS228, met2::cen7(cdeIII:XbaI).LEU2*	This Study

Supplemental Table S2.2. Primers used in this study

Primer Number	Sequence	Use
71	CCAATAGATGAGCTTCCGCTT	Hybridizes upstream of *MET2*. Used to confirm integration of *CEN7* and *LEU2* at the *MET2* locus. Also, to clone the 5' half of *met2::CEN7.LEU2* to pUC18-*KanMX-ARS228* vector.
72	GTGCGGTCAAAATGTGGAAA	Hybridizes downstream of *MET2*. Used to confirm integration of *CEN7* and *LEU2* at the *MET2* locus. Also to clone the 3' half of *met2::CEN7.LEU2* to pUC18-*KanMX-ARS228* vector.
88	TGGCAAAACGACGATCTTCT	Hybridizes in the 3' end to *LEU2*. Used to confirm integration of *CEN7* and *LEU2* at the *MET2* locus.
126	ATCGCCACTATCTTGTCTGCT	Hybridizes in the 5' end to *LEU2*. Used to confirm integration of *CEN7* and *LEU2* at the *MET2* locus.
133	TTGGCCTCTTCAAGATTATGG	Hybridizes downstream of the EcoRV site in *LEU2*. Used to clone the 5' half of *met2::CEN7.LEU2* to pUC18-*KanMX-ARS228* vector.
134	TGGCGATAGGGTCAACCTTAT	Hybridizes upstream of the EcoRV site in *LEU2*. Used to clone the 3' half of *met2::CEN7.LEU2* to pUC18-*KanMX-ARS228* vector.
147	ATTTTTGTTTTTGCCTTCTAGAAAGA AAATAGTTTTTTGTATTACGG	Hybridizes to *CEN7*. Used to mutate CDEIII in pTP18 via site directed mutagenesis.
148	CCGTAATACAAAAAACTATTTTCTTT CTAGAAGGCAAAAACAAAAAT	Hybridizes to *CEN7*. Used to mutate CDEIII in pTP18 via site directed mutagenesis.

Appendix B: Supplement for Chapter 3

SUPPLEMANTAL FIGURE 3.1

Late *CEN V* Replication Profiles

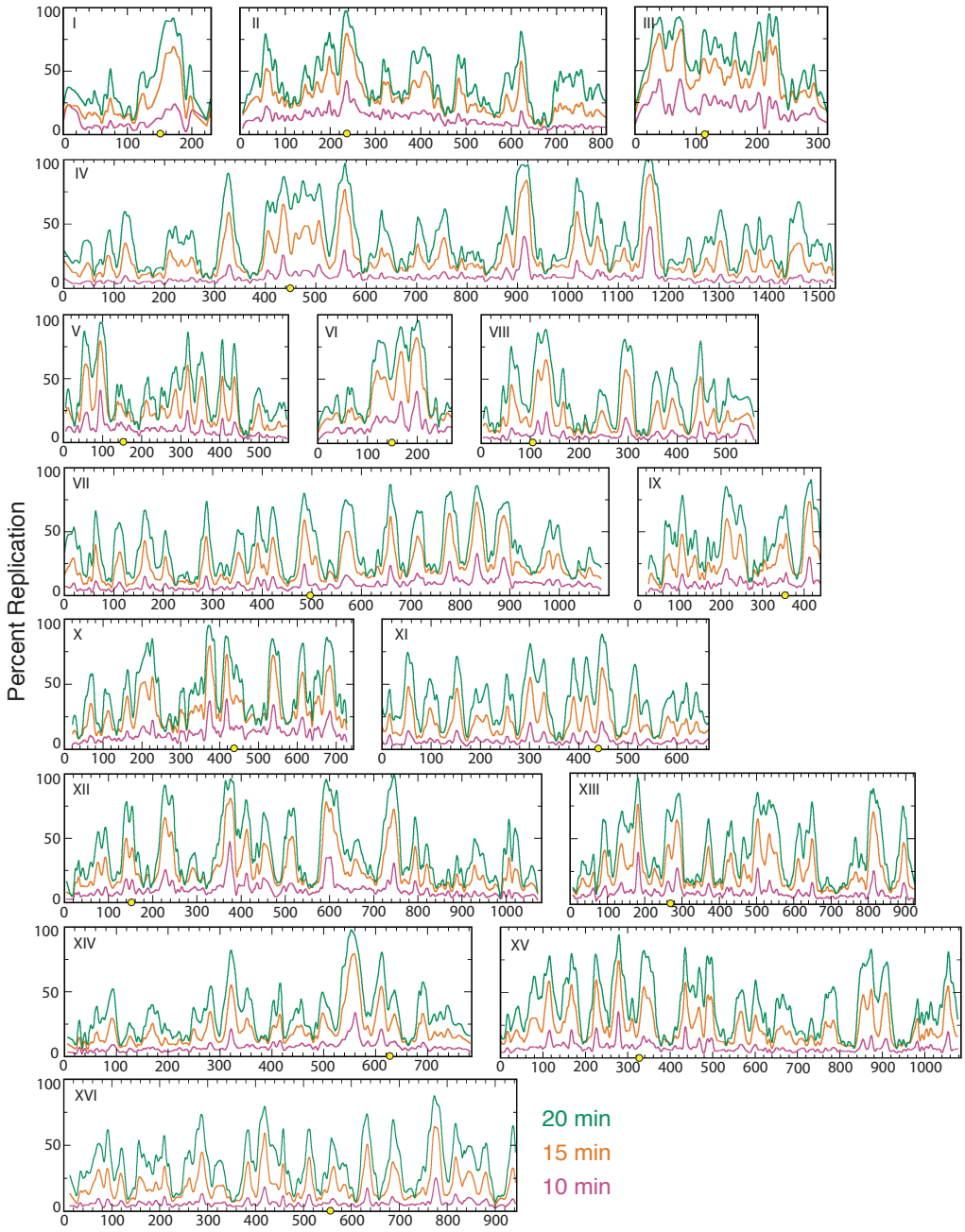


Figure S3.1: Replication kinetic profiles for the late centromere V strain. Microarray analysis was conducted on the 10 (magenta), 15 (orange), and 20 (green) minute samples. Smoothed data are plotted for each of the 16 *S. cerevisiae* chromosomes.

SUPPLEMANTAL FIGURE 3.2

WT Replication Profiles

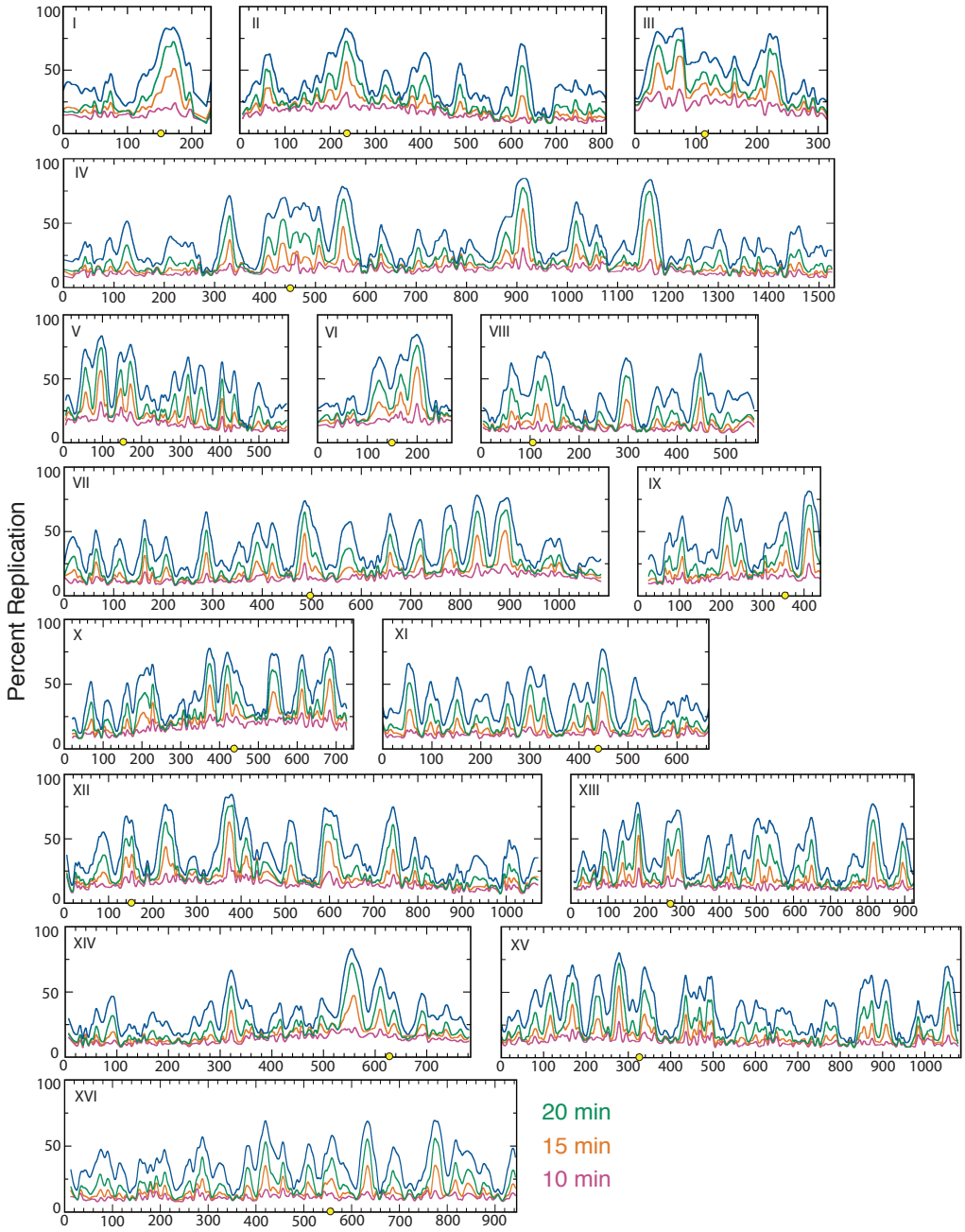


Figure S3.2: Replication kinetic profiles for the WT strain. Microarray analysis was conducted on the 10 (magenta), 15 (orange), and 20 (green) minute samples. Smoothed data are plotted for each of the 16 *S. cerevisiae* chromosomes.

FIGURE S3.3

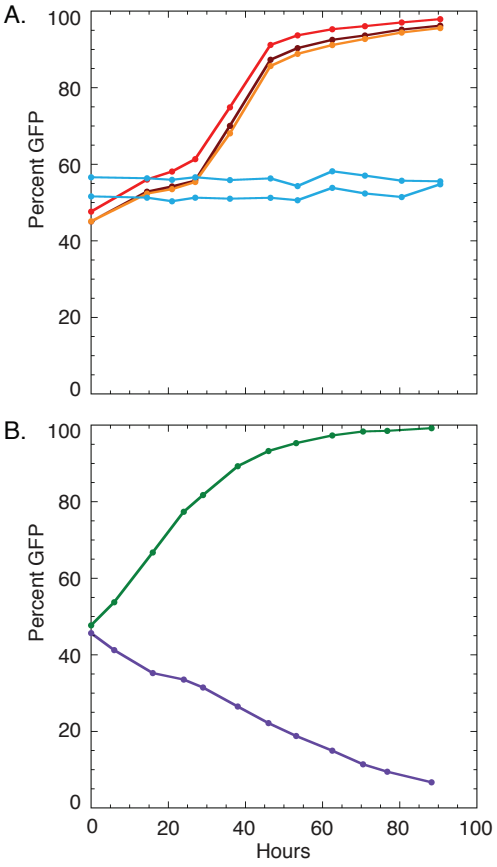

A.
Percent GFP
100
80
60
40
20
0
B.
Percent GFP
100
80
60
40
20
0
0
20
40
60
80
100
Hours

Figure S3.3: Competitive growth analysis of haploids that contain early or late replicating centromere V. (A) The strains used to generate this figure are as follows: The GFP+ and GFP- WT strains that were used in figure 3.4A (GFP+_A and GFP-_A, respectively), along with independent isolates of these two strains (GFP+_B and GFP-_B, respectively), and the two late centromere V strains that were used in figure 3.4A (GFP-_lateCEN).

Plot key:
Maroon = GFP+_A vs. GFP-_B.
Red = GFP+_B vs. GFP-_A.
Orange = GFP+_B vs. GFP-_B.
Light blue = GFP+_B vs. GFP-_lateCEN

(B) The strains used for this figure are all WT for centromere replication time. GFP+ *MAD2* cells were competed against GFP- *mad2Δ* cells (green). GFP+ *mad2Δ* cells were competed against GFP- *MAD2* cells (purple).

Supplemental Table S3.1. Yeast strains and plasmids used in this study. All strains are A364A background unless otherwise noted.

Name	Genotype	Source
YHLTP3	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 ars510::KanMX ars511:NAT*	This Study
YHLTP4	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 ars510::KanMX ars511:NAT*	This Study
YTP40	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 ars510::KanMX/ARS510 ars511:NAT/ARS512 URA3/ura3-52 CAN1/can1-100*	This Study
YTP41	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 KanMX/ARS510 NAT/ARS511 URA3/ura3-52 CAN1/can1-100*	This Study
YTP42	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 ars510::KanMX/ARS510 ars511:NAT/ARS512 URA3/ura3-52 CAN1/can1-100*	This Study
YSL1_c	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 KanMX CloNAT*	This Study
SL14-3α	*MATα cdc7-1 bar1 trp1-289 leu2-3,112 ura3-52 his6 can1-100*	This Study
YSL2	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 KanMX ARS510/ARS510 NAT ARS511/ARS511 URA3/ura3-52 CAN1/can1-100 mad2Δ:HYG/mad2Δ:HYG*	This Study
YSL3	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 ars510::KanMX/ARS510 ars511:NAT/ARS511 URA3/ura3-52 CAN1/can1-100 mad2Δ:HYG/mad2Δ:HYG*	This Study
YSL4	*MATa/MATα cdc7-1/cdc7-1 bar1/bar1 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his6/his6 ars510::KanMX/ARS510 ars511:NAT/ARS511 URA3/ura3-52 CAN1/can1-100 mad2Δ:HYG / mad2Δ:HYG*	This Study
SL14-3α_mad2Δ	*MATα cdc7-1 bar1 trp1-289 leu2-3,112 ura3-52 his6 can1-100 mad2Δ:HYG*	This Study
YSL1_c_mad2Δ	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 KanMX CloNAT mad2Δ:HYG*	This Study
YHLTP3_mad2Δ	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 ars510::KanMX ars511:NAT mad2Δ:HYG*	This Study
YHLTP4_mad2Δ	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 ars510::KanMX ars511:NAT mad2Δ:HYG*	This Study
YSL1_c_GFP	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 KanMX CloNAT ho::GFP_KanMX*	This Study
YSL1_c_mad2Δ_GFP	*MATa cdc7-1 bar1 trp1-289 leu2-3,112 his6 KanMX CloNAT mad2Δ:HYG ho::GFP_KanMX*	This Study
T7107	*MATa, RAD5, BUD4, leu2, ura3, trp1, ade2, his,* in W303 background	(191)

10650_a1	*Mata mutars919 mutars920 RAD5 BUD4 leu2 ura3 trp1 ade2 his3*, in W303 background	Derived from (191)
10650_a2	*Mata mutars919 mutars920 RAD5 BUD4 leu2 ura3 trp1 ade2 his3*, in W303 background	Derived from (191)

Supplemental Table S3.2. Primers used to confirm ARS knockouts

Primer Number	Sequence	Use
217	CCGTGACTTATCCTCACGAAA	Confirmed replacement of *ARS510* with *KanMX*
219	GAGCGATTGCTTCCTTAAAAT	Confirmed replacement of *ARS510* with *KanMX*
220	CTGTCCCATTCGATTTCAAA	Confirmed replacement of *ARS510* with *KanMX*
221	ATTTGCACATAAAGGGTGCC	Confirmed replacement of *ARS510* with *KanMX*
229	AGTACGAGACGACCACGAAGC	Confirmed replacement of *ARS511* with *ClonNAT*^{R}
230	TGACCGTCGAGGACATCGAG	Confirmed replacement of *ARS511* with *ClonNAT*^{R}
237	TTAACGACAGCCTTTGGAGC	Confirmed replacement of *ARS511* with *ClonNAT*^{R}
238	TAAACGCGCGAAGAAAGATG	Confirmed replacement of *ARS511* with *ClonNAT*^{R}

Appendix C: Supplement for Chapter 4

SUPPLEMENTAL FIGURE 4.1

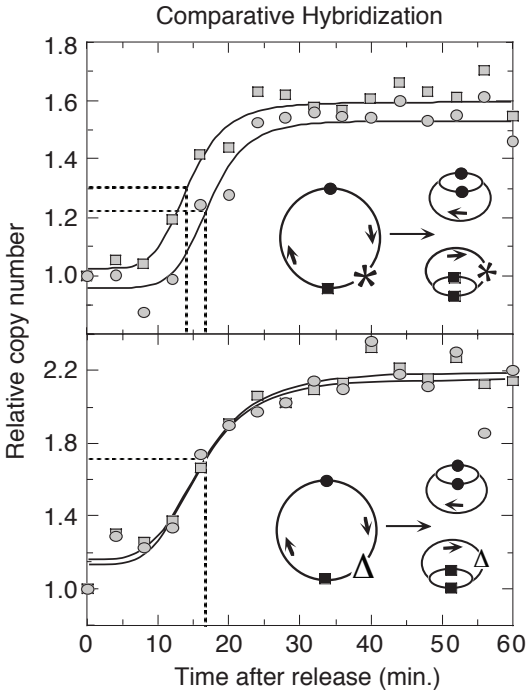


Figure S4.1: *ARS1*^{S} activation time analyzed by plasmid pop-out assay. The two plasmid copies of *ARS1* were separated from each other in vivo. Upper panel: separation from plasmid pN&Sdir. Lower panel: separation from plasmid pN&SΔ9dir derived from pN&Sdir by deleting the *Nsi*I-*Nco*I fragment. Replication kinetics determined by the comparative hybridization method (see Materials and Methods) are plotted for each separated, single-ARS plasmid: plasmid pS' or pS'Δ (squares) and plasmid pN' (circles). Recombinase target sites are indicated by arrows on the cartoons and (Δ) indicates the *Nsi*I-*Sma*I deletion of the bias element (*). The T_{rep} value for each plasmid is shown as a dotted line intersecting the X-axis.

SUPPLEMENTAL FIGURE 4.2

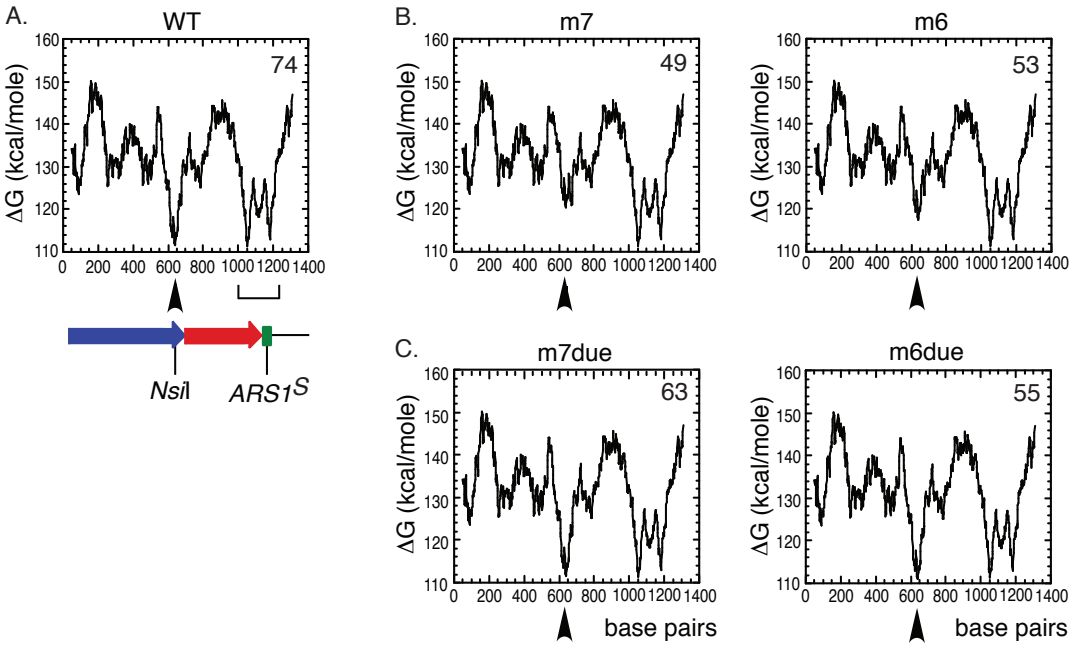


Figure S4.2: Examining the importance of a DUE at the bias determinant. Helical stability plots depicting the energy of unwinding (ΔG) of duplex DNA for the region containing the bias determinant (black arrow head). The percent of *ARS1*^{S} usage is depicted in the upper right corner. The cartoon depicts the sequence that was analyzed: the 3' end of *URA3* (blue arrow), the 3' end of *TRP1* (red arrow) and *ARS1*^{S} (green box). **(A)** Helical stability for the WT bias determinant and its surrounding region. The bracket indicates the *ARS1* DUE. **(B)** Helical stability for bias determinant mutants m6 and m7. **(C)** Helical stability for bias determinant mutants (see Materials and Methods) that maintain the presence of DUE but lack *ARS1*^{S} bias.

VITA

Thomas J. Pohl was born in Albuquerque, New Mexico on November 1, 1982 to Debra Pohl (nee Piro) and Max (Macario) Pohl. In 2001, he graduated as Valedictorian from Jemez Valley High School in Jemez Pueblo, New Mexico, whereafter he attended the University of New Mexico in Albuquerque, New Mexico. In 2004, Thomas identified a new species of *Andrena callandrena* in the laboratory of Dr. Leah L. Larkin. In 2005, he entered the Initiatives for Maximizing Student Development program under the guidance of Dr. Maggie Werner-Washburne where he conducted research studying the mechanisms of double strand break repair in the laboratory of Dr. Jac A. Nickoloff. In 2007, Thomas obtained a Bachelor of Science in Biology and a minor in Psychology from the University of New Mexico, becoming the first person in his family to graduate college. In 2007, Thomas began his graduate work at the University of Washington in Seattle, Washington. During this time he studied the mechanisms that regulate DNA replication in the laboratory of Dr. Bonita J. Brewer and Dr. M. K. Raghuraman. In 2013, he graduated from the University of Washington with a Doctor of Philosophy in Molecular and Cellular Biology.