The Regulated Loading and Distribution of the Mcm2-7 Helicase During G1

by

Sara Kathleen Powell

University Program in Genetics and Genomics
Duke University

Date:_______________________
Approved:

____
David MacAlpine, Supervisor

____
Kenneth Kreuzer

____
Beth Sullivan

____
Bernard Mathey-Prevot

____
Sue Jinks-Robertson

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the University Program in
Genetics and Genomics in the Graduate School
of Duke University

2014

ABSTRACT

The Regulated Loading and Distribution of the Mcm2-7 Helicase During G1

by

Sara Kathleen Powell

University Program in Genetics and Genomics
Duke University

Date:_______________________
Approved:

____
David MacAlpine, Supervisor

____
Kenneth Kreuzer

____
Beth Sullivan

____
Bernard Mathey-Prevot

____
Sue Jinks-Robertson

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the University Program in Genetics and Genomics in the
Graduate School
of Duke University

2014

Copyright by
Sara Kathleen Powell
2014

Abstract

DNA replication is the process of synthesizing an exact copy of a genome during S-phase. DNA replication must only occur once and only once in the cell cycle. Therefore, the DNA replication program is a highly regulated process that is established prior to S-phase. In G1, the protein complexes that define an origin of replication are assembled stepwise onto chromatin. Potential origins of replication are first bound by ORC, origin recognition complex. ORC then recruits other factors that load the Mcm2-7 helicase onto the chromatin to assemble a pre-RC, pre-replication complex. Paradoxically, there is a vast excess of Mcm2-7 relative to ORC assembled onto chromatin in G1. These excess Mcm2-7 are broadly distributed on chromatin, exhibit little co-localization with ORC or replication foci, and can function as dormant origins. I used biochemical and genomic approaches to dissect the mechanisms regulating the assembly and distribution of the Mcm2-7 complex in *Drosophila* tissue culture cells. I found that Mcm2-7 loading occurs in two distinct phases during G1. In the first phase, limiting amounts of Mcm2-7 are loaded at ORC binding sites in a cyclin E/Cdk2 independent manner. Subsequently, there is a cyclin E/Cdk2 kinase activity dependent phase of Mcm2-7 loading that results in a 15-fold increase in chromatin associated Mcm2-7 and a dramatic genome-wide reorganization of the distribution of Mcm2-7 that is shaped by active transcription. Thus, increasing cyclin E/Cdk2 kinase activity over the course of G1 is not only critical for Mcm2-7 loading, but also the distribution of the Mcm2-7 helicase prior to S-phase entry.

The assembly of the pre-RC is not only required for DNA replication, but it has been implicated in being required for cohesin loading. The cohesin complex imparts cohesion between sister chromatids as they are replicated and remains in place until the sister chromatids are separated in mitosis. I assessed if pre-RC assembly is required for cohesin loading using genomic and biochemical approaches in *Drosophila* tissue culture cells. I found that pre-RC components co-localize with cohesin subunits throughout the *Drosophila* genome. I was unable to detect any cohesin loading onto chromatin mediated by pre-RC assembly or components *in vivo*. However, this result does not mean that they are not coordinated.

Any errors during DNA replication can cause genomic instability through rereplication, fragile sites, or stalled forks. In addition, other processes like sister chromatid cohesion that are coordinated with DNA replication can also introduce genomic instability. Aneuploidy is a potential consequence of sister chromatid cohesion defects resulting in unequal multiples of a genome within a cell. Aneuploidy can be detrimental to a cell or organism due to copy number variation (CNV) causing differences in expression of genes. However, cells are able to compensate for CNV between the sexes due to the differences in the number of sex chromosomes. I used genomic approaches to characterize three aneuploid *Drosophila* cell lines for the modENCODE project. I further characterized the S2 *Drosophila* cell line using immunofluorescence microcopy approaches to identify the chromosomal rearrangements that were mapped by *de novo* assembly of the genome. Both approaches showed that the S2 cell line has highly rearranged chromosomes. The S2 cell line was also analyzed to address if cells can compensate for CNV on autosomes using genomic approaches. In collaboration with the Oliver group (NIH) we found that S2 cells are able to compensate for CNV of autosomal genes by buffering gene expression.

In summary, my research explored mechanisms that a cell can employ to maintain genomic stability: assembly of dormant origins, chromosome segregation, and CNV compensation.

Dedication

This work is dedicated to my beloved handsome chocolate devil who has made the bad times worse and the good times better. This work is also dedicated to those who left us too soon, P.S. and my beloved furry Spanish harlot, Carmen.

Contents

Abstract iv

List of Figures xi

Abbreviations Used xiii

Acknowledgements xiii

1. Introduction 1

  1.1 DNA Replication 1

  1.2 Origins of DNA Replication 1

  1.3 Initiators of DNA Replication 2

  1.4 Origins of Replication are Organized within the Nucleus 6

  1.5 Cyclin E/ Cdk2 Kinase Activity Drives the G1-S Transition 7

  1.6 DNA Replication Initiation 9

    1.6.1 Assembly of the Pre-IC and Helicase Activation 9

    1.6.2 DNA Polymerase Loading 10

    1.6.3 Timing of DNA Replication Initiation 12

  1.7 The 'MCM Paradox' 14

  1.8 Genomic Instability 16

2. Dynamic Loading and Distribution of Mcm2-7 Throughout the Cell Cycle 20

  2.1 Introduction 20

  2.2 Mcm2-7 Loading Occurs in Two Distinct Phases During G1 22

  2.3 Cyclin E Partially Rescues Mcm2-7 Loading Independently of Kinase Activity In Vivo 27

  2.4 Maximal Loading of Mcm2-7 Coincides with Entry into S-phase 29

  2.5 The Cyclin E/Cdk2 Kinase Dependent Phase of Mcm2-7 Loading Requires the Canonical Pre-RC Assembly Pathway 31

2.6 Genome-wide Distribution of Mcm2-7 is Determined by Cyclin E/Cdk2 Activity and Local Transcription 33

2.7 Discussion 42

2.7.1 Precise Regulation of Cyclin E/Cdk2 Kinase Activity is Critical for Pre-RC Assembly and Genomic Stability 44

2.7.2 Genome-wide Redistribution of Mcm2-7 46

2.7.3 Defining an Origin of Replication 48

2.8 Future Directions: Progression into S-phase with Minimal Chromatin Bound Mcm2-7 Levels. 49

2.9 Materials and Methods 53

3. Testing the Role of Pre-RC Assembly in Cohesin Loading 60

3.1 Cohesin Co-localizes with Pre-RC Components Throughout the Drosophila Genome 60

3.2 Cohesin can Load Independently of the Pre-RC in Vivo 61

3.3 Cohesin can Load Independently of ORC in Vivo 61

3.4 Cohesin can Load in the Absence of Nipped-B in Vivo 62

3.5 No in Vitro Interaction was Detected between Pre-RC Components and Cohesin 67

3.6 Summary 67

3.7 Materials and Methods 69

4. Characterization of Aneuploidy in *Drosophila* Cell Lines 72

4.1 Drosophila S2 Cells have an Aberrant Karyotype 72

4.2 Each Cell Line has a Unique Gross Copy Number Variation Along All of the Chromosomes 73

4.3 Determination of CNV in Three Drosophila Cell Lines 73

4.4 Copy Number does not Correlate 1-to-1 with Expression 77

4.5 Drosophila S2 Cells have Rearranged Chromosomes 78

4.6 Summary 83

4.7 Materials and Methods 83

5. Conclusions 88

Bibliography 93

Biography 113

List of Figures

Figure 1: The Initiators Required for Assembly of a Pre-RC in G1. 3

Figure 2: The Activated Mcm2-7 Helicase Travels Away from an Origin as Part of CMG and Replisome in S-Phase. 12

Figure 3: The 'MCM Paradox'.. 15

Figure 4: Mcm2-7 Loading Occurs in Two Distinct Phases during G1. 23

Figure 5: Mcm2-7 Chromatin Bound Levels are Similar in Dup/Cdt1 RNAi and G2 Arrested Cells. 25

Figure 6: FLAG-cyclin E Overexpression in the Absence of Cdk2 Kinase Activity can Stimulate Mcm2-7 Loading. 28

Figure 7: Mcm2-7 Loading in an Unperturbed G1-S Transition. 29

Figure 8: The Cyclin E/Cdk2 Kinase Dependent Phase of Mcm2-7 Loading Requires the Canonical Pre-RC Assembly Pathway. 31

Figure 9: Mcm2-7 Loading Occurs at ORC Binding Sites in Cyclin E RNAi Cells. 34

Figure 10: The Mcm2-7 Chromatin Distribution in G1 is Dependent on Cyclin E/Cdk2 Activity. 35

Figure 11: The Mcm2-7 Chromatin Distribution in G1 is Dependent on the Transcription Machinery. 36

Figure 12: Bimodal Enrichment of Mcm2-7 at Active and Inactive Transcripts is not Dependent on Early Origin Activation. 38

Figure 13: Mcm2-7 Enrichment in Non-transcribed Genes is Indistinguishable from Intergenic Levels. 39

Figure 14: Mcm2-7 are Displaced from Chromatin by DNA Replication. 40

Figure 15: Model of Pre-RC Assembly and Redistribution of Mcm2-7 in G1. 42

Figure 16: Cyclin E Independent Amount of Chromatin Bound Mcm2-7 are Sufficient for Entry into S-phase. 49

Figure 17: Dacapo Overexpression Level of Chromatin Bound Mcm2-7 are Insufficient to Support Entry into S-phase. 50

Figure 18: Chromatin Fractionation Optimization in *Drosophila* for Mcm2-7 and ORC. . 55

Figure 19: Cohesin and Pre-RC Components Colocalize Along the *Drosophila* Genome. 63

Figure 20: Cohesin can Load Independently of Pre-RC Assembly *in Vivo*. 64

Figure 21: ORC Depletion Does Not Inhibit Cohesin Loading onto Chromatin *in Vivo*.. 65

Figure 22: Cohesin Components are Chromatin Bound in the Absence of Nipped-B. 66

Figure 23: No Stable Interaction Between Mcm2-7 and Cohesin Subunits was Detected.68

Figure 24: Karyotype of a *Drosophila* S2 Cell. 74

Figure 25: Each Cell Line has a Unique Pattern of CNV. 75

Figure 26: CGH and NGS CNV Analysis Comparison. 76

Figure 27: Copy Number and Transcript Levels Comparison between Two *Drosophila* Cell Lines is not a 1-to-1 Relationship. 79

Figure 28: S2 Cells have Rearranged Chromosomes. 80

Figure 29: S2 Cells have Multiple Chromosome Break Points. 81

Figure 30: *Drosophila* S2 Cells have Complex Chromosome Rearrangements. 82

Abbreviations Used

ACS- ARS consensus sequence

ARS- autonomously replicating sequence

BAC- Bacterial artificial chromosome

CDK- cyclin dependent kinase

CGH- comparative genomic hybridization

ChIP- chromatin immunoprecipitation

CHORI- Children's Hospital Oakland Research Institute

CNV- copy number variation

DDK- Dbf4 dependent kinase

DMSO- Dimethyl sulfoxide

FISH- fluorescent *in situ* hybridization

GINS- Go, Ichi, Nii, and San

HU- hydroxyurea

IP- immunoprecipitation

NES- nuclear export sequence

NGS- next generation sequencing

NLS- nuclear localization sequence

Mcm2-7- minichromosome maintenance 2-7

OGREs- origin G-rich repeated elements

ORC- origin recognition complex

Pre-RC- pre-replication complex

SCC- sister chromatid cohesion

WCE- whole cell extract

Acknowledgements

The bioinformatic analyses throughout this work were made possible by Matt Eaton Ph.D. and Mr. Joey Prinz.

1. Introduction

1.1 DNA Replication

A pioneering theory, in 1963, on the mechanism of replication in prokaryotes postulated a simple replicator-initiator model (Jacob et al., 1963). The replicator or 'operator of replication' is a sequence element that the 'initiator' acts upon to promote DNA replication. The 50 years of research since have validated that this simple model is accurate for the replication of the 1 million base pairs of DNA in prokaryotes to the ~3 billion base pairs of DNA found in the human genome. Specific replicators and initiators have been identified in prokaryotes and eukaryotes [reviewed in (O'Donnell et al., 2013)]. Despite the functional conservation of replicators and initiators between prokaryotes and eukaryotes, the mechanisms that specify replicators in the chromosomal context and also govern the activity and localization of initiators are poorly understood in higher eukaryotes. These additional layers of regulation ensure that DNA replication occurs once and only once in a cell cycle as inappropriate DNA replication may result in genomic instability.

1.2 Origins of DNA Replication

Eukaryotic cis-acting sequence elements (replicators) that define origins of DNA replication are best understood in *S. cerevisiae*. Autonomously replicating sequences (ARSs) or replicators were identified by their ability to support the episomal maintenance of a plasmid (Stinchcomb et al., 1979). ARSs were further characterized by comparing several ARSs and performing deletion analysis to identify the ACS (ARS consensus sequence) (Broach et al., 1983; Celniker et al., 1984).

In higher eukaryotes, the identification of a common replicator present throughout the genome had eluded researchers for years. Several different experimental and computational methods failed to define a replicator [reviewed in (Leonard and Mechali, 2013)]. Interestingly, a possible metazoan replicator, the G-quadruplex, was recently described. Studies in mice and *Drosophila* cells identified origin G-rich repeated elements (OGREs) that are correlated with replication origins (Cayrou et al., 2012). OGREs also predict the presence of G-quadruplex structures at nearby origins (Cayrou et al., 2012). A functional analysis of two G-quadruplex motifs suggests that they are necessary, but not sufficient for promoting DNA replication (Valton et al., 2014).

Despite the identification of replicators in eukaryotes their regulation is still not clearly understood. Further experiments demonstrated that only a small fraction of ACS motif matches in *S. cerevisiae* are utilized in a single S-phase. A computational approach identified that only 350 out of 10,000 ACS sites are utilized in a single S-phase (Breier et al., 2004). Therefore, the ACS is necessary, but not sufficient to act as an origin of replication. These results suggest that there is additional regulation for the use and selection of replicators by the initiators.

1.3 Initiators of DNA Replication

In the simple model, a replicator is acted upon by an initiator to promote DNA replication. In eukaryotes, many initiator proteins have been identified and characterized. Further, research has shown that the initiators function in a highly regulated cell cycle dependent manner to promote DNA replication at the replicator.

The ACS is the DNA binding site for ORC (origin recognition complex (Orc1-6)), one of many initiator proteins (Bell and Stillman, 1992). However, in higher eukaryotes a specific ORC binding motif has not been identified. Early studies showed that ORC readily binds to DNA in a sequence independent manner *in vivo* and *in vitro* in *Drosophila* and humans (Vashee et al., 2003; Remus et al., 2004). Whole genome studies have provided more insights into ORC binding sites. Whole genome studies in *Drosophila* cell lines found ORC enriched in euchromatin or 'open' chromatin using ChIP-chip and ChIP-seq (MacAlpine et al., 2010). Recent studies in mice and *Drosophila* cells identified that OGREs are correlated with ORC binding sites (Cayrou et al., 2012). Together, these results suggest that multiple factors impact the location of ORC binding in higher eukaryotes.

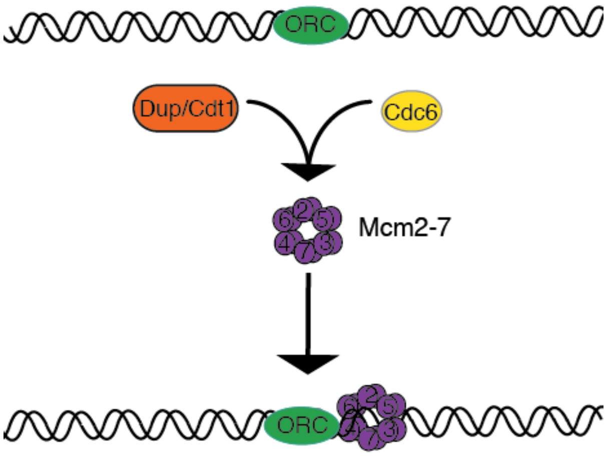

Figure 1: The Initiators Required for Assembly of a Pre-RC in G1.

Cartoon diagram of chromatin bound ORC recruiting Dup1/Cdt1 and Cdc6 to mediate a double hexamer of Mcm2-7 loading onto dsDNA.

ORC binding acts to recruit other initiator proteins to DNA. The next sequence of steps has been studied in *S. cerevisiae* using mutants and further validated using an *in vitro* assay **(Figure 1)**(Tsakraklides and Bell, 2010). ORC binds to DNA in an ATP-dependent manner (Bell and Stillman, 1992). DNA footprint and ChIP assays suggest that ORC is bound to the ACS throughout the cell cycle (Aparicio et al., 1997; Fujita et al., 1998). Subsequently in G1, Cdc6 (Cocker et al., 1996) is recruited and stably binds to chromatin bound ORC in an ATPase dependent manner (Speck et al., 2005). Two Cdt1-Mcm2-7 complexes are recruited and interact with the chromatin bound ORC-Cdc6 complex via two binding sites on Orc6 (Chen et al., 2007). The two recruited molecules of Mcm2-7 are loaded as a double hexamer encircling dsDNA through its central channel (Evrin et al., 2009; Remus et al., 2009; Evrin et al., 2013). Cdc6 ATPase activity loads Mcm2-7 onto dsDNA and also releases Cdt1 from the ORC-Cdc6 complex (Randell et al., 2006). Lastly, ORC ATP hydrolysis enables the loading of additional Mcm2-7 (Bowers et al., 2004). The assembled complex of ORC and a double hexamer of Mcm2-7 constitutes an assembled pre-RC (pre-replication complex). A pre-RC is assembled at each potential origin of replication and is required for DNA replication.

In higher eukaryotes Orc1-5, Cdc6, Mcm2-7, and Cdt1 are conserved; Orc6 and Cdt1 are less conserved [reviewed in (Bell and Kaguni, 2013)]. Studies, in *Drosophila*, *Xenopus*, and non-transformed human cells, indicate that their respective Cdt1 homologues are also required for pre-RC assembly and replication (Maiorano et al., 2000; Whittaker et al., 2000; Rialland et al., 2002). The pre-RC assembly steps outlined above, in *S. cerevisiae*, are similar in higher eukaryotes as well. In *Drosophila*, ORC, Cdc6, and Cdc6 ATPase activity are required for *in vitro* DNA replication using embryonic extracts (Svitin and Chesnokov, 2010). ORC, Cdc6, Cdt1, ATP hydrolysis, and Mcm2-7 all act to assemble pre-RCs and are all required for *in vitro* DNA replication, in *Xenopus* embryonic extracts (Gillespie et al., 2001). Experiments in *Xenopus* indicated that an Mcm2-7 double hexamer is loaded during pre-RC assembly (Gambus et al., 2011).

The assembly of the pre-RC defines a potential origin of replication. Therefore, if pre-RCs are aberrantly assembled outside of G1 re-replication can occur. Organisms have evolved a variety of mechanisms to regulate the assembly of pre-RCs. The mechanisms employed to regulate pre-RC assembly are different between metazoans and *S. cerevisiae* [reviewed in (Siddiqui et al., 2013)]. However, all the mechanisms have the common goal to limit DNA replication to once and only once during a cell cycle. In *S. cerevisiae*, CDK activity prohibits pre-RC assembly via four different targets: ORC, Cdc6, Mcm2-7, and Cdt1. Phosphorylation of Orc2 and Orc6 by Cdc28 prevents pre-RC assembly outside of G1 (Nguyen et al., 2001). Orc6 phosphorylation specifically blocks its interaction with Cdt1 (Chen and Bell, 2011) and prevents a phosphorylation event by Clb5/Cdc28 that promotes pre-RC assembly (Wilmes et al., 2004; Chen and Bell, 2011). Cdc6 is also phosphorylated in a cell cycle dependent manner by Clb/Cdc28 that targets Cdc6 for degradation at the end of G1 (Elsasser et al., 1999; Drury et al., 2000) Additionally, Cdc6 is stabilized after S-phase, but remains inactive because it is in a complex with Clb2/Cdc28 (Mimura et al., 2004). Cln/Cdc28 and Clb/Cdc28 both promote the exclusion of Mcm2-4, Mcm6, and Mcm7 from the nucleus as S-phase progresses (Labib et al., 1999; Nguyen et al., 2000). Cdc28 was further shown to mediate the export of Mcm2-4 and Mcm7 by direct phosphorylation of the NLS-NES (nuclear localization sequence-nuclear export sequence) module of Mcm3 (Liku et al., 2005). Finally, Cdt1 is excluded from the nucleus before S-phase in a Cdc28 dependent manner (Tanaka and Diffley, 2002).

The role of CDK regulation of pre-RC components is less clear in higher eukaryotes. In *Drosophila*, *in vitro* analysis of ORC phosphorylation indicated that Orc1 and Orc2 are both possible substrates for cyclin E/Cdk2 and Cyclin B/Cdk1 (Remus et al., 2005). Additionally, phosphorylation of Orc1 and Orc2 inhibit their DNA binding ability *in vitro* (Remus et al., 2005). In human tissue culture cells, CDK phosphorylated

Cdt1 is ubiquitinated by SCF-Skp2 E3 ligase that targets Cdt1 for degradation (Liu et al., 2004; Nishitani et al., 2006). However, studies in *Xenopus* and human cells indicate that the majority of Cdt1 is degraded in S-phase mediated via a PCNA coupled Cul4-Ddb1 mechanism (Arias and Walter, 2006; Jin et al., 2006; Senga et al., 2006).

In higher eukaryotes, ORC subunits dissociate from DNA to inhibit pre-RC assembly outside of G1. In the *Xenopus in vitro* system, Orc1 and Orc2 are removed from the chromatin and localize to the cytoplasm during mitosis (Romanowski et al., 1996b). These data suggest that the localization of Orc1 and Orc2 to the cytoplasm serves to inhibit aberrant pre-RC assembly because ORC is required for pre-RC assembly (Romanowski et al., 1996b). In humans, Orc1 is phosphorylated by cyclin A/CDK2, ubiquitinated, and degraded during S-phase (Kreitz et al., 2001; Li and DePamphilis, 2002; Mendez et al., 2002). In *Drosophila*, Orc1 is degraded in mitosis via the anaphase promoting complex (APC) (Araki et al., 2003).

In metazoans, Dup/Cdt1, a pre-RC assembly factor, is regulated by mechanisms different than those employed in *S. cerevisiae*. In metazoans, Dup/Cdt1 is targeted for ubiquitination and degradation during S-phase by geminin binding (Wohlschlegel et al., 2000; Arias and Walter, 2006; Jin et al., 2006; Senga et al., 2006).

1.4 Origins of Replication are Organized within the Nucleus

DNA replication, in eukaryotes, begins from multiple origins of replication along each chromosome [reviewed in (Leonard and Mechali, 2013)]. There are about 300 functional origins in *S. cerevisiae* (Nieduszynski et al., 2006) and about 30,000 predicted in mammals (Cadoret et al., 2008).

Eukaryotic origins of replication are spatially organized along chromosomes and within the nucleus. The organization of replicated tracts of DNA or replicons were first visualized by pulse labeling mammalian cells with radiolabeled nucleotides (Huberman and Riggs, 1968). Multiple replicons of equal lengths were spaced along on a single DNA fiber (Huberman and Riggs, 1968). These data indicate that origins of replication are often activated at the same time as their neighbors.

Replicons are further assembled into replication factories or foci. Replication foci were first observed by tracking actively replicating regions with BrdU, a nucleotide analogue, in mammalian cells. The BrdU labeled regions or replication foci increased in size and became ring-like over time (Nakamura et al., 1986). When similar experiments were performed with higher resolution microscopy each replication foci was estimated to contain from one to ten replicons in mammalian cells (Jackson and Pombo, 1998; Ma et al., 1998). Replication foci were demonstrated to contain the same replicons over several cell cycles (Jackson and Pombo, 1998; Ma et al., 1998). Together, origins of replication, replicons, and foci are deliberately organized and structured within the nucleus.

1.5 Cyclin E/ Cdk2 Kinase Activity Drives the G1-S Transition

The major drivers of the cell cycle are cyclin-dependent kinases (CDKs) and their regulatory proteins, the cyclins (Mahbubani et al., 1997). The expression, activity, and regulation of CDK and cyclin pairs are at the heart of controlling the cell cycle (Novak et al., 1999; Alberts et al., 2002; Murray, 2004). The target specificity of a cyclin/CDK pair is mediated by the cyclin.

Cyclin E/Cdk2 kinase activity is the key driver of eukaryotic cells through the G1-S transition (Dulic et al., 1992; Ohtsubo and Roberts, 1993; Knoblich et al., 1994). Cyclin E/Cdk2 levels can also modulate the timing of S-phase. An increase in cyclin E/Cdk2 activity results in a shorter G1-S transition (Ohtsubo and Roberts, 1993;

Wimmel et al., 1994; Ohtsubo et al., 1995) and the initiation of more origins in S-phase (Krasinska et al., 2008; Thomson et al., 2010). Cyclin E/Cdk2 promotes the G1-S transition by relieving Rb inhibition on E2F, a transcription factor required for transcription of G1 to S-phase critical genes (Lees et al., 1992).

Cyclin E protein levels are highly regulated in a cell cycle dependent manner. Live cell imaging detected cyclin E/Cdk2 activity gradually increasing through G1 and into S-phase before being down regulated shortly before mitosis using a Cdk2 phosphorylation target as a biosensor (Spencer et al., 2013). The control of cyclin E /Cdk2 kinase activity is regulated by degradation of cyclin E and inhibition of the cyclin E/Cdk2 complex. Cdk2 can phosphorylate cyclin E when they are bound together and promote their dissociation (Clurman et al., 1996). Phosphorylated and free cyclin E is ubiquitinated by Cdc4/Ago/Fbw7, a conserved E3 ligase, and is degraded by the proteasome (Clurman et al., 1996; Koepp et al., 2001; Moberg et al., 2001; Strohmaier et al., 2001; Szuplewski et al., 2009). In higher eukaryotes, cyclin E/Cdk2 kinase activity is also inhibited by the binding of p27/Dacapo (Polyak et al., 1994; Toyoshima and Hunter, 1994; Su et al., 1995). P27/Dacapo overexpression results in a G1 arrest (Lane et al., 1996).

If cyclin E levels are not properly regulated genomic instability can result. Deregulation of cyclin E causes an excessive number of origins to fire and an increase of DNA damage (Jones et al., 2012). Consistent with an increase of the number of origins fired, cyclin E deregulation depletes available nucleotide pools (Bester et al., 2011). Cyclin E deregulation has also been associated with cancer [reviewed in (Donnellan and Chetty, 1999; Schraml et al., 2003)]. Together these results suggest that cyclin E/Cdk2 deregulation is detrimental to a cell and organism.

1.6 DNA Replication Initiation

1.6.1 Assembly of the Pre-IC and Helicase Activation

In S-phase, sites of pre-RC assembly act as licensed origins of replication. The double hexamer of Mcm2-7 is converted into two active helicases that unwind the DNA ahead of the replication fork (Bochman and Schwacha, 2008). Mcm2-7 helicase activation and replication initiation require the recruitment and loading of additional proteins to assemble the pre-IC (pre-initiation complex). The steps to Mcm2-7 helicase activation have been worked out in *S. cerevisiae* using an *in vitro* assay [reviewed in (Tanaka and Araki, 2013)]. First, Mcm2-7 subunits (Mcm2, 4, and 6) are phosphorylated by DDK (Dbf4-dependent kinase) which then promotes Sld3, Sld7, and Cdc45 interaction with Mcm2-7 (Kamimura et al., 2001; Kanemaki and Labib, 2006; Randell et al., 2010; Heller et al., 2011; Tanaka et al., 2011). The second step is dependent on S-CDK (Clb5, Clb6/Cdc28) that phosphorylates Sld2 and Sld3 (Masumoto et al., 2002; Tanaka et al., 2007; Zegerman and Diffley, 2007; Heller et al., 2011) to promote the formation of the complex including Mcm2-7, Sld2, Dpd11, GINS (Go, Ichi, Nii, and San), Sld3, and Cdc45 (Masumoto et al., 2002; Kanemaki and Labib, 2006; Tak et al., 2006; Muramatsu et al., 2010). Together these steps assemble the holo-replicative helicase or CMG complex (Cdc45, Mcm2-7, GINS) (Gambus et al., 2006). During these steps, Mcm2-7 is converted from a dsDNA to a ssDNA encircling conformation perhaps via ATPase activity and conformational changes of the ring like protein structure (Bochman and Schwacha, 2010). *In vitro Xenopus* experiments demonstrated that the Mcm2-7 helicase travels on the leading strand (Fu et al., 2011). Together, these steps assemble the pre-IC and activate the CMG complex to unwind DNA at the origin.

The majority of the proteins involved in CMG activation are conserved among eukaryotes [reviewed in (Tanaka and Araki, 2013)]. The metazoan homologues with different names than those described in *S. cerevisiae* are: RecQ4 (Sld2), Treslin/Ticrr (Sld3), TopBP1 (Dpd11), and Cyclin A/Cdk2 (S-CDK) [reviewed in (Tanaka et al., 2011)]. The CMG is conserved in higher eukaryotes (Masai et al., 2006; Aparicio et al., 2009; Moyer et al., 2006; Ilves et al., 2010). The steps involved in CMG activation in metazoans are similar to those in *S. cerevisiae*. In *Xenopus*, mice, and humans Cdc45 association with Mcm2-7 is dependent on DDK phosphorylation of Mcm2-7 subunits (Jares and Blow, 2000; Takahashi and Walter, 2005; Yoshizawa-Sugata et al., 2005; Masai et al., 2006; Tsuji et al., 2006). In *Drosophila*, the Mcm2-7 helicase is activated when Cdc45 and GINS become associated (Ilves et al., 2010).

1.6.2 DNA Polymerase Loading

Following helicase activation a series of regulated steps are required to load polymerases onto DNA for replication [reviewed in (Johansson and Dixon, 2013)]. The detailed steps to load the polymerases were first worked out using a Simian virus 40 (SV40) DNA replication assay. The assay was first described to support the replication of an exogenous plasmid in monkey cells infected with SV40 (Li and Kelly, 1984) and has been adapted to use cellular extracts from multiple species (Li and Kelly, 1985). Further optimization of the SV40 replication assay used purified SV40 T-antigen and template DNA containing a SV40 origin of replication (Stillman and Gluzman, 1985). This assay was instrumental in isolating the eukaryotic factors necessary for DNA replication and their stepwise assembly. The factors (RFC, RPA, PCNA, Pol δ, Pol ε, and Pol α/Primase) and the sequence of events required to load the polymerases are conserved down through prokaryotes [reviewed in (Hedglin et al., 2013a)].

The role and order of activity of each factor was elucidated using the SV40 assay [reviewed in (Johansson and Dixon, 2013)]. First, RPA (single-stranded DNA binding protein) is recruited to the unwound origin and stabilizes the single-stranded DNA (Wold and Kelly, 1988; Tanaka and Nasmyth, 1998). Next Pol α/ Primase are recruited to provide RNA primers for the DNA polymerases (Roth, 1987). The next step is the ATP dependent recruitment of RFC, a multi-subunit complex containing several ATPases, to the RPA coated DNA/RNA primer template junction (P/T junction) (Lee et al., 1991; Tsurimoto and Stillman, 1991). Next PCNA (proliferating cell nuclear antigen) is recruited to RFC and loaded onto DNA (Tsurimoto and Stillman, 1990; Tsurimoto and Stillman, 1991). Polymerases are not processive without a clamp (PCNA) that tethers them to DNA. Experiments with purified components indicate that without PCNA, polymerases only replicate a few base pairs before dissociating from the DNA (Prelich et al., 1987; Prelich and Stillman, 1988; Lee et al., 1989). The crystal structure of *S. cerevisiae* PCNA demonstrated that it is a ring-like structure with a central opening large enough for DNA (Krishna et al., 1994a; Krishna et al., 1994b). The crystal structure of the RCF/PCNA/DNA complex from *S. cerevisiae* shows how PCNA is loaded like a screw cap at the P/T junction (Bowman et al., 2004). More recent studies in human cells and single yeast cells using FRET suggest that RFC can also unload PCNA if polymerases are not loaded (Kumar et al., 2010; Hedglin et al., 2013b). The two polymerases are used for the bulk of DNA replication. Pol ε synthesizes the leading strand and Pol δ synthesizes the lagging strand [reviewed in (Johansson and Dixon, 2013)].

Together, the CMG, PCNA, and polymerase are collectively referred to as the replisome that travels away from the origin of replication and is at the replication fork **(Figure 2)** (Gambus et al., 2006; Pacek et al., 2006; Gambus et al., 2011).

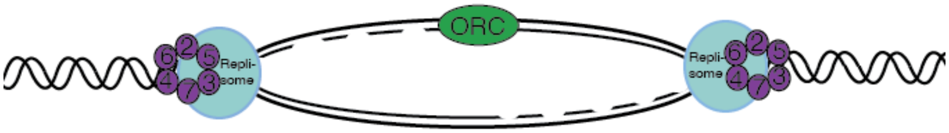

Figure 2: The Activated Mcm2-7 Helicase Travels Away from an Origin as Part of CMG and Replisome in S-Phase.

A cartoon depicting a single Mcm2-7 helicase traveling with the replisome at each replication fork away from ORC binding sites.

1.6.3 Timing of DNA Replication Initiation

Origins of replication are activated in a temporal manner in order to efficiently copy the entire genome within the confines of S-phase. The timing of origin initiation is reproducible between cell types and stable over several S-phases [reviewed in (Rhind and Gilbert, 2013)]. Additionally, chromosomal regions conserved between species are replicated at similar points in S-phase. Together, these results suggest that the timing of replication follows a programmed initiation pattern or replication timing program.

A determinant of the timing of replication initiation is the transcriptional activity of a given genomic region. Early labeling studies demonstrated that the two X chromosomes in a female mammalian cell are replicated differentially in S-phase; one replicates early and the other replicates late (Taylor, 1960; Latt, 1973). This timing information along with mouse genetics experiments indicating that one female X chromosome is genetically inactive suggested that the genetically active X is replicated earlier (Lyon, 1961). More recent studies have used whole genome technologies to assay the correlation between the replication timing program and transcriptional activity. Synchronized mammalian cells were pulsed with BrdU, a nucleotide analogue, at different time points in S-phase to incorporate into newly synthesized DNA (Goldman et al., 1984). The early and late replicating labeled regions of the genome were isolated and probed with actively transcribed cDNAs. The probes of actively transcribed cDNAs hybridized to early replicating regions of the genome. Similar timing studies found that replication timing was consistent over several cell cycles (Zink et al., 1999; Sadoni et al., 1999). A whole genome study, in metazoans, showed a correlation between replication timing and transcription. Pulsed cells were sorted into different S-phase populations by FACS and the BrdU labeled parts of the genome were hybridized onto expression microarrays (Schubeler et al., 2002). A genome-wide microarray based approach identified a correlation between origins of replication and transcription start sites in mouse ES cells (Sequeira-Mendes et al., 2009). Specifically, 85% of origins of replication were associated with genes and half of the origins of replication were also at transcription start sites (Sequeira-Mendes et al., 2009). Overall euchromatic or actively transcribed genomic regions are replicated early in S-phase.

The timing of replication initiation of a genomic region can be altered when the chromatin environment is changed. A study in *S. cerevisiae* showed that relocating a late replicating ARS away from the heterochromatic telomere changed its timing to earlier in S-phase (Ferguson and Fangman, 1992). Conversely, an early replicating ARS becomes later replicating when relocated proximal to a telomere, in *S. cerevisiae* (Friedman et al., 1996). The entire genome of *S. cerevisiae* replicated earlier when histone acetylation was increased, by genetic deletion of Rpd3, a histone deacetylase (Vogelauer et al., 2002). The same effect of histone acetylation on replication timing was found in *Drosophila* follicle cells and human erythroid cells (Aggarwal and Calvi, 2004; Goren et al., 2008). Together these results suggest that the local chromatin environment is a determinant of the timing of origin initiation.

Origins are also regulated by neighboring origins. In *S. cerevisiae*, replication was only initiated once on plasmid containing two identical ARSs suggesting that initiation of one origin inhibited the initiation of the other (Brewer and Fangman, 1994). Additional studies in *S. cerevisiae* showed that origin inhibition may be caused by passive replication, or the passing of a neighboring replication fork (Santocanale et al., 1999).

Mechanisms that control the replication timing have been proposed. It was shown, in *S. cerevisiae*, that Cdc45 associates with Mcm2-7 at early origins in G1 (Zou, 1998). Further experiments, in *S. cerevisiae*, found that Sld3 is only at early origins in G1 (Kanemaki and Labib, 2006). These results suggest that association of both Cdc45 and Sld3, in G1, primes an origin for CMG activation early in S-phase (Kanemaki and Labib, 2006). Another possible mechanism to prime origins to fire early in S-phase is a higher concentration of Mcm2-7. Whole genomic studies in *S. cerevisiae* and *Drosophila* suggest that genomic regions with dense Mcm2-7 association replicate early (Yang et al., 2010; Cayrou et al., 2011). Together, it is possible that a localized Mcm2-7 concentration increases the probability to recruit Cdc45 and Sld3. Subsequently, an origin with Cdc45 and Sld3 associated is primed for CMG activation upon entry into S-phase.

1.7 The 'MCM Paradox'

Every initiated origin of replication has two bidirectional replication forks traveling away from it. Therefore, each origin would require two Mcm2-7 helicases; one for each fork or a single double hexamer loaded. However, ORC loads multiple Mcm2-7 onto template DNA *in vitro* (Gillespie et al., 2001; Bowers et al., 2004; Evrin et al., 2013). Estimates from *Xenopus* experiments suggest that one molecule of ORC is chromatin bound every 10kb and one Mcm2-7 complex is bound every 700-1500 bps (Mahbubani et al., 1997; Gillespie et al., 2001). Quantitative western blot analysis from *Xenopus* suggest that 20-40 more Mcm2-7 complexes are bound to bait DNA than ORC (Edwards et al., 2002). These results suggest that a vast excess of Mcm2-7 are present on the chromatin than are needed to license an origin.

Consistent with an excess of Mcm2-7 loaded, Mcm2-7 are present on the chromatin throughout the nucleus and are localized away from ORC where they were initially loaded. Immunofluorescence microscopy studies show that Mcm2-7 are bound throughout the chromatin before replication (Madine et al., 1995; Krude et al., 1996; Romanowski et al., 1996a). ChIP experiments, in *Xenopus*, found that Mcm2-7 are located away from the single ORC binding site (Harvey and Newport, 2003). More recent electron microscopy studies demonstrated that multiple Mcm2-7 can be loaded by ORC and they are distributed throughout the template DNA (Evrin et al., 2009; Remus et al., 2009). Together these results suggest that ORC loads multiple Mcm2-7 that can translocate along dsDNA away from ORC.

Surprisingly, no cell cycle phenotype is evident when chromatin bound levels of Mcm2-7 are reduced to undetectable levels by RNAi in human cells (Ibarra et al., 2008). Similar Mcm2-7 reduction experiments in *Drosophila*, humans, and *C. elegans* also exhibit no cell cycle phenotype when whole cell Mcm2-7 protein and mRNA are reduced to undetectable levels (Woodward et al., 2006; Crevel et al., 2007; Ge et al., 2007).

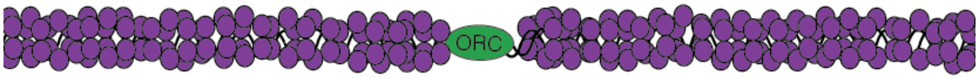

Figure 3: The 'MCM Paradox'.

A cartoon depicting the excess Mcm2-7 present on the chromatin relative to ORC.

Together, these seemingly paradoxical observations regarding the location and quantity of Mcm2-7 were termed the 'MCM Paradox' **(Figure 3)**. The excess loading and distribution of Mcm2-7 are important for maintaining genome stability during replicative stress. The ‘extra’ Mcm2-7 that are distally located away from ORC may be used as dormant origins as ORC is dispensable for S-phase after loading Mcm2-7 (Donovan et al., 1997; Rowles et al., 1999). The inter origin distances along DNA fibers are shorter when cells are challenged with replicative stress, in *Xenopus* and human tissue culture cells, suggesting the activation of dormant origins (Woodward et al., 2006; Ge et al., 2007). However, cells with reduced amounts of Mcm2-7 have a marked reduction in the utilization of dormant origins and an increase in cell death when S-phase is challenged by replicative stress (Woodward et al., 2006; Ge et al., 2007; Ibarra et al., 2008). Interestingly, dormant origins specifically in the replication foci that are experiencing stress are activated while other foci are inhibited (Ge and Blow, 2010).

Studies in mice indicate that the Mcm2-7 level is important for the overall health of an organism without experimentally induced replicative stress. Mice homozygous for the Mcm4^{Choas3} allele have a 40% reduction of Mcm2-7 and develop early onset tumors (Shima et al., 2007; Chuang et al., 2010). Thus, the amount of Mcm2-7 is critical for an organism suggesting that cells rely on dormant origins to maintain genomic stability.

1.8 Genomic Instability

Defects in the DNA replication program can drive genomic instability: rereplication, under replication, fragile sites, stalled replication forks, or mutations. One of many possible causes of genomic instability associated with the process of DNA replication are defects in sister chromatid cohesion.

Sister chromatids are paired together by cohesion at the time of their replication. Sister Chromatid Cohesion (SCC) ensures that the sister chromatids remain together until they are separated in mitosis. Cohesion creates tension between the chromosomes and the spindles to aid in chromatid separation during mitosis (Tanaka et al., 2000). SCC is mediated by the cohesin complex, a highly conserved complex among eukaryotes, consisting of Smc1, Smc3, Scc1/Mcd1/Rad21, and Scc3/SA (Anderson et al., 2002; Haering et al., 2002). The tripartite ring formed by the cohesin subunits (Smc1, Smc3, and Scc1) has been proposed to either encircle the two sister chromatids [reviewed in (Nasmyth, 2011)] or that two cohesin rings stably interact while encircling a single sister chromatid (Huang et al., 2005; Zhang et al., 2008).

The cohesin complex is loaded onto chromatin in G1, by Nipped-B (Scc2) and Mau-2 (Scc4), the adherin complex [reviewed in (Peters and Nishiyama, 2012)]. In *S. cerevisiae*, when Scc2 and Scc4 are mutated cohesin is not loaded (Toth et al., 1999; Ciosk et al., 2000). Also, *Drosophila* Nipped-B mutants have a reduced amount of sister chromatid cohesion (Rollins et al., 2004). *In vitro* studies in *Xenopus* implicate that the assembly of the pre-RC in G1 is necessary for cohesin loading (Gillespie and Hirano, 2004; Takahashi et al., 2004). This is further supported by evidence from *Xenopus* and human cells that indicate interactions between pre-RC components and cohesin components (Takahashi et al., 2008; Guillou et al., 2010).

Loading of the cohesin complex onto chromatin does not impart sister chromatid cohesion. Sister chromatid cohesion is established during DNA replication (Uhlmann and Nasmyth, 1998). Eco1, a highly conserved protein, acetylates Smc3 during S-phase to establish cohesion [reviewed in (Peters and Nishiyama, 2012)]. Studies in *S. cerevisiae* showed that mutants of the Eco1 homologue had separated sister chromatids prior to anaphase (Skibbens et al., 1999; Toth et al., 1999). More interestingly, the yeast Eco1 homologue interacts with PCNA, a replication fork component (Skibbens et al., 1999; Mayer, 2001; Kenna and Skibbens, 2003).

A potential consequence of defects in sister chromatid cohesion is aneuploidy [reviewed in (Torres et al., 2008)]. Aneuploidy is defined as a cell having unequal multiples of its genome [reviewed in (Storchova and Kuffer, 2008)]. Aneuploidy has been shown to be detrimental to an organism or a cell. Some well-known human syndromes are caused by aneuploidy (Craddock et al., 2010), for example, Down's syndrome patients are trisomic for chromosome 21 (Dierssen et al., 2009), Di George syndrome is caused by the loss of part chromosome 22, and Cri du Chat is caused by a deletion on chromosome 5 (Freeman et al., 2006; Zhang et al., 2009). In these syndromes a small amount of aneuploidy causes a significant phenotype. The phenotype is most likely a consequence of altered gene expression due to copy number variation (CNV) of important genes during development.

However, aneuploidy is present in cancer cells, which are highly proliferative [reviewed in (Schvartzman et al., 2010)]. A genome-wide high throughput analysis of 8000 cancer genomes revealed that some copy number variations are common among cancers (Kim et al., 2013). These results suggest a correlation between aneuploidy and cancer. However, the true nature of the relationship between aneuploidy and tumorigenesis, whether it is detrimental or beneficial, is still unclear [reviewed in (Holland and Cleveland, 2009)].

More strikingly, cells have evolved mechanisms to compensate for the copy number differences of genes between the sexes. In mammals, a female has two X chromosomes while a male has a single X chromosome. Female mammals undergo Xinactivation (heterchromatinization) to bring the levels of the active X chromosome gene expression equal to that of males. In *Drosophila*, male flies up regulate transcription from their single X chromosome to equal the transcription level of the two female X chromosomes in a process called dosage compensation [reviewed in (Gelbart and Kuroda, 2009)]. The mechanisms to adjust for sex chromosome gene dosage differences are well studied. Together these studies indicate that organisms possess mechanisms to potentially compensate for gene copy number differences introduced by aneuploidy of autosomes. However, it is unclear if cells are able to compensate for CNV on autosomes.

2. Dynamic Loading and Distribution of Mcm2-7 Throughout the Cell Cycle[1]

2.1 Introduction

The duplication of a eukaryotic genome within the confines of S-phase is a remarkable event. First, the sheer scale of the process -- tens to thousands of million base pairs of DNA need to be precisely copied once and only once within just a few hours. In addition, DNA replication forks must initiate within and progress through diverse local chromatin environments including accessible euchromatin and repressive heterochromatin. And finally, the process has to be dynamic and capable of responding to environmental and developmental cues. Thus, every cell cycle, thousands of DNA replication start sites (origins) must be selected and activated in a regulated and coordinated manner to ensure that the entire genome is faithfully duplicated [reviewed in (Masai et al., 2010)].

Each potential origin of replication is marked by the origin recognition complex (ORC) (Bell and Stillman, 1992; Rao and Stillman, 1995). In G1, the Mcm2-7 complex, the replicative helicase, is loaded as a double hexamer at ORC binding sites in an ORC, Cdc6, and Cdt1 dependent manner to form the pre-replicative complex (pre-RC) (Evrin et al., 2009; Remus et al., 2009; Gambus et al., 2011). The assembly of the pre-RC in G1 'licenses' the origin for potential activation in the subsequent S-phase. As a cell enters S-phase, cyclin dependent kinase (CDK) and Dbf4-dependent kinase (DDK) activate the Mcm2-7 helicase by the recruitment of Cdc45 and the GINS complex to form the CMG (Cdc45, Mcm2-7, and Gins) holo-helicase complex [reviewed in (Tanaka and Araki, 2013)]. After initiation of DNA replication, the Mcm2-7 helicase, as part of the CMG complex, travels with and unwinds the DNA ahead of the replication fork (Aparicio et al., 1997; Labib, 2000; Pacek and Walter, 2004; Pacek et al., 2006; Sekedat et al., 2010).

---

[1] Under revision for EMBO.

In theory, a double hexamer of Mcm2-7 loaded at each origin should be sufficient to replicate the genome, with a single hexamer of Mcm2-7 traveling with each bidirectional DNA replication fork. However, nuclear Mcm2-7 protein levels are in vast excess relative to ORC or the number of replication origins (Burkhart et al., 1995; Lei et al., 1996; Donovan et al., 1997; Mahbubani et al., 1997; Edwards et al., 2002). *In vitro*, multiple double hexamers of Mcm2-7 are able to be loaded at ORC binding sites by reiterative rounds of ATP-hydrolysis (Bowers et al., 2004; Evrin et al., 2013). Electron microscopy revealed that these double hexamers are distributed throughout the DNA template (Remus et al., 2009; Ervin et al., 2009) suggesting the ability of the Mcm2-7 complex to translocate away from ORC binding sites. Similarly, *in vivo* immunofluorescence and chromatin association studies have revealed that Mcm2-7 are broadly distributed throughout the nucleus and exhibit little co-localization with ORC (Madine et al., 1995; Krude et al., 1996; Romanowski et al., 1996a; Ritzi et al., 1998; Edwards et al., 2002; Harvey and Newport, 2003). Together, these seemingly paradoxical observations regarding the location and quantity of Mcm2-7 have been termed the 'MCM Paradox' (Takahashi et al., 2005).

Although the mechanisms regulating the loading of multiple Mcm2-7 complexes and their distribution throughout the genome are unclear, increasing data suggests that the excess Mcm2-7 are important for maintaining genome stability during replicative stress. Mcm2-7 levels can be depleted by more than 90% with little, if any, impact on progression through S-phase (Woodward et al., 2006; Crevel et al., 2007; Ge et al., 2007; Ibarra et al., 2008). However, in the absence of the full complement of Mcm2-7, there is a marked reduction in the utilization of dormant origins and an increase in cell death when cells encounter replicative stress during S-phase (Woodward et al., 2006; Ge et al., 2007; Ibarra et al., 2008). Thus, the full complement of Mcm2-7 is critical for the activation of dormant replication origins and functions to preserve genome integrity during replicative stress.

We set out to examine this 'MCM Paradox' from a biochemical and genome-wide perspective by using chromatin association assays and genome-wide chromatin immunoprecipitation (ChIP) to quantify the loading and distribution of Mcm2-7 at different points in the cell cycle of *Drosophila* Kc cells. Specific questions we set out to address included: i) How does Mcm2-7 loading onto chromatin progress during G1? ii) Do Mcm2-7 load adjacent to ORC? iii) Where are Mcm2-7 distributed throughout the genome? and iv) Do other chromosomal features (eg. transcription units) impact the distribution of Mcm2-7?

The Mcm2-7 complex is loaded onto DNA via an ORC, Dup/Cdt1, and Cdc6 dependent mechanism at ORC binding sites. Mcm2-7 complex loading has been shown to begin in G2 and continue through G1. Due to the nature of the helicase and the 'MCM Paradox' the localization and distribution of theses complexes at different cell cycle time points is not known on a genome wide level. To this end, we wanted to comprehensively locate Mcm2-7 complexes immediately prior to S-phase activation using *Drosophila* Kc cells.

2.2 Mcm2-7 Loading Occurs in Two Distinct Phases During G1

Mcm2-7 chromatin association is dynamic throughout the cell cycle. Mcm2-7 are loaded onto chromatin in G1, travel ahead of the replication fork in S-phase, and are removed from the chromatin as S-phase progresses [reviewed in (Masai et al., 2010)]. In order to examine the dynamics of Mcm2-7 chromatin association, we developed approaches to arrest *Drosophila* Kc cells at specific points in the cell cycle when Mcm2-7 would presumably be associated or not associated with ORC at potential origins of replication. We were particularly interested in identifying a defined cell cycle point immediately before Mcm2-7 helicase activation and its movement with the DNA replication fork.

We reasoned that inhibition of cyclin E/Cdk2 kinase activity would be sufficient to arrest cells in G1 immediately prior to origin activation. We first depleted cyclin E and Cdk2 separately using RNAi in *Drosophila* Kc cells. In *Drosophila*, cyclin E is the regulatory subunit of the cyclin dependent kinase 2 (Cdk2), both of which, like their mammalian homologues, function together to drive cells out of G1 and into S-phase (Dulic et al., 1992; Ohtsubo and Roberts, 1993; Knoblich et al., 1994). We also overexpressed Dacapo, a p27 homolog and potent inhibitor of cyclin E/Cdk2 kinase activity (Lane et al., 1996), from a copper responsive metallothionein promoter. As controls, cells were arrested in early G1 by RNAi depletion of Dup/Cdt1, a key replication licensing factor required for Mcm2-7 loading (Whittaker et al., 2000) and at the G1/S transition by treatment with 1 mM hydroxyurea (HU). Dup/Cdt1 RNAi, cyclin E RNAi, Cdk2 RNAi, Dacapo overexpression (+Dacapo), and HU treatment all resulted in a cell cycle arrest with the majority of cells having a 2C DNA content by flow cytometry **(Figure 4A)**.

We next assessed the relative amounts of chromatin associated Mcm2-7 in each of these conditions by chromatin fractionation **(Figure 4B)**. A polyclonal antibody specific to *Drosophila* Orc2 (Austin et al., 1999) and a monoclonal antibody (AS1.1) that recognizes all six Mcm2-7 subunits (Chen et al., 2007) were used to assess the chromatin

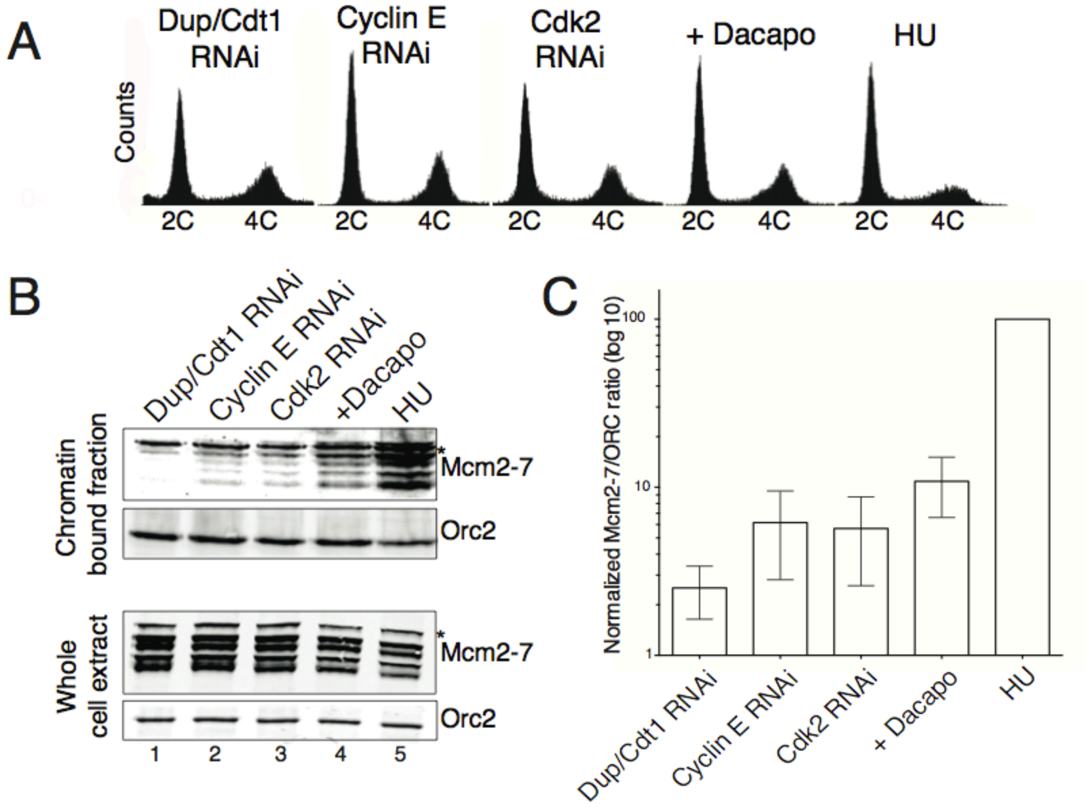

Figure 4: Mcm2-7 Loading Occurs in Two Distinct Phases during G1.

A. FACs profiles of DNA content for cells arrested at different points in the cell cycle by Dup/Cdt1 RNAi, cyclin E RNAi, Cdk2 RNAi, Dacapo overexpression (+Dacapo), and 1 mM HU. **B.** Analysis of Mcm2-7 chromatin association at different points in the cell cycle. Chromatin bound fractions and whole cell extracts were assayed for Orc2 and Mcm2-7 by western blot. Orc2 was used as a loading control and as an indicator of ORC levels. A non-specific band is indicated by an asterisk (*). **C.** Quantification of the ratio of chromatin bound Mcm2-7 relative to Orc2 ($\log_{10}$ scale) was determined from at least three independent experiments.

associated and whole cell extract levels of ORC and Mcm2-7. As expected, cells lacking the licensing factor, Dup/Cdt1, exhibited no detectable chromatin associated Mcm2-7 **(Figure 4B; lane 1)**. The background levels of Mcm2-7 chromatin association in the Dup/Cdt1 depleted cells were similar to the levels observed in G2 arrested cells where the presence of geminin prevents pre-RC formation (McGarry and Kirschner, 1998) **(Figure 5)**. In contrast, cells arrested at the G1/S transition by HU treatment had a robust association of Mcm2-7 with chromatin **(Figure 4B; lane 5)**. Interestingly, cyclin E and Cdk2 RNAi cells exhibited an intermediate phenotype **(Figure 4B; lanes 2 and 3)** with considerably less chromatin bound Mcm2-7 than HU arrested cells. We observed a slight increase in the amount of chromatin bound Mcm2-7 in +Dacapo cells **(Figure 4B; lane 4)** over both cyclin E and Cdk2 RNAi cells. However, the amount of chromatin bound Mcm2-7 in +Dacapo cells was still markedly reduced compared to levels seen in HU treated cells. In contrast to the cell cycle fluctuations in chromatin bound Mcm2-7 levels, ORC remained constant for each condition surveyed. Importantly, the cell cycle differences observed in Mcm2-7 chromatin association were not simply due to changes in protein levels as the levels of ORC and Mcm2-7 remained constant in the whole cell extract. The chromatin associated Mcm2-7/ORC ratio was calculated for each condition, normalized to the level in HU to obtain a fold difference and plotted on a $\log_{10}$ scale **(Figure 4C)**. A 40-fold difference in chromatin associated Mcm2-7 was observed between Dup/Cdt1 RNAi and HU and 9-to 15-fold differences in loading were observed between HU and cells with impaired cyclin E/Cdk2 kinase activity (cyclin E RNAi, Cdk2 RNAi, and +Dacapo).

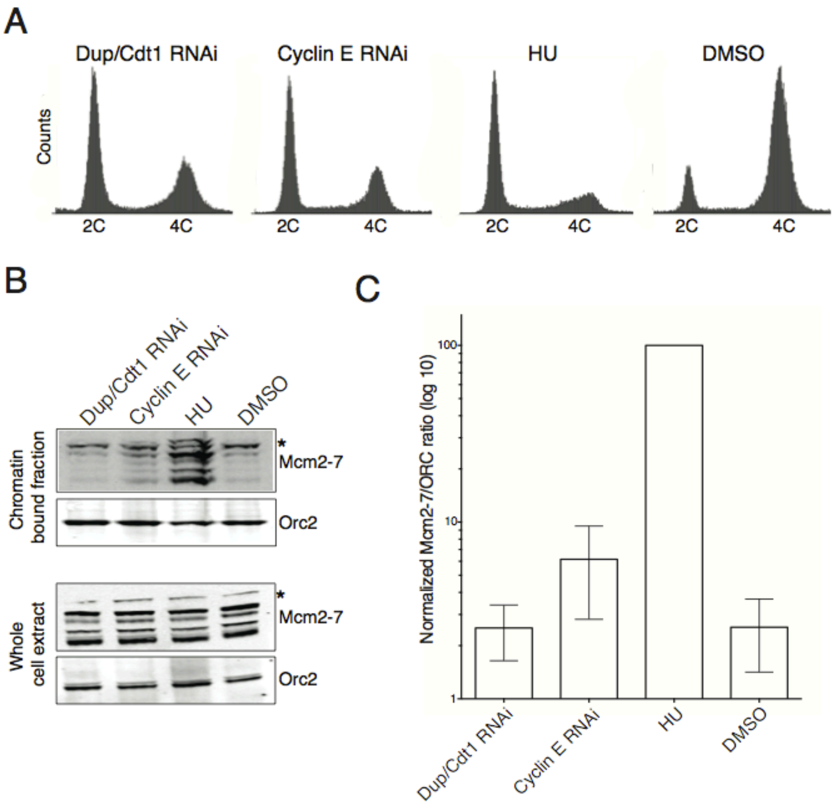

Figure 5: Mcm2-7 Chromatin Bound Levels are Similar in Dup/Cdt1 RNAi and G2 Arrested Cells.

A. DNA content FACs profiles of each cell population used to determine Mcm2-7 chromatin loading. **B.** Whole cell extracts and chromatin bound fractions were assayed for Orc2 and Mcm2-7 by western blot. A non-specific band is indicated by an asterisk (*). **C.** The Mcm2-7/ ORC ratios from each condition were normalized to the HU ratio for each independent blot and plotted on a $\log_{10}$ scale.

2.3 Cyclin E Partially Rescues Mcm2-7 Loading Independently of Kinase Activity In Vivo

Our results suggest that cyclin E/Cdk2 kinase activity is required to load the full complement of Mcm2-7 in mitotically cycling *Drosophila* Kc cells. In mammalian cells exiting quiescence, there is provocative data that may suggest a role for cyclin E, independent of any kinase activity, in the assembly of the pre-RC (Geng et al., 2007). Our prior experiments were not designed to uncouple a potential cyclin E independent role from cyclin E/Cdk2 kinase activity in Mcm2-7 loading. By depleting cells of Cdk2, we also likely destabilized cyclin E (Clurman et al., 1996; Won and Reed, 1996). To examine the role of cyclin E independently of Cdk2 kinase activity in *Drosophila* pre-RC assembly, a metallothionein regulated FLAG-tagged cyclin E construct was stably transfected into *Drosophila* Kc cells. This allowed us to deplete cells of Cdk2 by RNAi and then subsequently overexpress FLAG-cyclin E by the addition of copper to the medium. Copper induced overexpression of FLAG-cyclin E was sufficient to drive entry into S-phase in control cells treated with non-specific RNAi (pUC) **(Figure 6A)**. In contrast, overexpression of FLAG-cyclin E in the absence of Cdk2 expression was not sufficient to drive cells into S-phase.

Using this approach, we tested whether FLAG-cyclin E overexpression, in the absence of Cdk2 expression, would be sufficient to promote Mcm2-7 chromatin loading. Overexpression of FLAG-cyclin E in Cdk2 RNAi cells resulted in a slight (~3-fold), but reproducible, increase in chromatin bound Mcm2-7 over Cdk2 RNAi alone **(Figure 6B)**. To ensure that there was no residual or compensatory CDK activity associated with FLAG-cyclin E overexpression in Cdk2 depleted cells, we performed an *in vitro* kinase assay using FLAG immunoprecipitated extracts and histone H1 **(Figure 6C)**. Robust phosphorylation of histone H1 was observed only when FLAG-cyclin E was

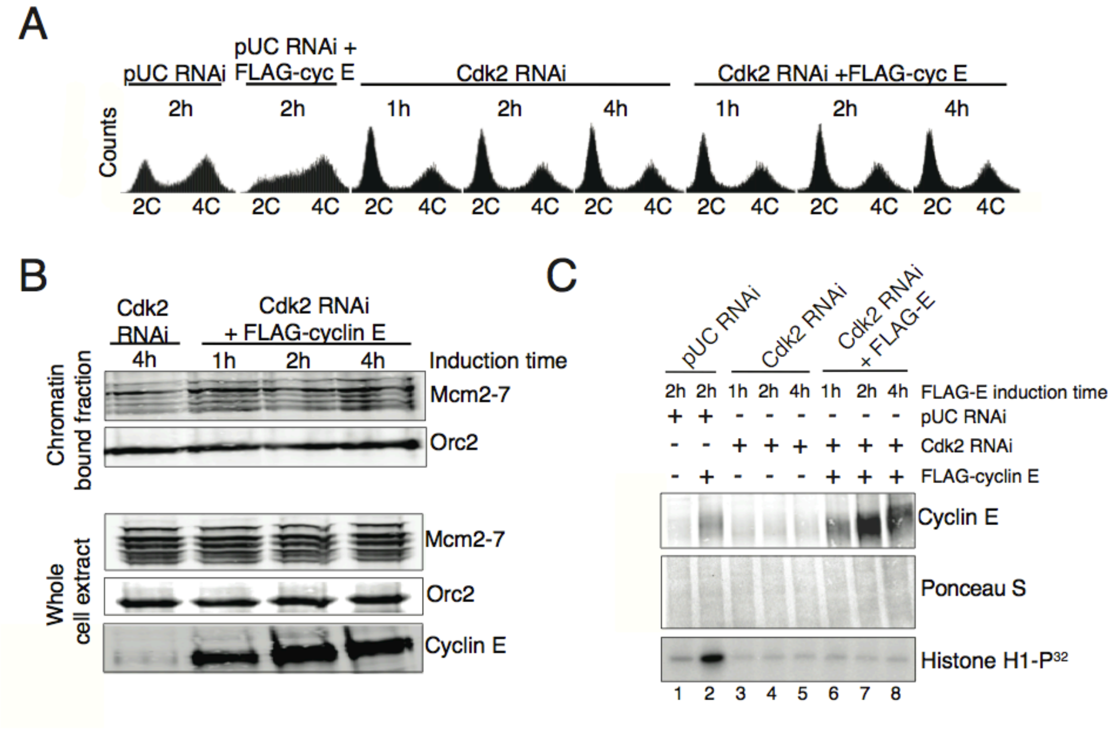

Figure 6: FLAG-cyclin E Overexpression in the Absence of Cdk2 Kinase Activity can Stimulate Mcm2-7 Loading.

A. FACs profiles of cell populations used to determine histone H1 kinase activity and Mcm2-7 loading in absence of Cdk2 while overexpressing FLAG-cyclin E. **B.** Whole cell extracts were assayed for Mcm2-7, Orc2, and cyclin E and the chromatin bound fractions were assayed for Orc2 and Mcm2-7 by western blot. The asterisk (*) indicates a nonspecific band. **C.** Histone H1 kinase assay using FLAG-cyclin E immunoprecipitates to examine cyclin E mediated kinase activity indicated by P^{32} signal. Whole cell lysates were balanced by Ponceau S staining and probed for cyclin E overexpression by western blot.

overexpressed and Cdk2 was present. We conclude that cyclin E may have a very limited Cdk2 independent role in promoting pre-RC assembly. However, cyclin E overexpression in the absence of Cdk2 was clearly not sufficient to load the full complement of Mcm2-7 on the chromatin. Together, our results suggest that loading of the full complement of Mcm2-7 occurs in two phases; a cyclin E independent phase followed by a second phase that requires cyclin E/Cdk2 kinase activity.

2.4 Maximal Loading of Mcm2-7 Coincides with Entry into S-phase

Our results suggest that Mcm2-7 levels continue to increase during G1 concomitant with increasing Cdk2 kinase activity prior to entry into S-phase. However, it remained possible that the maximal Mcm2-7 loading we observed was due to replicative stress resulting from HU treatment (Anglana et al., 2003). Thus, we wanted to assess Mcm2-7 loading without perturbing S-phase or activating the intra-S-phase checkpoint.

To monitor pre-RC assembly as cells synchronously enter S-phase from G1, we induced overexpression of Dacapo to block cyclin E/Cdk2 kinase activity and prevent the full complement of Mcm2-7 from being assembled on the chromatin in G1 **(Figure 7)**. We then released the cells from the Dacapo arrest by removing the inducer (Cu2+) of Dacapo expression. Entry into the next cell cycle was prevented by the addition of the mitotic inhibitor, colcemid, to the medium. As cells progressed into S-phase, we monitored DNA content, Dacapo protein levels, and the chromatin associated levels of Mcm2-7 and ORC. We found that most cells entered S-phase by 3 hours and had completed DNA replication by 9 hours. Entry into S-phase at 3 hours coincided with a decrease in Dacapo expression and a marked increase (~9-fold) in Mcm2-7 association

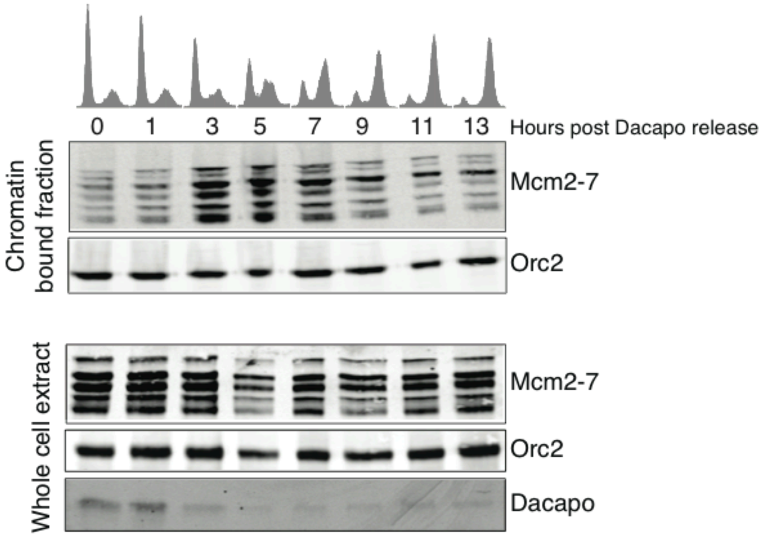

Figure 7: Mcm2-7 Loading in an Unperturbed G1-S Transition.

Cell cycle analysis by FACs of cells arrested in G1 by Dacapo overexpression followed by release back into the cell cycle for 13 h (top). Western blot analysis of Mcm2-7 and Orc2 for the chromatin bound fractions and whole cell extracts of Mcm2-7, Orc2, and Dacapo.

with the chromatin. The increase in Mcm2-7 chromatin association at the onset of S-phase entry was transient as Mcm2-7 were likely removed from chromatin with passage of the DNA replication fork. Together, these results demonstrate that the cyclin E/Cdk2 kinase dependent loading of the full complement of Mcm2-7 is a regulated process that occurs during a normal cell cycle.

2.5 The Cyclin E/Cdk2 Kinase Dependent Phase of Mcm2-7 Loading Requires the Canonical Pre-RC Assembly Pathway

It is well documented that ORC, Cdc6, and Cdt1 are able to load multiple Mcm2-7 double hexamers onto DNA templates in *vitro* (Bowers et al., 2004; Evrin et al., 2009; Remus et al., 2009). It was evident from our biochemical studies that Dup/Cdt1 was required to load minimal amounts of Mcm2-7 in a cyclin E independent manner **(Figures 5B)**. However, it remained unclear whether the full complement of Mcm2-7 loading we observed, which is stimulated by cyclin E/Cdk2 kinase activity, was dependent on the canonical pre-RC assembly pathway (eg. Cdc6 and Dup/Cdt1). To determine if the full complement of Mcm2-7 loading we observed in HU arrested cells is dependent on the established pre-RC assembly pathway, we depleted cells of Cdc6 and Dup/Cdt1 immediately after the cyclin E/Cdk2 independent pre-RC assembly step **(Figure 8A)**. Specifically, we arrested cells in early G1 by RNAi depletion of cyclin E for 24 hours followed by treatment with either control (pUC) or Cdc6 and Dup/Cdt1 dsRNA for 24 hours. Cyclin E was subsequently restored by overexpression of RNAi resistant FLAG-cyclin E. To assess Mcm2-7 loading at a static point in the cell cycle immediately after entry into S-phase, the cells were released into medium containing 1 mM HU **(Figure 8A)**. As expected, the majority of cells from each condition arrested with a 2C DNA content **(Figure 8B)**.

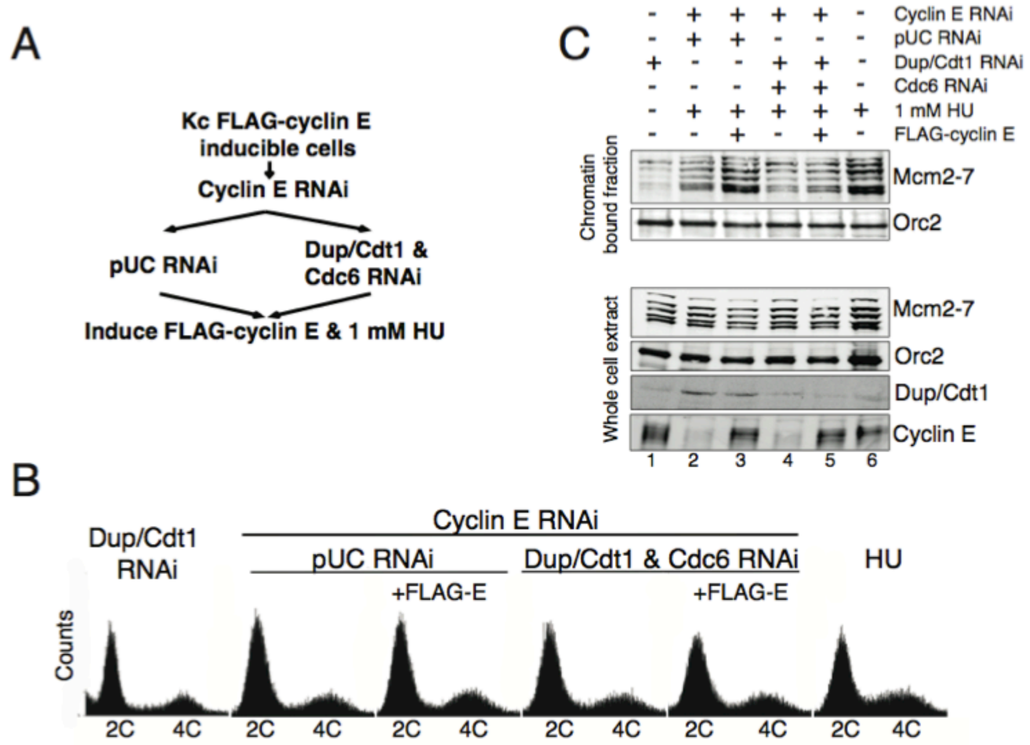

Figure 8: The Cyclin E/Cdk2 Kinase Dependent Phase of Mcm2-7 Loading Requires the Canonical Pre-RC Assembly Pathway.

A. Schematic of the experiment. **B.** FACs profile of DNA content for each cell population assayed. **C.** Western blot analysis for Orc2 and Mcm2-7 in the chromatin bound fractions and Orc2, Mcm2-7, Dup/Cdt1, and cyclin E in the whole cell extract for each condition assayed.

We found that cells depleted of Dup/Cdt1 and Cdc6 were unable to load additional Mcm2-7 when cyclin E expression and progression through G1 were restored by FLAG-cyclin E overexpression **(Figure 8C; lanes 4 and 5)**. In contrast, control RNAi (pUC) cells were able to load additional Mcm2-7 when driven through G1 by FLAG-cyclin E overexpression **(Figure 8C; lane 3)**. The increase of chromatin bound Mcm2-7 was not due to a difference in FLAG-cyclin E overexpression. These results indicate that Dup/Cdt1 and Cdc6 are needed to load the full complement of Mcm2-7 onto chromatin throughout G1 in a regulated manner.

2.6 Genome-wide Distribution of Mcm2-7 is Determined by Cyclin E/Cdk2 Activity and Local Transcription

Earlier immunofluorescence based studies noted that Mcm2-7 do not strictly co-localize with ORC in the nucleus during S-phase progression (Madine et al., 1995; Krude et al., 1996; Romanowski et al., 1996a). The dichotomy in Mcm2-7 chromatin association that we observed between cells arrested by cyclin E RNAi and cells arrested at the G1/S transition by HU treatment prompted us to investigate the genome-wide distribution of the Mcm2-7 relative to ORC. Specifically, we used chromatin immunoprecipitation to address where Mcm2-7 was localized in relation to ORC in cyclin E and HU arrested cells.

We used genome-wide mapping experiments to localize ORC in asynchronous cells (MacAlpine et al., 2010; Roy et al., 2010) **(Figure 9A)** and Mcm2-7 in cyclin E RNAi **(Figure 9B)** treated cells. We identified 5,135 ORC peaks and 3,792 Mcm2-7 peaks throughout the *Drosophila* Kc genome. The concordance between ORC and Mcm2-7 in cyclin E RNAi treated cells was greater than 93%, consistent with the critical role of ORC

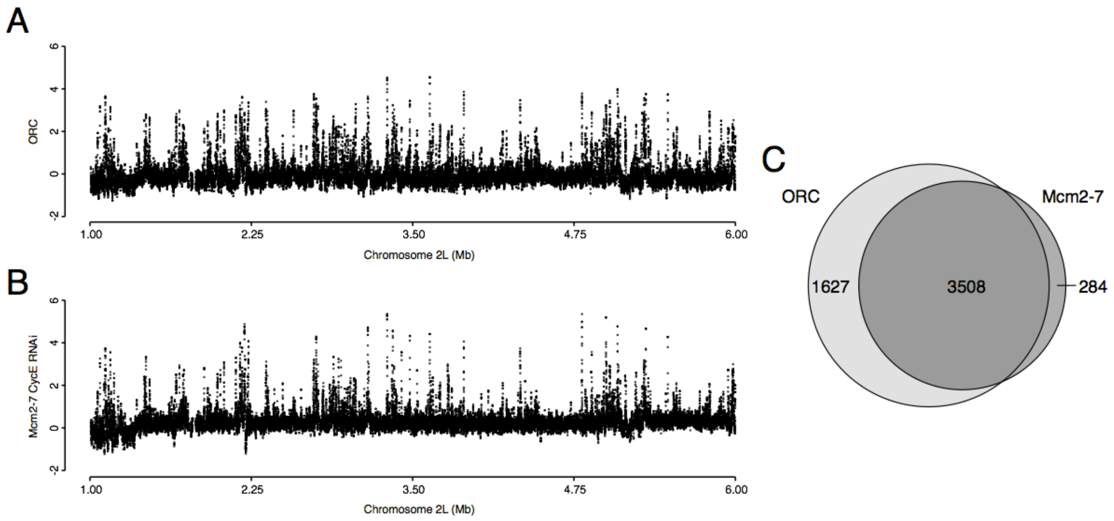

Figure 9: Mcm2-7 Loading Occurs at ORC Binding Sites in Cyclin E RNAi Cells.

A. Genome-wide analysis of ORC localization by ChIP-chip. ORC enrichment from asynchronous cells is depicted for a 4 Mb section of chromosome 2L. **B.** Genome-wide analysis of Mcm2-7 localization in early G1 by ChIP-chip. Mcm2-7 enrichment from cyclin E RNAi depleted cells is depicted for a 4 Mb section of chromosome 2L. **C.** Venn diagram depicting the overlap between ORC and Mcm2-7 peaks.

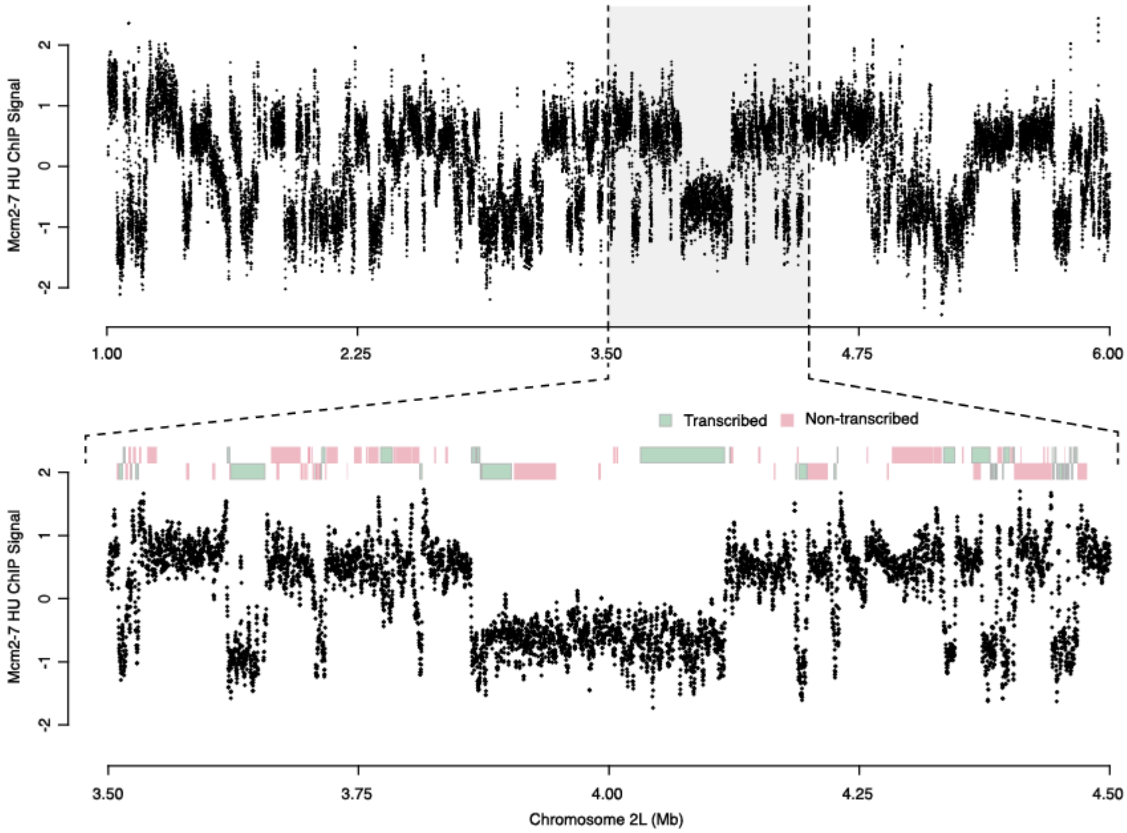

Figure 10: The Mcm2-7 Chromatin Distribution in G1 is Dependent on Cyclin E/Cdk2 Activity.

Genome-wide analysis of Mcm2-7 localization at the G1/S transition by ChIP-chip. Mcm2-7 enrichment from HU arrested cells is depicted for a 4 Mb section of chromosome 2L. Inset: transcribed (green) and non-transcribed genes (red) are indicated above with genes on the positive strand on the top and those on the negative on the bottom.

in pre-RC assembly **(Figure 9C)**. Strikingly, we found a very different pattern of Mcm2-7 localization at the G1/S transition during an HU arrest **(Figure 10)**. In contrast to the tight co-localization with ORC observed in the cyclin E depleted cells, we observed a 'binary' pattern of Mcm2-7 localization across the genome. Specifically, we observed broad chromosomal regions containing Mcm2-7 signal punctuated by the absence of Mcm2-7 localization. Together with our biochemical results, these data suggest that the full complement of Mcm2-7 we observed at the G1/S transition has re-distributed from ORC binding sites. We do not believe that Mcm2-7 are completely coating the DNA, but rather that we are observing the likelihood of detecting Mcm2-7 signal within specific genomic features (see below).

The binary distribution of Mcm2-7 across the genome prompted us to investigate genomic features that may be associated with the broad regions of high or low Mcm2-7 levels along the chromosome. The Mcm2-7 localization pattern relative to annotated genomic features suggested that Mcm2-7 may be displaced by actively transcribed genes **(Figure 10)**. To quantitatively assess the Mcm2-7 distribution relative to transcription units, we generated histograms of Mcm2-7 enrichment for transcribed and non-transcribed genes **(Figure 11; left panel)** and found a bimodal pattern of Mcm2-7 enrichment. Specifically, active genes had no or very little Mcm2-7 signal whereas inactive or non-transcribed genes exhibited an elevated Mcm2-7 signal ($p < 1.02e-257$; $t=40.16$). We also considered that the bimodal distribution of Mcm2-7 enrichment between active and inactive genes might be due to the activation of early origins that are enriched near actively transcribed genes. However, we found that the bimodal distribution for Mcm2-7 was not dependent on early origin activity, but instead was a specific feature of annotated transcripts **(Figure 12)**.

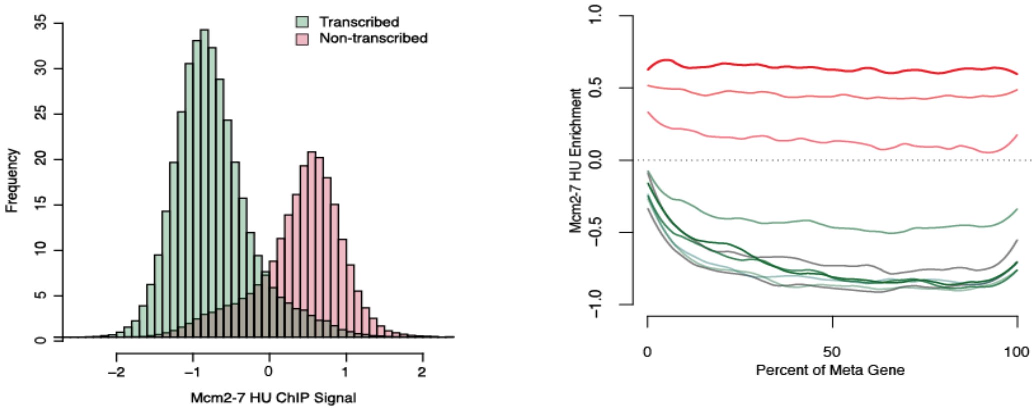

Figure 11: The Mcm2-7 Chromatin Distribution in G1 is Dependent on the Transcription Machinery.

LEFT. Bimodal distribution of Mcm2-7 enrichment over transcribed and non-transcribed genes in HU arrested cells. Histogram showing the distribution of probe scores found within transcribed (green) and non-transcribed (red) genes. **RIGHT.** 'Meta' gene analysis of Mcm2-7 enrichment for different deciles of gene expression and their aggregated probe intensities were plotted.

The enrichment of Mcm2-7 signal at non-transcribed genes was indistinguishable from intergenic levels **(Figure 13)** suggesting that active transcription was responsible for displacing Mcm2-7 from chromatin. We reasoned that if active transcription was sufficient to displace Mcm2-7 from chromatin, then in the absence of additional Mcm2-7 loading we should expect that both low and high transcription levels should be able to displace Mcm2-7 from chromatin. To address this, we binned the 14,594 genes into deciles based on their relative gene expression levels and plotted the Mcm2-7 enrichment relative to a 'meta' gene body **(Figure 11, right panel)**. We found that Mcm2-7 levels were depleted over the entire transcription unit for actively transcribed genes. The seven most expressed deciles (green and gray lines) exhibited the same degree of Mcm2-7 depletion suggesting that any amount of transcription is sufficient to displace Mcm2-7 from the chromatin. Together, these results suggest that individual Mcm2-7 complexes can be displaced by active transcription and that after the G1/S transition they cannot be re-established or translocate into these regions.

The Mcm2-7 helicase complex is inherently dynamic throughout the cell cycle; it is loaded onto chromatin, travels with the replication fork, and is rapidly removed from the chromatin by replication (Madine et al., 1995; Krude et al., 1996; Romanowski et al., 1996b). Our biochemical and genomic data indicate that the full complement of Mcm2-7 is loaded onto chromatin in late G1 and redistributes throughout non-transcribed genes and intergenic sequences by the G1/S transition. If this broad distribution of chromatin associated Mcm2-7 at the G1/S transition represents the true biological distribution of the Mcm2-7 complex, then we would expect that these Mcm2-7 complexes would be been displaced from early replicating regions of the genome and only remained at those

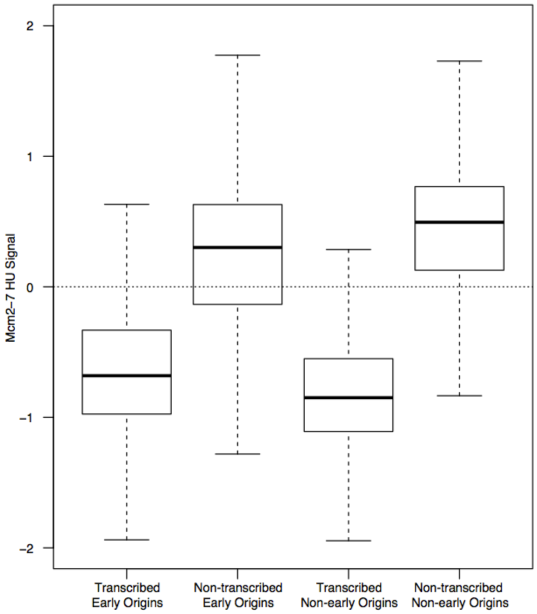

Figure 12: Bimodal Enrichment of Mcm2-7 at Active and Inactive Transcripts is not Dependent on Early Origin Activation.

Box-plots of Mcm2-7 enrichment for actively transcribed and non-transcribed genes grouped by their overlap with early origins of DNA replication.

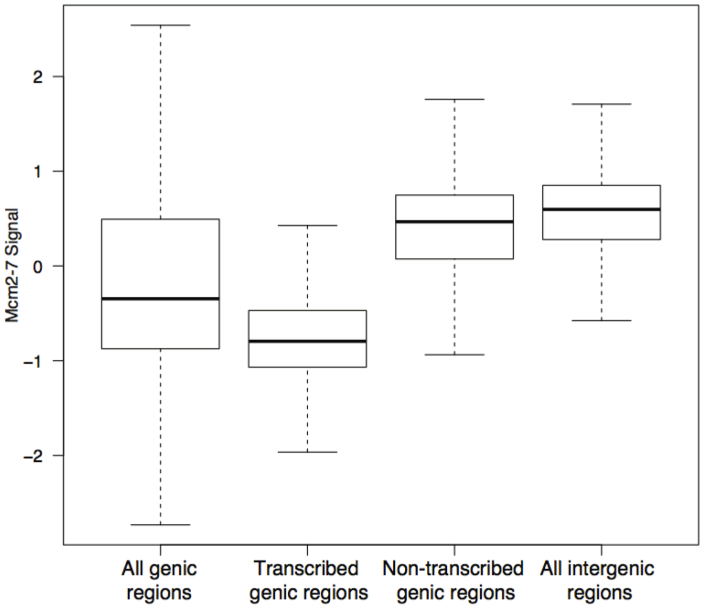

Figure 13: Mcm2-7 Enrichment in Non-transcribed Genes is Indistinguishable from Intergenic Levels.

Box-plots of Mcm2-7 signal over genic regions and intergenic regions. Non-transcribed genic regions have a similar level of Mcm2-7 enrichment signal as intergenic regions.

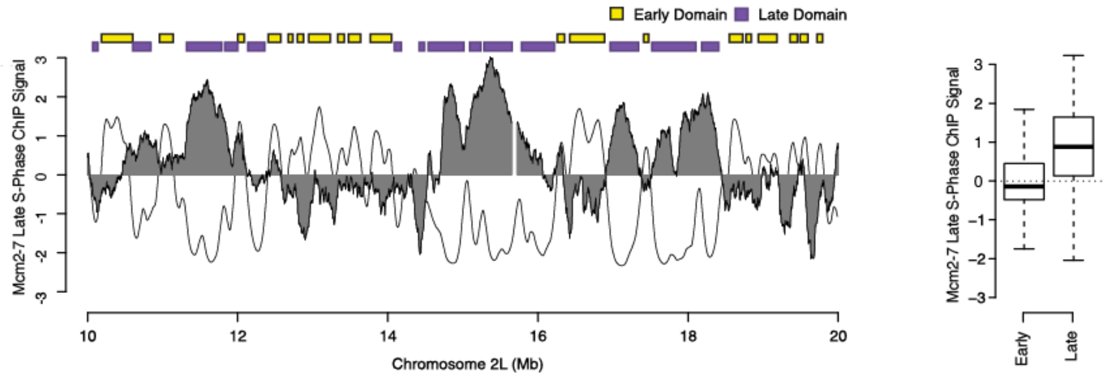

Figure 14: Mcm2-7 are Displaced from Chromatin by DNA Replication.

LEFT. Genome-wide analysis of Mcm2-7 localization in late S-phase (6 h post HU release) by ChIP-chip. Mcm2-7 enrichment is depicted for a 4 Mb section of chromosome 2L (filled gray), replication timing profile (black line) and early (yellow) and late (purple) replication timing domains. **RIGHT.** Box-plots representing late S-phase Mcm2-7 ChIP signal found within early or late replication domains.

sequences copied in late S-phase **(Figure 14; left panel, black line and purple segments)**.To quantify these results we used our prior segmentation of the genome into displaced by replication during S-phase. To test this hypothesis, we arrested cells at the G1/S transition by treatment with 1 mM HU and then released them synchronously into S-phase. We surveyed the genome-wide distribution of Mcm2-7 near the end of S-phase (6 hours after HU release) **(Figure 14; left panel, dark gray)**. We found that Mcm2-7 had early (yellow) and late (purple) replicating domains (Lubelsky et al., In press) and compared it to the distribution of Mcm2-7 at the end of S-phase. We found a significant enrichment of Mcm2-7 associated with late replicating domains relative to early domains **(Figure 14, right panel;** $p < 2.37e-68$; $t=18.22$**).** Together, these data argue that the broad distribution of Mcm2-7 we observe at the G1/S transition represents the full and functional complement of Mcm2-7 on the chromatin.

2.7 Discussion

In order to ensure the fidelity of the DNA replication program and genome integrity, there are many more Mcm2-7 helicase complexes loaded onto DNA than potential start sites marked by ORC. We have examined this 'MCM Paradox' using both genomic and biochemical approaches to understand the mechanisms by which excess Mcm2-7 are loaded and distributed throughout the genome to preserve genomic integrity. We found that during G1, Mcm2-7 are loaded onto chromatin in two distinct phases -- both of which rely on the canonical pre-RC assembly pathway including Dup/Cdt1 and Cdc6 **(Figure 15)**. In the first phase of pre-RC assembly, a minimal level of Mcm2-7 is loaded throughout the genome at ORC binding sites in a cyclin E and Cdk2 independent manner. Subsequently, there is a second wave of pre-RC assembly that is dependent on

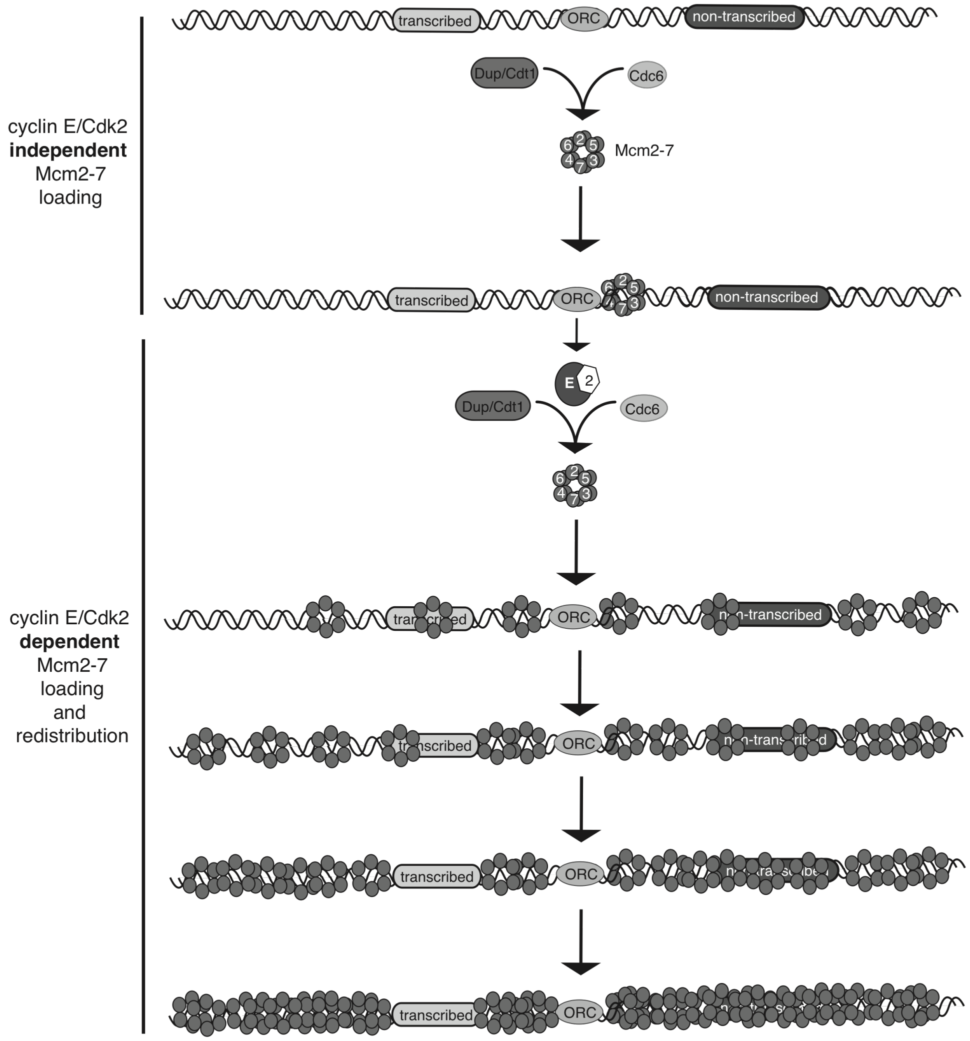
transcribed
ORC
non-transcribed
Dup/Cdt1
Cdc6
Mcm2-7
6 2 5 4 7 3
cyclin E/Cdk2
independent
Mcm2-7
loading
E
2
cyclin E/Cdk2
dependent
Mcm2-7
loading
and
redistribution

Figure 15: Model of Pre-RC Assembly and Redistribution of Mcm2-7 in G1.

The first phase of Mcm2-7 loading proximal to ORC binding sites occurs independently of cyclin E/Cdk2. The second phase of Mcm2-7 loading is dependent on cyclin E/Cdk2 kinase activity and results in the full complement of Mcm2-7 being loaded onto the chromatin. Prior to the G1/S transition the full complement of Mcm2-7 has been redistributed along the chromosomes and displaced from transcribed genes.

cyclin E/Cdk2 kinase activity and which is required for the loading of the full complement of Mcm2-7 onto chromatin. Strikingly, the full complement of Mcm2-7 when loaded onto chromatin is not restricted to sequences immediately adjacent to ORC binding sites, but rather distributed throughout the genome and shaped by active transcription.

2.7.1 Precise Regulation of Cyclin E/Cdk2 Kinase Activity is Critical for Pre-RC Assembly and Genomic Stability

Pre-RC assembly begins in telophase (Dimitrova et al., 1999) continues through G1 and culminates just prior to entry into S-phase (Symeonidou et al., 2013). Here, we have shown that there are distinct phases of pre-RC assembly both independent and dependent on cyclin E/Cdk2 kinase activity. The precipitous drop in CDK activity as cells exit mitosis likely leads to the limited amounts of Mcm2-7 being assembled onto chromatin that we detect specifically at ORC binding sites. As the cells progress into G1, increasing levels of Cdk2 activity contribute to the second phase of cyclin E/Cdk2 dependent Mcm2-7 loading. Indeed, a gradual increase of Cdk2 activity occurs following exit from mitosis in cycling human cells (Spencer et al., 2013). Conversely, at some point increasing CDK activity leads to the inhibition of pre-RC assembly and ultimately origin activation [reviewed in (Masai et al., 2010)]. In *S. cerevisiae* CDK activity targeted towards pre-RC components directly impairs pre-RC assembly (Drury et al., 1997; Nguyen et al., 2001; Elsasser et al., 1999); however, in higher eukaryotes the evidence for a direct role is less clear. In human cell culture, phosphorylation of Cdt1 by CDK actively promotes its degradation by the SCF-Skp2 E3 ligase (Liu et al., 2004; Nishitani et al., 2006). However, this appears to be a minor pathway as the bulk of Cdt1 is degraded in S-phase by PCNA coupled Cul4-Ddb1 destruction (Arias and Walter, 2006; Jin et al., 2006; Senga et al., 2006).

How does cyclin E/Cdk2 activity promote the loading of the full Mcm2-7 complement? Prior experiments performed on quiescent mammalian cells re-entering the cell cycle suggested that cyclin E/Cdk2 kinase activity stabilizes Cdc6 and Cdc7, two factors critical for pre-RC assembly and initiation, respectively (Mailand and Diffley, 2005; Chuang et al., 2009). Although Cdc6 stabilization may, in part, play a role in promoting pre-RC assembly in our mitotically cycling *Drosophila* Kc cells, we note that there must be sufficient Cdc6 to allow a single round of Mcm2-7 loading at almost all the ORC binding sites in the *Drosophila* genome. Thus, Cdc6 would have to be specifically targeted for degradation immediately following the first round of pre-RC assembly at ORC binding sites. Instead, we propose that the Mcm2-7 complexes which are loaded proximal to the ORC binding sites in a cyclin E and Cdk2 independent manner may serve to promote additional Mcm2-7 loading upon exposure to increasing cyclin E/Cdk2 kinase activity prior to S-phase by a direct Mcm2-7 - cyclin E - Cdt1 interaction (Geng et al., 2007). The transition from a cyclin E/Cdk2 kinase independent to dependent mode of Mcm2-7 loading would facilitate a feed forward mechanism to amplify Mcm2-7 loading as G1 progresses.

The precise regulation of cyclin E/Cdk2 kinase activity during G1 is critical for pre-RC assembly and genome stability. Cyclin E/Cdk2 kinase activity is required to load the full Mcm2-7 complement, yet deregulation of cyclin E/Cdk2 kinase activity results in a shortening of G1, a lengthening of S-phase, and genomic instability (Ohtsubo and Roberts, 1993; Resnitzky et al., 1994; Wimmel et al., 1994; Spruck et al., 1999). The genomic instability and lengthening of S-phase appear to be due to a defect in pre-RC assembly (Ekholm-Reed et al., 2004). We speculate that deregulation of cyclin E/Cdk2 kinase activity would result in the cells transitioning from mitosis being immediately exposed to high levels of cyclin E/Cdk2 kinase activity instead of the gradual increase that normally occurs over the course of G1 (Spencer et al., 2013). Thus, insufficient pre-RCs would be assembled on chromatin prior to S-phase entry and may not properly redistribute throughout the genome. *In vivo*, mutations that impact Mcm2-7 levels and pre-RC assembly have been linked to genome instability and tumorigenesis (Pruitt et al., 2007; Shima et al., 2007; Chuang et al., 2010).

2.7.2 Genome-wide Redistribution of Mcm2-7

Our chromatin immunoprecipitation studies of Mcm2-7 localization provide the first genome-wide view of the Mcm2-7 distribution in a higher eukaryote, and also reveal a dramatic reorganization of Mcm2-7 during late G1. In the absence of cyclin E/Cdk2 kinase activity (early G1) we find that there is near perfect concordance between ORC and Mcm2-7 peaks. A similar distribution of overlapping ORC and Mcm2-7 peaks are also observed in *S. cerevisiae* (Wyrick et al., 2001; Eaton et al., 2010). Not surprisingly, these genome-wide datasets from *S. cerevisiae* reinforce the essential role of ORC in pre-RC assembly and suggest that Mcm2-7 loading is a local phenomena restricted to the adjacent sequences. In contrast, in *Drosophila* we see a dramatic change in the distribution of Mcm2-7 that occurs by the G1/S transition. Not only are more Mcm2-7 loaded onto the chromatin, but they are also broadly enriched along the entire chromosome and not restricted to ORC binding sites.

The distribution of the full Mcm2-7 complement along the genome is shaped by active transcription. Mcm2-7 signal is enriched in both intergenic and non-transcribed sequences, but depleted from actively transcribed genes. The displacement of Mcm2-7 is not localized to exons, but rather occurs over the entirety of the transcript suggesting that local sequence bias (eg. increased GC content in exons) is not a contributing factor. We do not envision that the full complement of Mcm2-7 is coating these intergenic and non-transcribed regions, but rather represent an increased probability of finding an Mcm2-7 complex in static regions of the genome relative to those that are actively undergoing transcription.

We propose that Mcm2-7 residing in transcribed regions are displaced by the passage of RNA Pol II during transcription. The displacement of Mcm2-7 during transcription is analogous to the removal of Mcm2-7 from inactive origins during passage of the replication fork in S-phase (Madine et al., 1995; Krude et al., 1996; Romanowski et al., 1996b). The bimodal distribution of Mcm2-7 at intergenic and non-transcribed genes relative to active genes may only be established after the transition into S-phase. Prior to S-phase and the cessation of pre-RC assembly, Mcm2-7 are likely in a cycle of loading and subsequent eviction in transcribed regions; however, once pre-RC assembly is inhibited they are no longer able to re-occupy transcribed regions following eviction by active transcription.

The mechanism(s) by which the Mcm2-7 enrichment transitions from ORC specific localization to a broad and distributed pattern along the chromosome is not clear. We envision at least two possibilities -- active translocation along the DNA away from ORC or the loading of Mcm2-7 at distal sites by chromosomal looping. *In vitro* reconstitution of Mcm2-7 loading on circular DNA templates revealed Mcm2-7 double hexamers distributed randomly on the template irrespective of the location of the ACS or ORC binding site (Evrin et al., 2009; Remus et al., 2009). Thus, *in vitro*, Mcm2-7 double hexamers can be loaded and are free to translocate along the dsDNA template.

However, it is very difficult to imagine how, *in vivo*, Mcm2-7 complexes could translocate along the DNA given the chromatin obstacles such as nucleosomes, DNA binding proteins, and active transcription. Nucleosomes would have to be displaced and re-assembled throughout the genome requiring ATP-dependent chromatin remodeling activities and histone chaperones (Alabert and Groth, 2012). Alternatively, the loading of Mcm2-7 at sites distal to ORC may be achieved by chromatin looping mediated by cohesin complexes that are close to ORC in *Drosophila* (MacAlpine et al., 2010). In mammalian cells, cohesin has been shown to be required for looping chromatin at replication factories (Guillou et al., 2010).

2.7.3 Defining an Origin of Replication

The search for sequence-based replicators, similar to the ARS element in *S. cerevisiae*, has been a holy grail for the mammalian replication field. *In vitro*, any sequence can be replicated in *Xenopus* extracts and plasmid based assays have also exhibited very promiscuous replication (Krysan and Calos, 1991). Recently, the analysis of mammalian replication intermediates by next-generation sequencing identified G4-quadruplex sequences as potential replicators (Cayrou et al., 2011; Besnard et al., 2012; Valton et al., 2014). Despite the identification of G4 quadruplex structures as origins of replication there is little concordance between datasets, indicating that these degenerate structures, which occur every few kilobases, are not sufficient for origin specification. Work from the Hamlin group using the gold standard of two-dimensional gel electrophoresis demonstrated that many genomic fragments have the potential to harbor inefficient replication origins (Mesner et al., 2006). We propose that the broad distribution of the full complement of Mcm2-7 we observe contributes to the apparent promiscuity of metazoan origin selection. *In vitro* experiments in *Xenopus* have shown that ORC is not required for initiation after pre-RC assembly (Hua and Newport, 1998; Rowles et al., 1999). Thus, those Mcm2-7 helicases located distally from ORC would still have the potential to be activated and could likely function as dormant origins in the presence of replicative stress.

In our prior studies, we have demonstrated that early origins of replication, which are resistant to HU, are enriched for ORC binding sites (MacAlpine et al., 2010). These ORC binding sites also serve as focal points for the cyclin E independent Mcm2-7 loading. Together, these data suggest that Mcm2-7 loaded proximal to ORC binding sites have an increased likelihood of activating during early S phase. In *S. cerevisiae*, Cdc45 specifically associates with early activating origins in G1 of the cell cycle (Aparicio et al., 1999). Thus, during early G1 in *Drosophila*, Cdc45 may specifically associate with the few Mcm2-7 complexes loaded proximal to ORC binding sites, increasing the likelihood of activation of this subset of pre-RCs during entry into S-phase.

2.8 Future Directions: Progression into S-phase with Minimal Chromatin Bound Mcm2-7 Levels.

The cellular amount of Mcm2-7 can be lowered significantly without any immediate phenotype or change in the rate of S-phase (Woodward et al., 2006; Crevel et al., 2007; Ge et al., 2007; Ibarra et al., 2008). However, if a cell with depleted amounts of Mcm2-7 is challenged with replicative stress, then S-phase is slowed because the number of dormant origins is reduced (Lei et al., 1996; Woodward et al., 2006; Ge et al., 2007; Ibarra et al., 2008). I wanted to assess if the two minimal amounts of chromatin bound Mcm2-7 (cyclin E independent and +Dacapo) were sufficient for S-phase. I first assessed if cells with the cyclin E independent level of chromatin bound Mcm2-7 could enter S-phase. Cells were treated as in **Figure 8**, however, cells were not treated with 1 mM HU so, they could progress into S-phase. Cells were driven through G1 phase and into S-phase by inducing FLAG-cyclin E overexpression for either 0, 2, or 4 hours. Cells with the cyclin E independent amount of chromatin bound Mcm2-7 were able to enter S-phase and had similar FACs profiles as the control **(Figure 16)**. These results indicate that the cyclin E independent amount of chromatin bound Mcm2-7 is sufficient for S-phase entry and progression. Interestingly, these cells were not sensitive to replicative stress **(data not shown)**. These results agree with another study, in *Drosophila*, that showed that cells with a 95-99% reduction of Mcm2-7 had no cell cycle phenotype (Crevel et al., 2007).

The previous experiment relied on the overexpression of FLAG-cyclin E to drive cells into S-phase. I next tested if minimal levels of chromatin bound Mcm2-7 obtained through a less perturbing mechanism (overexpression of cyclin E can cause genomic instability (Bester et al., 2011; Jones et al., 2012)) were sufficient for S-phase progression. I first overexpressed Dacapo followed by depletion of Dup/Cdt1 and Cdc6 by RNAi to prevent additional Mcm2-7 loading. Cells were then released from the Dacapo arrest, harvested at 6 hours and 18 hours after release, and assessed by FACs. Control cells were able to continue to cycle, however, cells pre-treated by Dacapo overexpression, which results in reduced Mcm2-7 chromatin levels, failed to enter S-phase and appeared to die **(Figure 17)**.

In the future, I would like to address the differences between the experimental design and possible biological relevance of these two contrasting results. The overexpression of cyclin E has been shown to drive cells into S-phase, but leads to

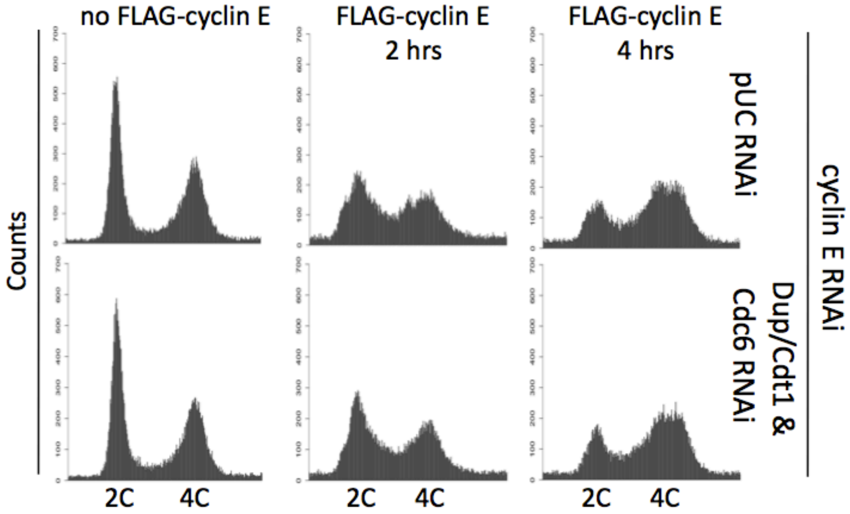

Figure 16: Cyclin E Independent Amount of Chromatin Bound Mcm2-7 are Sufficient for Entry into S-phase.

FACs profiles of KC cells treated similarly as in Figure 8A, except HU was not added to the medium upon FLAG-cyclin E overexpression. Cells stably transfected with an inducible FLAG-cyclin E were driven through the G1-S transition with overexpression of FLAG-cyclin E with the cyclin E independent amount of chromatin bound Mcm2-7.

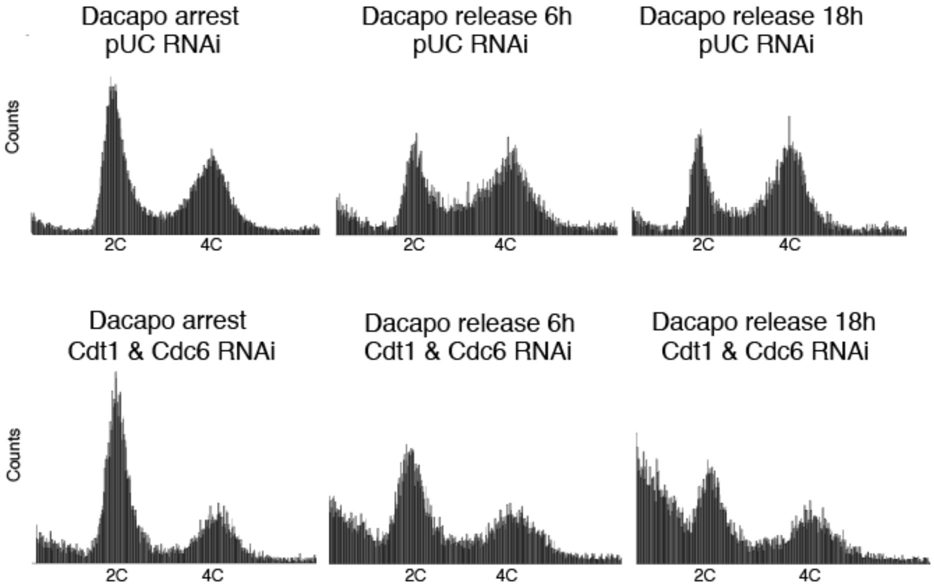

Figure 17: Dacapo Overexpression Level of Chromatin Bound Mcm2-7 are Insufficient to Support Entry into S-phase.

FACs profiles of Dacapo arrested cells released with the ability to load additional Mcm2-7 (pUC RNAi) or unable to load additional Mcm2-7 (Dup/Cdt1 and Cdc6 RNAi).

genomic instability (Bester et al., 2011). Therefore, assessing if the cyclin E independent amount of Mcm2-7 is sufficient for S-phase could be better tested over varying degrees of cyclin E overexpression. It is also possible that Dacapo has a role similar to its other mammalian homologue p21 (de Nooij et al., 1996). It has recently been shown that in M-phase, p21 modulates whether or not cells will continue to cycle (Spencer et al., 2013). If Dacapo levels are increased through M-phase it could poise cells to senesce. Thus, they would be unprepared to enter the following S-phase. I could begin to address if an increase of Dacapo in G2 and M-phase affects the subsequent S-phase by arresting cells in G2 with 3% DMSO while over expressing Dacapo, releasing cells by plating in fresh medium, and assessing the kinetics of S-phase entry. Further, the differences between the two experiments could be due to differences in Mcm2-7 phosphorylation. Mcm2-7 subunits are phosphorylated in other organisms by DDK to promote pre-IC formation (Jares and Blow, 2000; Takahashi and Walter, 2005; Yoshizawa-Sugata et al., 2005; Masai et al., 2006; Tsuji et al., 2006). However, a DDK homologue has not been described in *Drosophila* and perhaps cyclin E/Cdk2 activity phosphorylates Mcm2-7 in *Drosophila*. Together these two experiments indicate that the cyclin E RNAi and +Dacapo chromatin bound Mcm2-7 populations are possibly different. These differences can be explored to gain more understanding about the regulation of Mcm2-7 loading.

2.9 Materials and Methods

Drosophila cell culture

Cells were cultured in 150-mm plates at an approximate density of $1X10^6$ cells/mL in Schneider's Insect Cell Medium (Invitrogen) supplemented with 10% heat inactivated FBS (Hyclone) and 1% penicillin/streptomycin /glutamine (Invitrogen). Dacapo was overexpressed for 48 h in the presence of 500 μM copper sulfate. FLAG-tagged cyclin E was overexpressed in the presence of 500 μM copper sulfate in the medium. To arrest at G1/S, cells were incubated with 1 mM HU for 24 h (unless otherwise noted). G2 arrested cells were treated with 3% DMSO and assayed after 48 h.

dsRNA synthesis

Primers were designed with a 5′ T7 sequence (TTAATACGACTCACTATAGGGAGA) to amplify a 500-900 bp region in the gene of interest. The region was amplified using standard PCR followed by gel extraction (Qiagen Qiaquick gel extraction) and the PCR product was subsequently used as a template for dsRNA production (Promega T7 Ribomax express large scale expression). The dsRNA was purified by phenol/chloroform extraction followed by ethanol precipitation. Primers used for each target: Dup/Cdt1 dsRNA , F: 5′ctatcagtatcaagaacaggcg, R: 5′tgctttccaccagactg; Cyclin E dsRNA, F: 5' gccatccgtcacataagca, R:5' atcgtggaagcaagcagac; Cdk2 dsRNA, F: 5′tgtggccctcaaaaagattc, R: 5′gaagtaagcgtgctgcagtg; Cdc6 dsRNA, F:5′gccacagcacctgatcagtccttcgcg, R:5′gcagtttacaagaaactatgcacc; pUC dsRNA, F: 5′ agctcactcaaaggcggtaa, R: 5′ gcctacatacctcgctctgc.

RNAi in *Drosophila* Kc cells

Cells were washed with serum-free Schneider's insect medium (Invitrogen), resuspended to 1X10^6 cells/mL in serum free insect medium, and then replated. DsRNA (15 μg/1X10^6 cells/mL) was added to the medium and incubated for 1 h at 25°C. Medium with 2X serum was added to make the total medium 1X in respect to serum. Cyclin E and Cdk2 RNAi cells were assessed 48 h after RNAi, Dup/Cdt1 and Cdc6

RNAi cells were assessed 24 h after RNAi, and pUC RNAi cells were assessed 24 h after RNAi.

FACs preparation and analysis

Cells were harvested by centrifugation, washed with ice-cold 1X PBS, resuspended in residual 1X PBS, and fixed overnight at 4°C with ice-cold 100% ethanol. Next, the cells were centrifuged, the supernatant removed, and washed with 1X PBS- 1% fetal bovine serum. The cells were resuspended in a final volume of 800 μL 1X PBS- 1% fetal bovine serum with 50 mg/mL of propidium iodide and 0.1 mg/mL RNaseA. The cells were incubated at 37°C for 1 h. Equal numbers of cells per condition were assayed using a FACSCanto machine running BD FACSDiva software and R bioconductor flowCore package to generate histograms.

Chromatin fractionation

Fractionation of the chromatin was optimized for *Drosophila* cell lines **(Figure 18)** from established protocols (Hancock, 1974; Ritzi et al., 1998). Cells were harvested and then washed with ice cold 1X PBS twice. Cells were resuspended in hypotonic buffer (10 mM HEPES-KOH (pH 7.5), 20 mM KCl, 0.25 mM EDTA, and a Roche protease mini tab) kept on ice for 4 mins then centrifuged for 4 mins at 700g at 4°C. The pellet was resuspended in lysis buffer (10 mM HEPES-KOH (pH 7.5), 70 mM NaCl, 20 mM KCl, 5 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5% NP-40, and a Roche protease inhibitor tab) and underlayed with 1 volume of 30% sucrose. The lysate was centrifuged for 3 mins at 3000g at 4°C. The supernatant was set aside as the whole cell lysate sample. The pellet was resuspended in lysis buffer and centrifuged with a sucrose cushion. The pellet was washed with salt buffer (20 mM HEPES/KOH (pH 7.5), 0.5 mM $MgCl_2$, 0.3 M Sucrose, KCl, and Roche protease inhibitor tab) with or without a KCl gradient (0 – 480 mM) and centrifuged at 4°C for 3 mins at 3000g. The pellet was washed again in salt buffer and resuspended in 2X SDS+ßME loading buffer.

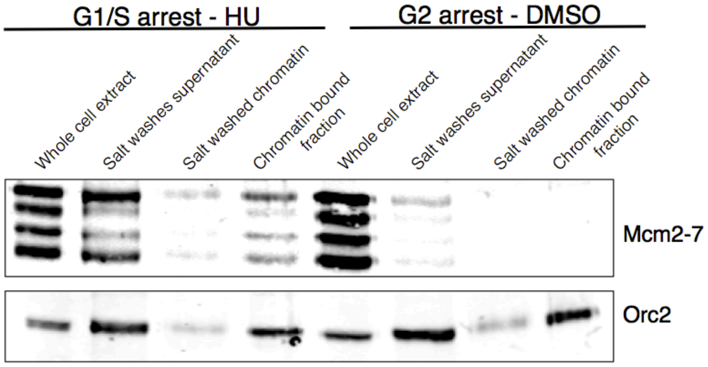
G1/S arrest - HU
G2 arrest - DMSO
Whole cell extract
Salt washes supernatant
Salt washed chromatin
Chromatin bound fraction
Whole cell extract
Salt washes supernatant
Salt washed chromatin
Chromatin bound fraction
Mcm2-7
Orc2

Figure 18: Chromatin Fractionation Optimization in *Drosophila* for Mcm2-7 and ORC.

1 mM HU and 3% DMSO treated cells were subjected to chromatin fractionation (Whole cell extract and Chromatin bound fraction) and chromatin fractionation followed by sequential salt washes of the isolated chromatin (Salt wash supernatant and Salt washed chromatin). Chromatin bound Mcm2-7 and ORC were released from the chromatin by the salt washes indicating that they are soluble. Importantly, both conditions had similar amounts of salt soluble ORC released from the chromatin bound fraction.

Western blot analysis

Mcm2-7 monoclonal mouse antibody (AS1.1) was used at 1:100 (Chen et al., 2007), Orc2 rabbit polyclonal antibody was used at 1:3,000 (Austin et al., 1999), Dacapo (NP-1) used at 1:1,000 (Iowa Hybridoma Bank), Dup/Cdt1 guinea pig raised antibody was used at 1:5,000, cyclin E antibody raised in rabbits was used at 1:500 (d-300 Santa Cruz sc-33748). Secondary antibodies: goat anti-mouse IRDye 800CW IgG (LiCor), goat anti-rabbit Alexa Fluor 680 IgG (Invitrogen), and goat anti-guinea pig IRDye 800 (Rockland immunochemicals cat# 606-132-129) were all used at 1:10,000. Westerns were visualized and quantified using LiCor infrared technology and gray scaled in ImageJ.

ChIP-chip sample preparation

Samples were prepared as described previously (MacAlpine et al., 2010). Mcm2-7 (AS1.1) antibody was used at a 1/25 dilution. All experiments were performed in duplicate.

ChIP-Chip analysis

Within-replicate probe intensities were determined and between-slide intensities normalized via the R (Team, 2012) package limma (Smyth, 2005). Replicated probe intensities were determined via MA2C (Song et al., 2007). Subsequently, all analysis was done in R using the combined replicate probe intensities. Gene bodies were obtained from (Graveley et al., 2011). All genomic data is publicly available at NCBI GEO data repository with the following accessions numbers: GSE17282, GSE17283, GSE41349, and GSEXXXXX.

FLAG-cyclin E stable transfection

SLIC cloning (Li and Elledge, 2012) was used to introduce a FLAG fusion tag in frame at either the N-terminus or C-terminus of cyclin E cDNA (clone LD22682) in the pMK-CTAP plasmid (pMK-CTAP was a gift from Artavanis lab). Briefly, the plasmid was digested with a single restriction site enzyme, N term= Xho1, C term= Spe1. The 5′ ends were treated with T4 DNA polymerase (NEB) for 45 mins at room temperature. Approximately 150 ng of vector in a 1:1 ratio was annealed to FLAG primers in 1X ligation mix (Invitrogen) at 37°C for 30 mins. The reactions were transformed into max efficiency DH5α cells (Invitrogen). Effectene transfection reagent kit (Qiagen) protocol was followed for suspension cells in 100 mm dish, with the only exception being that 150 μL Effectene reagent was added. Cells were plated at $1X10^6$ cells/mL and incubated at 25°C for 2 days. Hygromycin B (Sigma) was then added to 0.125 μg/mL.

Histone H1 kinase assay

Cells were harvested, washed with 1X PBS and resuspended in lysis buffer (Geng et al., 1999) (40 mM HEPES (pH 7.5), 120 mM NaCl, 1 mM EDTA, 10 mM pyrophosphate, 10 mM glycophosphate, 50 mM NaF, 1.5 mM Na_3VO_4, 1% Triton-X-100, and a Roche protease inhibitor tab). Cells were incubated on ice for 30 mins then centrifuged at 24,000g for 15 mins at 4°C. The supernatant was removed, saved, and stored at -80°C. The protein concentration of the extract was determined using Bio-rad protein concentration reagent (500-0006). 100 μg of protein was incubated with FLAG-M2 antibody (Sigma 3165) overnight at 4°C. Gamma g sepharose beads (GE Healthcare) were added and incubated for 4 h at 4°C. The beads were washed 3 times with wash buffer (Geng et al. 1999) (20 mM Tris-HCl (pH 8), 10% glycerol, 5 mM $MgCl_2$, 100 mM

KCl, 0.1% Tween-20, and a Roche protease inhibitor tab). The beads were equilibrated with kinase assay buffer (Geng et al. 1999) (20 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 2.5 mM $MnCl_2$, 1 mM DTT, 0.013 mM ATP, 2 μCi P^{32} gamma ATP, 2.5 μg histone H1 (NEB M2501S). The reactions were incubated for 30 mins at 37°C and stopped by addition of 2X SDS+ßME buffer. The reactions were heated to 95°C for 5 mins then loaded onto a 15% polyacrylamide gel. The gel was imaged on a phosphorimager.

3. Testing the Role of Pre-RC Assembly in Cohesin Loading

Sister chromatids are paired together by cohesion at the time of their replication. Sister Chromatid Cohesion (SCC) ensures that the sister chromatids remain together until they are separated in mitosis. Cohesion creates tension between the chromosomes and the spindles to aid in chromatid separation during mitosis (Tanaka et al., 2000). SCC is mediated by the cohesin complex, a highly conserved complex among eukaryotes, consisting of Smc1, Smc3, Scc1/SA, and Scc1/Mcd1/Rad21 (Anderson et al., 2002; Haering et al., 2002).

The two processes of replication and sister chromosome cohesion (SCC) are both critical to maintain genomic integrity. *In vitro* studies in *Xenopus* suggest that the assembly of the pre-RC in G1 is necessary for cohesin loading (Gillespie and Hirano, 2004; Takahashi et al., 2004). This is further supported by evidence from *Xenopus* and human cells that demonstrate direct interactions between pre-RC and cohesin components (Takahashi et al., 2008; Guillou et al., 2010). I wanted to characterize a potential link between pre-RC assembly and cohesin loading in *Drosophila*. To this end, cohesin and pre-RC components were mapped throughout genome and cohesin complex loading was assessed when pre-RC components were depleted by RNAi in *Drosophila* Kc cells.

3.1 Cohesin Co-localizes with Pre-RC Components Throughout the Drosophila Genome

I first assessed the localization of cohesin and pre-RC components throughout the genome by ChIP-chip. I used *Drosophila* genome wide localization data describing the binding of cohesin complex members (from the Dorsett lab) and compared them with Orc2 localization data. Consistent with the *in vitro* results from *X. laevis* (Takahashi et al., 2004) these *in vivo* results demonstrate that members of the cohesin complex (Smc1 shown here) are enriched at ORC binding sites **(Figure 19).** These data were published (MacAlpine et al., 2010). Further, these data suggest that ORC and pre-RC assembly may be involved in cohesin loading in *D. melanogaster*.

3.2 Cohesin can Load Independently of the Pre-RC in Vivo

In order to determine if pre-RC assembly is necessary for cohesin loading, I used RNAi in *Drosophila* Kc tissue culture cells to inhibit pre-RC assembly. The chromatin bound levels of cohesin and pre-RC components were compared by chromatin fractionation and western blot analysis. Four populations of cells in different phases of the cell cycle were assessed for chromatin bound cohesin; an asynchronous population, a G1 arrest by Dup/Cdt1 RNAi, a G1-S arrest induced by 1 mM HU, and a G2 arrest by 3% DMSO **(Figure 20**). In Dup/Cdt1 RNAi arrested cells and G2 arrested cells there was no detectible chromatin bound Mcm2-7, as expected. In addition, there were no detectible differences in chromatin bound cohesin between any of the conditions assayed. These results suggest that cohesin can load independently of pre-RC assembly *in vivo* in *Drosophila*.

3.3 Cohesin can Load Independently of ORC in Vivo

ORC is another possible mediator of cohesin loading as they co-localize along the *Drosophila* chromosomes. My previous data suggested that cohesin chromatin levels do not differ throughout the cell cycle. However, differential ORC loading and its effect on cohesin loading were not assessed. In budding yeast, depleting Orc2 results in loss of SCC (Shimada and Gasser, 2007). Further, *Drosophila* ORC mutants were found to have

SCC defects (Pflumm and Botchan, 2001) suggesting that ORC may directly mediate cohesin loading.

To determine the role of ORC in cohesin loading, *in vivo*, RNAi and chromatin fractionation procedures were performed for Orc1, Orc2, and Orc6 simultaneously. Cells were assessed at 48, 72, and 96 hours after RNAi treatment. Cells depleted of ORC had cell cycle defects **(Figure 21A)** consistent with cells being unable to cycle properly and in agreement with previous reports (Loupart et al., 2000; Pflumm and Botchan, 2001). Histone H4 acetylation was used as a loading control. Western blot analysis of the chromatin fractionation showed that a chromatin bound cohesin component, Smc1, was unaffected by ORC depletion **(Figure 21B)**. My results suggest that ORC is dispensable for cohesin loading. In addition, HP1 levels were assessed to ascertain if depletion of ORC subunits affected heterochromatin. It has been shown in *Drosophila* that ORC and HP1 interact and colocalize in heterochromatin (Pak et al., 1997). In human cells, depletion of ORC subunits alters HP1 localization and vice versa (Prasanth et al., 2004; Prasanth et al., 2004). Further, in mammalian cells and fission yeast HP1 and cohesin components colocalize (Nonaka et al., 2002; Koch et al., 2008). The unaltered chromatin associated levels of HP1 suggests that heterochromatin was not affected by ORC depletion. Therefore, the lack of change in cohesin chromatin bound levels may be a result of heterochromatin remaining intact when ORC is depleted.

3.4 Cohesin can Load in the Absence of Nipped-B in Vivo

I next addressed if depletion of Nipped-B, a factor shown to be required for cohesin loading in other organisms, was required for cohesin chromatin association (Ciosk et al., 2000; Takahashi et al., 2004). Cells were arrested in G2 with 3% DMSO to synchronize cells. While arrested in G2, cells were treated with Nipped-B specific or

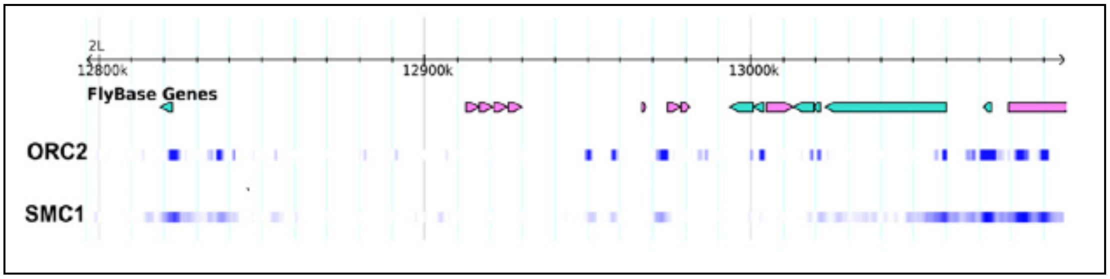

Figure 19: Cohesin and Pre-RC Components Colocalize Along the *Drosophila* Genome.

ChIP-chip analysis of pre-RC (Orc2) and cohesin (Smc1) enrichment along a region of *Drosophila* chromosome 2L.

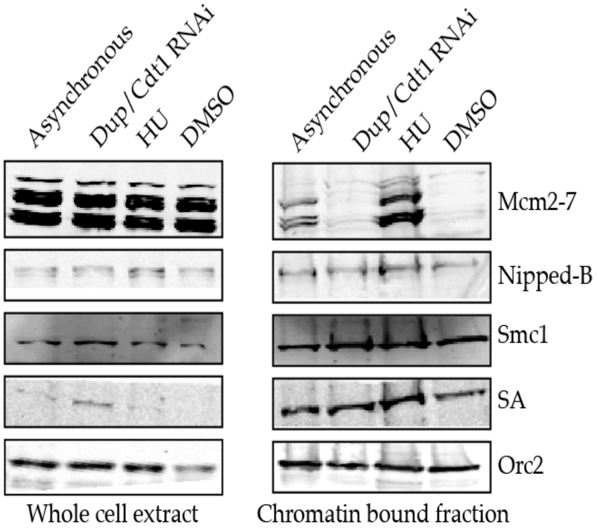
Asynchronous
Dup/Cdt1 RNAi
HU
DMSO
Asynchronous
Dup/Cdt1 RNAi
HU
DMSO
Mcm2-7
Nipped-B
Smc1
SA
Orc2
Whole cell extract
Chromatin bound fraction

Figure 20: Cohesin can Load Independently of Pre-RC Assembly *in Vivo*.

Western blot analysis of Mcm2-7, Nipped-B, Smc1, SA, and Orc2 in the whole cell extract and chromatin bound fraction from asynchronous cells, Dup/Cdt1 RNAi treated cells, 1mM HU treated cells, and 3% DMSO treated cells.

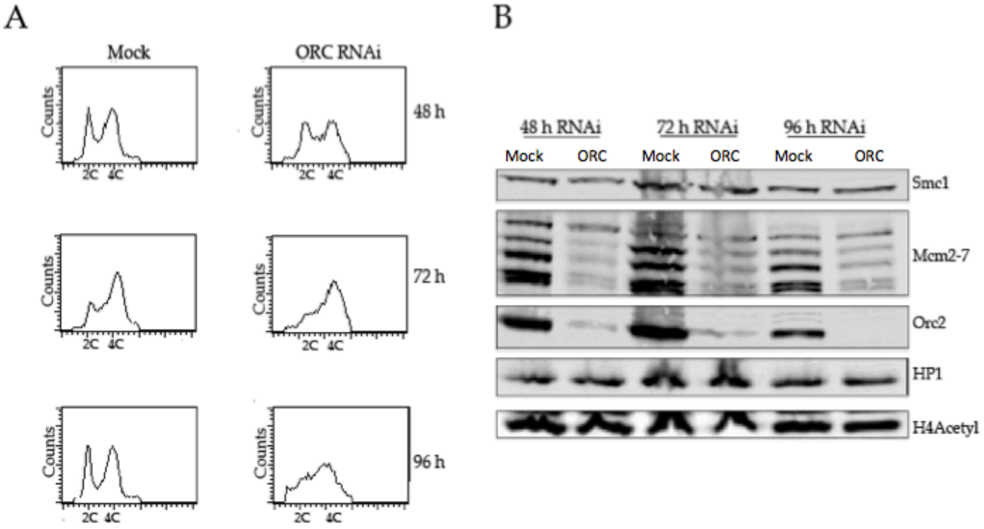

Figure 21: ORC Depletion Does Not Inhibit Cohesin Loading onto Chromatin *in Vivo*.

A. FACs profiles of Kc cells treated with ORC RNAi or mock treated for 48h, 72h, and 96h. **B.** Western blot analysis of Smc1, Mcm2-7, Orc2, HP1, and Histone H4ac in the chromatin bound fraction from cells treated with ORC RNAi or Mock treated cells for 48h, 72h, and 96h.

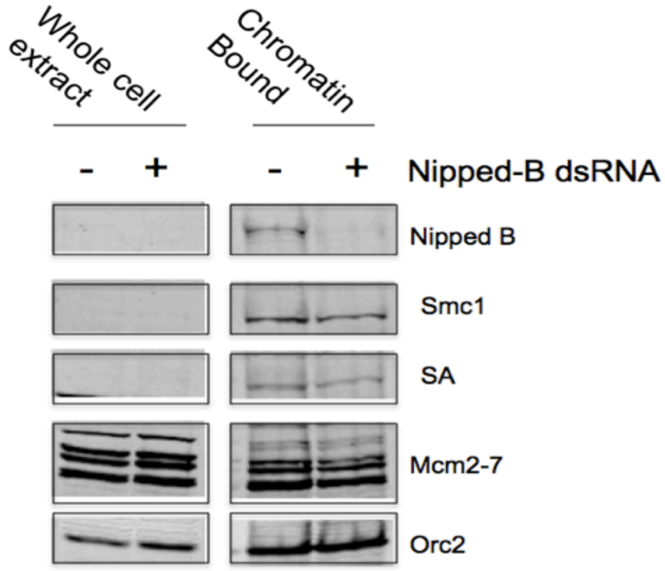
Whole cell extract
Chromatin Bound
-
+
-
+
Nipped-B dsRNA
Nipped B
Smc1
SA
Mcm2-7
Orc2

Figure 22: Cohesin Components are Chromatin Bound in the Absence of Nipped-B.

Western blot analysis of Nipped-B, Smc1, SA, Mcm2-7, and Orc2 in the whole cell extract and chromatin bound fraction of cells treated with and without Nipped-B RNAi.

control dsRNA for 24 h and then released for another 24 h into medium with dsRNA. Nipped-B levels were reduced to a level that was undetectable by western blot **(Figure 22)**. Further, there was no difference in chromatin bound cohesin levels **(Figure 22**). Depletion of Nipped-B caused no discernible cell cycle defect, indicating that there was still sufficient Nipped-B to enable sister chromatid cohesion **(data not shown)**. Overall, these results suggest that Nipped-B was not reduced to a low enough level to cause a phenotype.

3.5 No in Vitro Interaction was Detected between Pre-RC Components and Cohesin

An interaction between pre-RC components and cohesin subunits has been described in the literature in both humans and *Xenopus* (Takahashi et al., 2008; Terret et al., 2009). To establish if pre-RC components interact with cohesin complex components co-immunoprecipitations in 0-2 hour *Drosophila* embryonic extracts were performed. When extracts were incubated with Mcm2-7 antibodies conjugated to beads only Mcm2-7 was pulled down **(Figure 22)**. It is possible that the conditions for an interaction between Nipped-B and Mcm2-7 were not met in this experiment and thus, caution should be used when interpreting this negative result.

3.6 Summary

There is evidence in other organisms that shows a clear connection between DNA replication and sister chromatid cohesion through their conserved protein complexes (Takahashi et al., 2004). I found that pre-RC components and cohesin components co-localize along the *Drosophila* genome. In this chapter, I sought to address if pre-RC components are required to mediate cohesin loading in *Drosophila*. I found no indication that ORC or Mcm2-7 are required for cohesin loading *in vivo* or that they interact *in vitro*.

The failure to see a connection or interaction, between the pre-RC and cohesin, is not necessarily an indication that a link does not exist in *Drosophila*. Cohesin imparts other functions aside from SCC. Cohesin has been described as having at least five functions: sister chromatid cohesion, DNA damage response, heterochromatin organization, replication foci organization, and transcription regulation [reviewed in (Rudra and Skibbens, 2013)]. Each of the functional pools of cohesin has differences in their regulatory mechanisms. Further, the amount of cohesin imparting each function is not necessarily equal, therefore, making it difficult to assess one population.

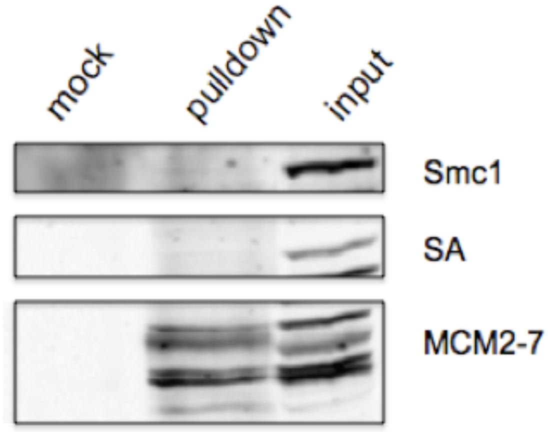

Figure 23: No Stable Interaction Between Mcm2-7 and Cohesin Subunits was Detected.

Western blot analysis for Smc1, SA, and Mcm2-7 in mock, pulldown, and input fractions from a co-immunoprecipitation using Mcm2-7 antibodies.

3.7 Materials and Methods

Western Blotting

Antibodies for cohesin subunits were a gift from D. Dorsett and were used at 1:5000 dilution.

dsRNA synthesis

Synthesized as previously described in chapter 2.9 with the target specific primers:

Nipped-B dsRNA

Forward: 5′TTAATACGACTCACTATAGGGAGAtatttgttccggtgggatgt

Reverse: 5′TTAATACGACTCACTATAGGGAGAgaaccaactcaagcacacca

SA dsRNA

Forward: 5′TTAATACGACTCACTATAGGGAGAtggctgacaaggattgtgaa

Reverse: 5′TTAATACGACTCACTATAGGGAGAtgatgttggatctggcgata

Smc1 dsRNA

Forward: 5′TTAATACGACTCACTATAGGGAGAccctttacatcaaggccaaa

Reverse: 5′TTAATACGACTCACTATAGGGAGAccatgaacttgcccaaaact

ORC dsRNA

To render the ORC complex non-functional 3 subunits were targeted. BACS were used as a template for PCR.

Orc1F: 5′TTAATACGACTCACTATAGGGAGAgatcctgcacatgtacgagc

Orc1R: 5′TTAATACGACTCACTATAGGGAGAcggcagatagttggcattga

Orc6F: 5′TTAATACGACTCACTATAGGGAGAcgaaaatgggcctaagggag

Orc6R: 5′TTAATACGACTCACTATAGGGAGAggggcttatgtgcctttttgc

Orc2R: 5’TTAATACGACTCACTATAGGGAGActggcatgcgagtgaaagta

Orc2F: 5’TTAATACGACTCACTATAGGGAGAgtccaagtcccaccaagaaa

Orc2F2:5’TTAATACGACTCACTATAGGGAGAggaggagtatcgatcgtact

Orc2R2:5’TTAATACGACTCACTATAGGGAGAggtatgtggcattgcccttg

Chromatin Fractionation

As described previously in chapter 2.9.

Drosophila cell culture

As described previously in chapter 2.9.

Drosophila cell culture RNAi

As described previously in chapter 2.9. ORC RNAi was performed for 48, 72, and 96 h. Nipped-B RNAi was performed during a G2 arrest for 24 h and repeated while cells were released for 24 h.

Mcm2-7 co-immunoprecipitation

Kc167 cells were washed with ice cold 1X PBS and resuspended in lysis buffer (50 mM Tris-HCl (pH 7.5), 125 mM NaCl, 5% glycerol, 0.2% NP-40, 1.5 mM $MgCl_2$, 1 mM dithiothreitol, 25 mM NaF, 1 mM Na_3VO_4, 1 mM EDTA and complete mini protease inhibitor tab (Roche)) and incubated on ice for 30 mins. The lysate was spun at 20,000g for 20 mins at 4ºC and the supernant was stored at -80ºC before use. Mcm2-7 antibody was added to lysate and incubated at 4ºC while rotating overnight. Gammabind G beads (Ge Life Sciences) were added and incubated for 6 h at 4ºC while rotating. The beads were washed with 1X PBS+0/1%NP-40 3 times. The inputs were saved and beads resuspended in 2X SDS+BME buffer, ran on polyacrylamide gel and subjected to western blot analysis.

FACs preparation

As described previously in chapter 2.9.

4. Characterization of Aneuploidy in *Drosophila* Cell Lines

Aneuploidy has been described as being detrimental to an organism. Small amounts of aneuploidy can have drastic phenotypes as seen in several human diseases presumably caused by differences in gene expression from copy number variation (CNV) (Torres et al., 2008; Dierssen et al., 2009). In these CNV diseases, cells are less proliferative than wild type. Contrastingly, aneuploidy is associated with cancer which is highly proliferative [reviewed in (Holland and Cleveland 2009)]. However, organisms have evolved mechanisms to adjust for differences in gene dosage from the sex chromosomes between the sexes. Together these suggest that a cell can tolerate and compensate for CNV.

To begin to better understand the mechanisms by which cells tolerate CNV, we initiated a study to examine the ploidy differences in three common *Drosophila* cell lines being characterized by the modENCODE consortium. The modENCODE consortium has generated a wealth of information on these cells lines from replication timing, transcription levels, and chromatin modifications. Therefore, it is important to understand the CNV in these cells to make biologically relevant conclusions from all the data collected.

4.1 Drosophila S2 Cells have an Aberrant Karyotype

The S2 cell line is a widely used cell line in *Drosophila* tissue culture research. In a typical diploid *Drosophila* male cell in metaphase, there are a total of 8 chromosome pairs; two chromosome 2, two chromosome 3, two chromosome 4, one chromosome X, and one Y chromosome. A cursory look at the morphology of the chromosomes in a single S2 cell metaphase spread suggests that each chromosome has two additional copies for a total of 16 chromosome pairs **(Figure 24)**.

4.2 Each Cell Line has a Unique Gross Copy Number Variation Along All of the Chromosomes

We wanted to further characterize the genomic content of S2 cells and two other widely used *Drosophila* cell lines, Kc and Bg3. To this end, we used these *Drosophila* cell lines which were characterized extensively for the modENCODE project. The modENCODE project was an NHGRI funded consortium to systematically and comprehensively identify all functional DNA elements in the *Drosophila* and *C. elegans* genome (Roy et al., 2010). My CNV characterization is an important piece of information to understand how CNV impacts the specific biological processes associated with each cell line.

To begin to characterize the cells and to determine the CNV in these cell lines we used whole genome Comparative Genomic Hybridization (CGH). We compared genomic DNA from the cell lines to that of W1118 *Drosophila* embryos. We found that the three cells lines Bg3, S2, and Kc had gross CNV throughout their chromosomes as demonstrated by the X chromosome **(Figure 25)**. There are instances of both decreases and increases of genomic content involving large regions and single loci. Overall, the cells are nearly tetraploid. Importantly, each cell line has its own signature of copy number variation suggesting that there is not a specific cell culture adaption.

4.3 Determination of CNV in Three Drosophila Cell Lines

We next wanted to determine the exact copy number for the entire genome of the *Drosophila* cell lines. To this end, we used next generation sequencing (NGS) to generate millions of 30 base pair sequence reads of each cell line’s genome. It can be assumed that

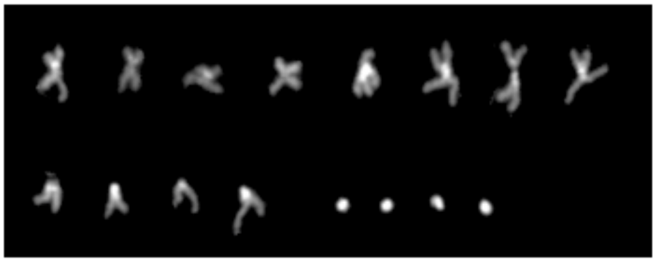

Figure 24: Karyotype of a *Drosophila* S2 Cell.

Microscope image of metaphase chromosomes from a representative S2 cell stained with DAPI arranged by metacentric, acrocentric, and dot chromosomes.

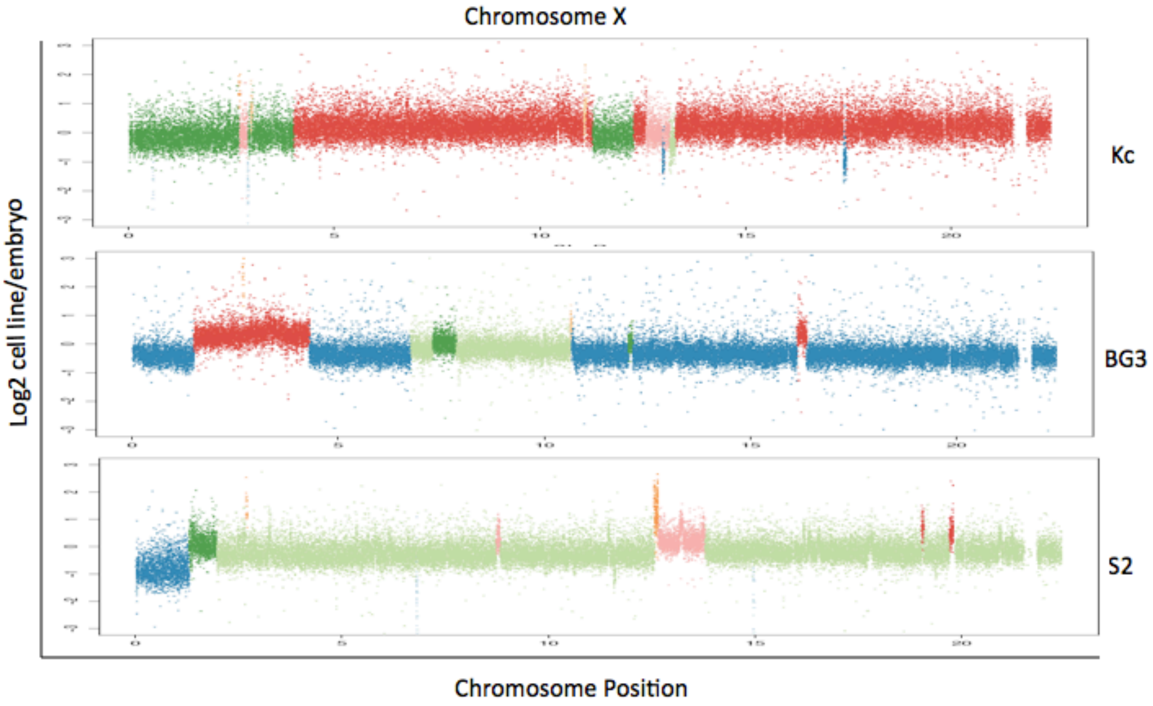

Figure 25: Each Cell Line has a Unique Pattern of CNV.

CGH analysis of total genomic DNA from Kc, BG3, and S2 cell lines compared to W1118 embryo genomic DNA along a portion of the *Drosophila* X chromosome. Chromosomal sections with the same copy number within each cell line are color-coded.

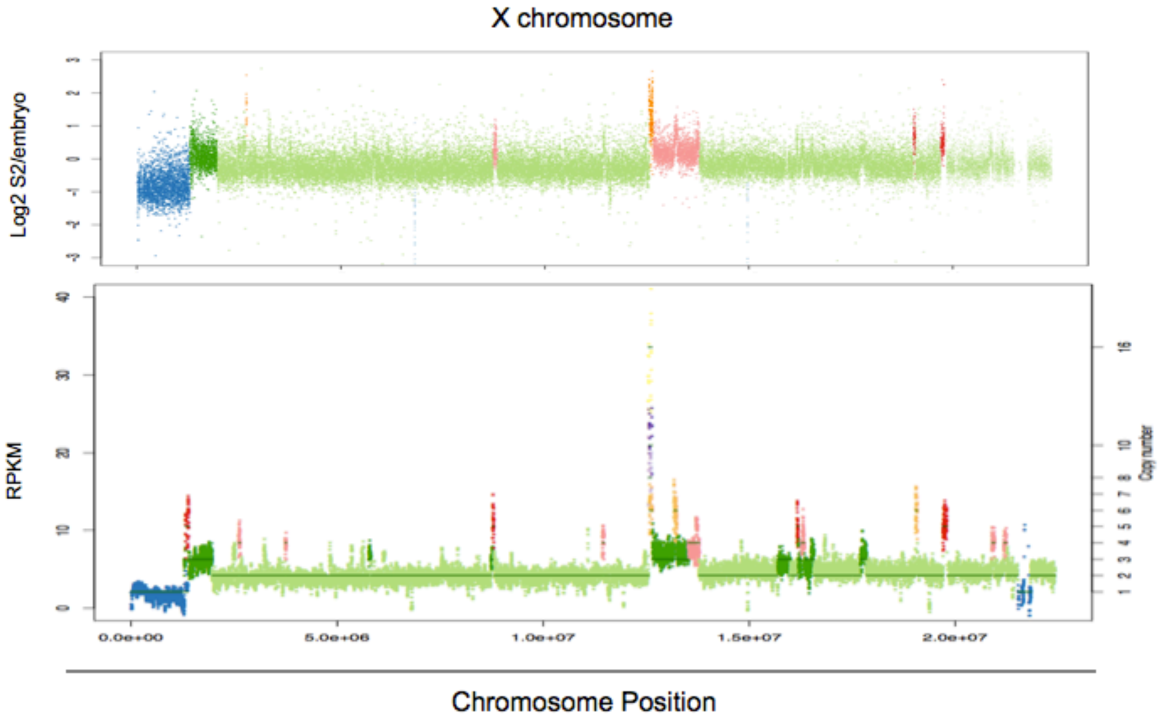

Figure 26: CGH and NGS CNV Analysis Comparison.

CGH analysis of genomic DNA from S2 cells compared to W1118 embryo genomic DNA along a portion of the *Drosophila* X chromosome with the chromosome sections with equal CNV color-coded (top panel). NGS data of S2 genomic DNA reads aligned to the same portion of the X chromosome as above with the chromosome sections with equal CNV color-coded (bottom panel).

the number of reads for a gene is correlated with the copy number of that gene. Conversely, regions with fewer reads have a lower copy number. This technique also provided higher resolution, at the base pair level, to determine the boundary points of CNV along chromosomes. The data obtained from next generation sequencing showed the same pattern of chromosomal segments differing in copy number as with CGH. The CGH **(Figure 26, top panel)** and next generation sequencing **(Figure 26, bottom panel)** replicates for a total of four independent experiments showed the same CNV patterns suggesting, at least in the short term, that these cells have a stable karyotype. Further, a different isolate of S2 cells was compared and both isolates had the same pattern further indicating that the karyotype is stable.

4.4 Copy Number does not Correlate 1-to-1 with Expression

We next reasoned that transcript levels would be most affected by CNV. We compared the transcript levels versus copy number differences using RNA-seq data generated from Dr. Sue Celniker's group as part of the modENCODE consortium **(Figure 27)**. The copy number differences versus transcript levels between two cell lines were plotted. As expected, CNV was correlated with transcript levels, however, there was not a 1-to-1 correlation **(Figure 27; black dashed line)**. Instead, the relationship between copy number and transcript levels was slightly less than 1:1 **(Figure 27; red line)**. We concluded the regions of altered copy number were buffered because the correlation was lower than a 1-to-1 relationship. These results suggest that cells can adjust for gene dosage by buffering in response to CNV. These results were published in collaboration with Dr. Brian Oliver (Zhang et al., 2010). Additional analysis is accepted for publication in Genome Biology as part of ModENCODE.

4.5 Drosophila S2 Cells have Rearranged Chromosomes

We previously showed that S2 cells are tetraploid with regions of chromosomes having drastic differences in copy number compared to neighboring chromosome regions. Therefore, we next wanted to assess how the CNV is arranged on the chromosomes. We first looked at a specific chromosome histone mark that is highly enriched on the male X chromosome, H4K16Ac. We found only three labeled acrocentric chromosomes instead of the expected four acrocentric X chromosomes in this nearly tetraploid cell line **(Figure 28)**. This result suggested that the chromosomes are rearranged. We next wanted to describe the chromosomal rearrangements throughout the genome. In collaboration with Dr. Elliot Margulies, we did a *de novo* assembly of the S2 cell line genome. Previous genome assemblies were applied to a fly backbone were not designed to identify chromosome rearrangements. This *de novo* method allowed us better understand how the chromosomes are rearranged and we can identify break points. Several rearrangements were found between the chromosomes **(Figure 29)**.

A few of the chromosomal rearrangements identified by the *de novo* assembly were verified by fluorescent *in situ* hybridization (FISH). FISH probes were selected for two neighboring genomic regions (one rearranged and the other not) and another one probe for a region close to the identified rearrangement. I tried to further characterize an identified translocation of a part of chromosome 3 to X **(Figure 30A)**. I was able to verify that the chromosomes were rearranged as indicated by non-adjacent labeling, however, the results were not as clear-cut as expected **(Figure 30B)**. Therefore, I was unable to actually reconstruct each of the four chromosome copies.

This result sheds light on how massively rearranged the chromosomes are in these cells. In theory, there are several possible combinations of breaks on a single

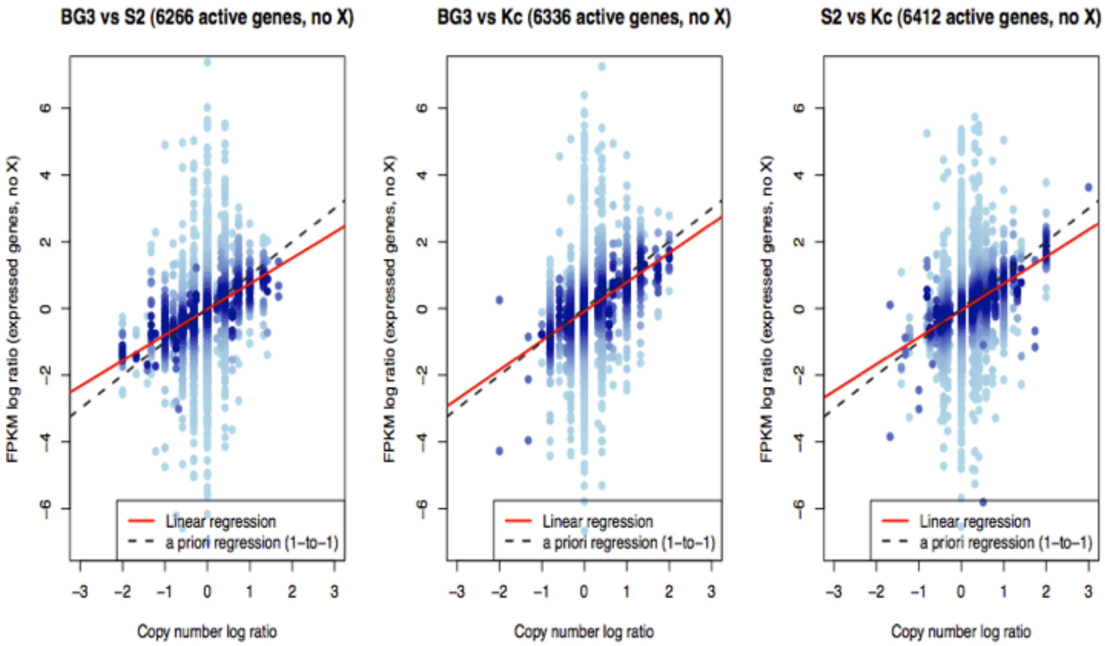

Figure 27: Copy Number and Transcript Levels Comparison between Two *Drosophila* Cell Lines is not a 1-to-1 Relationship.

The copy number log ratio between two cell lines was plotted against the transcript levels. A best fit line was found for each comparison (red line) to be slightly lower than a 1-to-1 relationship line (black dotted).

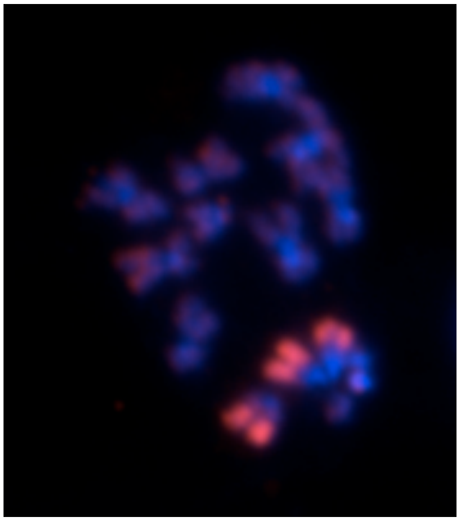

Figure 28: S2 Cells have Rearranged Chromosomes.

Representative immunofluorescent microscopy image of an S2 cell's metaphase chromosomes stained with DAPI (blue) and with an X chromosome specific histone mark, H4K16Ac (red).

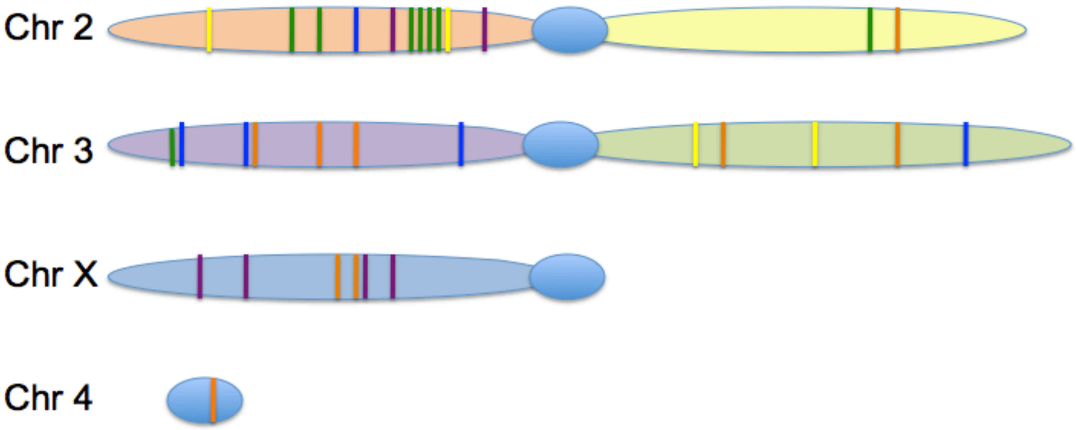

Figure 29: S2 Cells have Multiple Chromosome Break Points.

Diagram of the approximate locations of mapped chromosome break points from the *de novo* assembly of the *Drosophila* S2 cell line genome. Each chromosome arm is differentially colored to indicate the parental chromosome arm of each of the mapped break points.

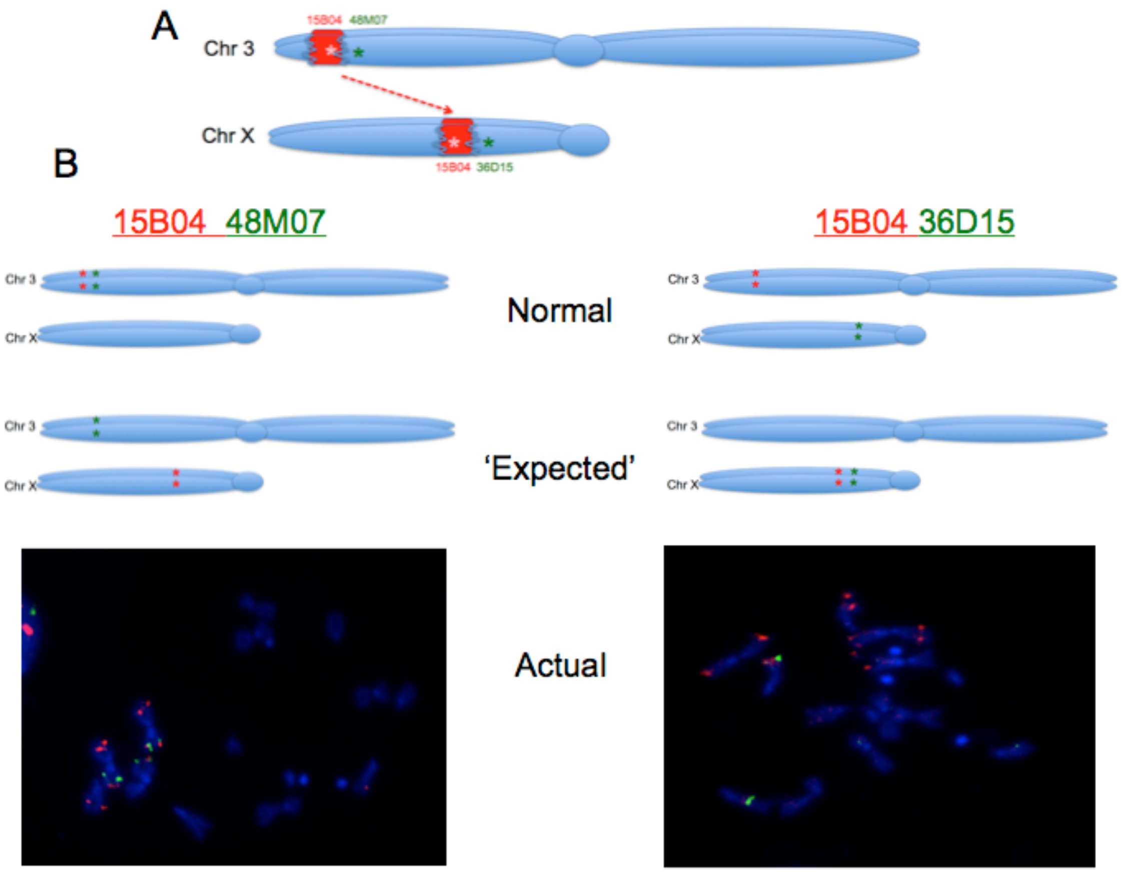
A
Chr 3
15B04
48M07
Chr X
15B04
36D15
B
15B04 48M07
15B04 36D15
Chr 3
Chr X
Normal
'Expected'
Actual

Figure 30: *Drosophila* S2 Cells have Complex Chromosome Rearrangements.

A. Diagram of a translocation indicated between chromosome 3 and the X chromosome by the *de novo* assembly of the S2 genome. The BAC probes used to identify specific chromosome regions are labeled in the color of the fluorophore used for FISH. **B.** Diagrams of the normal and the 'expected' chromosome labeling pattern for each set of BAC probes and the actual images of metaphase spreads labeled with both sets of BAC probes.

chromosome arm**.** My results suggest that each homologous chromosome is uniquely rearranged.

4.6 Summary

In summary, three aneuploid *Drosophila* cell lines were characterized. This characterization is critical for drawing biological conclusions from the data generated from these cell lines by the modENCODE project. Further, these results suggested that *Drosophila* cells may compensate for gene dosage differences in autosomal genes by buffering. In addition, I attempted to understand the pattern of chromosomal rearrangements. However, this line of research was abandoned due to the complexity of the rearrangements.

4.7 Materials and Methods

Drosophila cell culture

S2 cells were cultured in 150-mm plates at an approximate density of 1E6 cells/mL in Schneider's Insect Cell Medium (In*vitro*gen cat.# 11720-067 or Sigma cat.# S0146) supplemented with 10% heat inactivated FBS (Stratagene) and 1% penicillin/ streptomycin/ glutamine (In*vitro*gen cat.# 10378016). Bg3 cells were cultured in 150-mm plates at an approximate density of 1E6 cells/mL in Shields and Sang M3 media (Sigma cat.# S3652) was supplemented with 10μg/mL insulin (Sigma cat.# I9278). Cells were maintained at 25ºC.

Genomic DNA preparation from cell lines

Cells were harvested and resuspended in 10mM Tris pH 9.5. Pre-warmed NDS (0.5 M EDTA, 10 mM Tris (pH 9.5), 1% SDS) was added and the tube inverted. 2 mg Proteinase K was added and samples were incubated at 37ºC for 2 h. The DNA was purified using phenol/chloroform followed by ethanol precipitation. The RNA was then degraded with RNase A and incubated at 37ºC for 1.5 h followed by ethanol precipitation. The DNA was sonicated to give an average DNA fragment size of 1kb. The DNA was then ethanol precipitated.

Labeling of DNA with fluorescent nucleotides for CGH

600ng of DNA was labeled with either fluorescent Cy5- or Cy3-conjugated dUTP (Perkin Elmer), using Sequenase (US Biochemicals) and a random nonamer oligo (IDT). The DNA was dried in a speed-vac and resuspended in 10 µL of Sequenase primer mix (1× Sequenase buffer, 5 µg of random nonamer, 5 mM dUTP-Cy3 or 5). The samples were heat-denatured and cooled to 4°C in a thermocycler before adding 5 µL of Sequenase reaction mix (1× Sequenase buffer, 1.5 mM dATP, 1.5 mM dCTP, 1.5 mM dGTP, 0.75 mM dTTP, 500 ng of BSA, 3.5 mM DTT, and 13 U of Sequenase). The reaction temperature was slowly ramped to 37°C and incubated for 30 min. Following incubation, the sample was heat-denatured and cooled to 4°C, and fresh Sequenase (4 U) was added for the second and final round of labeling. Following the labeling, unincorporated nucleotides, oligo, and dye were removed using Microcon filters (Millipore).

CGH: Microarray hybridization

Labeled DNA was prepared per Agilent protocol. Briefly, the sample was mixed with a final 1x blocking solution and 1X hybridization buffer and heated to 95°C for 5 mins. The mixture was added to a *Drosophila* whole genome array slide and sealed with a gasket slide. The slide was then rotated at 65°C overnight. The slide was sequentially washed with Agilent Buffer 1 for 5 mins with stirring, Agilent Buffer 2 for 5 mins with stirring, acetonitrile for 30 secs with stirring, and then washed with Agilent Buffer 3 for30 seconds with stirring. Slides were then scanned with an Agilent reader. Subsequent analysis of data was done in R.

Whole genome sequencing

To sequence the cell lines Illumina next generation sequencing was used. The 'preparing samples for sequencing genomic DNA' protocol was followed using reagents from other vendors. Genomic DNA was prepared and sonicated to an average size of 1kb. 100 ng of DNA was used to make a library. The following substitution were made using enzymes from New England Biolabs: T4 DNA ligase buffer with 10 mM ATP, T4 DNA polymerase, Klenow DNA polymerase, T4 PNK, Klenow buffer replaced by NEB Buffer 2, Klenow exo, DNA Ligase buffer replaced by Quick ligase buffer, and DNA ligase replaced by Quick ligase. Also an addition gel extraction step was performed after 'enrich the adapter-modified DNA fragments by PCR' to size select 200-300bps products. The libraries were sequenced at the IGSP DNA sequencing facility.

Metaphase chromosome preparation (per Dr. Beth Sullivan protocol)

Cells were arrested with 0.5 μg/mL of colcemid (Sigma) by gently rocking the plate at room temperature for 2 h. Cells were harvested by centrifugation at 1200 rpm for 5 mins. Cells were resuspended in 0.8% sodium citrate and incubated at room temperature for 5 mins. Cells were then centrifuged at 900 rpm for 5 mins. Cells were then fixed with a 3:1 methanol to acetic acid solution by adding 5 mL drop wise. Cells were then centrifuged at 900rpm for 5 min and the supernatant was removed. The cells were then fixed and centrifuged for at least 6 more times. The cells were then resuspended in the remainder of supernatant and dropped onto a slide on a cold hard surface. Slides were aged at room temperature for at least 1 day.

BAC Preparation

BACs were obtained from CHORI. They were purified using the PrepEase® BAC Purification Kit (US Biochemicals)

Labeling BACS for FISH (per Dr. Beth Sullivan protocol)

100 ng of purified BAC DNA was labeled with either fluorescent Cy3-conjugated dUTP (Perkin Elmer) or ChromaTide® Alexa Fluor® 488-5-dUTP (In*vitro*gen), using Sequenase (US Biochemicals) and a random nonamer oligos (IDT). The DNA was dried in a speed-vac and resuspended in 10 μL of Sequenase primer mix (1× Sequenase buffer, 5 μg of random nonamer, 5 mM dUTP- Cy3 or 488). The samples were heat-denatured and cooled to 4°C in a thermocycler before adding 5 μL of Sequenase reaction mix (1× Sequenase buffer, 1.5 mM dATP, 1.5 mM dCTP, 1.5 mM dGTP, 0.75 mM dTTP, 500 ng of BSA, 3.5 mM DTT, and 13 U of Sequenase). The reaction temperature was slowly ramped to 37°C and incubated for 30 min. Following incubation, the sample was heat-denatured and cooled to 4°C at which time fresh Sequenase (4 U) was added for another round of labeling. Following the labeling, unincorporated nucleotides, oligo, and dye were removed using Microcon filters (Millipore).

FISH probe preparation

The purified labeled BAC probe was then ethanol precipitated, with 50ug ssDNA for FISH (Sigma), 3 M NaOAc and 3 μg of Cot1 DNA (Roche) and 2.5 V 100% Ethanol. The precipitated probe was resuspended in 10μl of hybridization mix (2X SSC, 50% Formamide, 10% dextran sulfate, 1% Tween-20). The probe was either used immediately or stored at -20°C. Immediately before use, the probe was heated to 70-75°C for 10 mins then incubated on ice for 1 min.

Metaphase chromosome slide preparation for FISH (per Dr. Beth Sullivan protocol)

Slides with dropped metaphase spreads were aged at least 1 day. Slides were incubated at 37°C in 2X SCC and 100μg/mL RNaseA for 1 h. Slides were dehydrated in an ethanol series (70%, 90% and 100%) for 2 mins on ice. Slides were air-dried followed by denaturation by incubating in 2X SSC/70% Formamide at 75°C for 1 min. Slides were then dehydrated in the same ethanol series. Slides were dried on a 50°C heat block. The probe was then added. The slide was sealed with a cover slip and rubber cement. The slide was incubated in a humidified chamber overnight at 37°C. Slides were washed 3 times in 2X SSC/50%Formamide at 37°C for 2 mins, washed 3 times 2X SSC for 2 mins at 37°C, and then washed twice with 0.1X SSC for 1 min at 37°C. Slides were rinse in PBST before staining with DAPI (Vector Laboratories), coverslipped, and sealed with clear nail polish. Slides were viewed using an Axio Imager microscope with Metamorph software.

5. Conclusions

Over the last 50 years, our understanding of the regulation of DNA replication initiators has expanded. However, the simple replicator-initiator model first used to describe DNA replication still applies. Our understanding has been expanded to include more initiators with complex regulation. This work has added to the knowledge of the regulation of initiators of DNA replication. I showed that there is regulated loading of the full complement of Mcm2-7 mediated by cyclin E/Cdk2 kinase activity during G1. In addition during the loading of the full complement of Mcm2-7 their localization is drastically changed from ORC proximal to broadly distributed throughout the genome by the G1-S transition. This result is consistent with the literature showing that an 'excess' amount of Mcm2-7 are loaded onto chromatin and that they act as dormant origins during replication fork stalling induced by replicative stress (Woodward et al., 2006; Ge et al., 2007). Together these results suggest that cyclin E/Cdk2 activity acts to establish distributed dormant origins throughout the genome to prevent genomic instability. However, genomic instability does still occur through a number of potential mechanisms: replication fork stalling, defects in sister chromatid cohesion, or rereplication, to name a few. I explored a potential outcome of sister chromatid cohesion defects, aneuploidy. I showed that cells may compensate for CNV resulting from aneuploidy by buffering (Zhang et al., 2010). Aneuploidy is common in solid tumors and whether it is tumorigenic or a survival mechanism by the cell is controversial [reviewed in (Holland and Cleveland, 2009)]. My results indicate that it is perhaps a survival mechanism because the cell may tolerate it.

In chapter 2, I showed that the loading of the full complement of Mcm2-7 occurs in two phases. The full complement of Mcm2-7 has been shown to be important for activating dormant origins during replicative stress. I found that both phases require the canonical Dup/Cdt1 and Cdc6 dependent pre-RC assembly pathway. However, I found the two phases differ in their requirement for cyclin E/Cdk2 kinase activity. The first phase of Mcm2-7 loading is cyclin E/Cdk2 kinase activity independent. The second phase is cyclin E/Cdk2 kinase activity dependent. I showed this differential regulation by cyclin E/Cdk2 kinase activity by depleting cyclin E and Cdk2 by RNAi and inhibiting their activity by Dacapo overexpression. In addition, the genome-wide distribution of Mcm2-7 changed. After the cyclin E/Cdk2 independent phase, Mcm2-7 was localized to ORC binding sites. However, after the second phase, Mcm2-7 was localized more broadly and absent from actively transcribed genes. Together, the full complement of Mcm2-7 distribution is dependent on cyclin E/Cdk2 kinase activity and shaped by transcription. This is the first time in mitotically cycling cells that cyclin E/Cdk2 kinase activity has been implicated in promoting Mcm2-7 loading. In addition, the stepwise regulation of Mcm2-7 loading was visualized for the first time.

In chapter 3, I explored the coordination of DNA replication and sister chromatid cohesion. However, no direct link was found beyond co-localization of components throughout the genome. I depleted pre-RC components by RNAi and assessed if cohesin loading was impaired though I was never able to detect a change in cohesin loading this negative result is not necessarily an indication that the two complexes have no direct link.

In chapter 4, I explored how eukaryotic cells are able to compensate for CNV resulting from aneuploidy. I characterized the CNV of three *Drosophila* cells lines that are part of the modENCODE project. The CNV data is an important piece of information in the overall description of these cells. In collaboration, with Brian Oliver, we found that, in addition to mechanisms to compensate sex chromosome gene dosage differences, cells can buffer autosomal CNV. Further analysis of one of the cell lines by *de novo* genome assembly and FISH showed that the chromosomes are highly rearranged.

This work also provides several lines of future research to understand how genomic instability may be prevented and tolerated.

1. How does cyclin E/ Cdk2 kinase activity promote the additional rounds of Mcm2-7 loading? There is evidence from the literature that suggests cyclin E/ Cdk2 kinase activity may stabilize Cdc6 transcripts, Cdc7 transcripts, or interact with the Cdt1-Mcm2-7 complex (Mailand and Diffley, 2005; Chuang et al., 2009). To test if cyclin E/Cdk2 kinase activity stabilizes Cdc6 transcripts, Cdc6 mRNA levels could be quantified by RT-PCR or qPCR in *Drosophila* Kc cells arrested by Dacapo overexpression compared to cells released from a Dacapo overexpression arrest. If cyclin E/Cdk2 kinase activity stabilizes Cdc6 transcripts, then one would expect low levels of Cdc6 transcripts in Dacapo overexpression cells. Subsequently, an increase in Cdc6 transcript levels would be seen when cells are released from Dacapo overexpression. *Drosophila* Kc cells stably transfected with a copper inducible FLAG-cyclin E can be used to test for a possible interaction between cyclin E and Cdt1-Mcm2-7. Overexpression of FLAG-cyclin E can be induced followed by a FLAG IP. If there is an interaction between cyclin E, Mcm2-7, and Dup/Cdt1, then Mcm2-7 and Dup/Cdt1 should be detected in the FLAG IP fraction by western blot.

2. How is it determined that ‘enough’ Mcm2-7 has been loaded or is there too much Mcm2-7? A recent paper indicates that cyclin E/Cdk2 kinase activity gradually increases through G1 (Spencer et al., 2013). These results suggest that there may be

regulated thresholds of cyclin E/Cdk2 kinase activity that modulate Mcm2-7 loading. This possibility could be tested in *Drosophila* Kc cells stably transfected with a copper inducible FLAG-cyclin E. The level of cyclin E/Cdk2 kinase activity could be modulated by the amount of inducer, copper, present in the medium. The level of chromatin bound Mcm2-7 can be compared to the levels of cyclin E/Cdk2 kinase activity assayed by a kinase assay.

3. How does Mcm2-7 become distally located from ORC? While there is evidence that Mcm2-7 can translocate along dsDNA (Evrin et al., 2009; Remus et al., 2009) it seems unlikely on chromatin given the amount of histones and DNA binding proteins. The removal of these to allow Mcm2-7 translocation would require a significant amount of energy. Another potential mechanism is through chromosome loops that are mediated by cohesin (presented in this work) that bring distal chromosome domains to ORC for Mcm2-7 loading. The possibility of cohesin mediating distal Mcm2-7 loading could be tested using conditions outlined in this work. Kc cells stably transfected with an inducible FLAG-cyclin E can be treated with cyclin E RNAi and subsequently treated with cohesin RNAi while arrested by cyclin E RNAi. If cohesin mediates distal Mcm2-7 loading, upon induction of FLAG-cyclin E there would be minimal additional loading in cells depleted of cohesin. Additionally, localization of Mcm2-7 can be analyzed by ChIP-chip

4. Is pre-RC assembly required for cohesin loading? I tried to address this question in chapter 3. However, I was possibly limited by differentially regulated and different pools of cohesin that function in DNA damage, heterochromatin organization, replication foci organization, transcriptional regulation, and SCC. Each of these pools could be in higher amounts than the SCC pool and regulated in cell cycle independent

manner. This could be addressed by creating more specific antibodies that only recognize a specific pool of cohesin, perhaps, by unique modifications. While Smc3 acetylation by Eco1-established cohesion [reviewed in (Peters and Nishiyama, 2012)], studies in *S. cerevisiae* found that acetylation of another residue of Smc3 promotes condensation of the chromosomes (Guacci and Koshland, 2012). Together with other possible modifications identified this suggests that a 'cohesin code' exists where different modifications promote the activity of a specific cellular process mediated by the cohesin complex [reviewed in (Rudra and Skibbens 2013)]. The ability to differentiate between the different pools of cohesin would greatly increase our knowledge of how this complex functions in diverse cellular processes.

The work presented here lays the foundation for a potential new cancer treatment angle. The cell cycle and DNA replication are deregulated in cancer. Mice homozygous for an $Mcm4^{choas3}$ allele that destabilizes the complex on the chromatin leads to early onset cancer (Shima et al., 2007; Chuang et al., 2010). These results indicate that the full complement of Mcm2-7 is required to prevent tumorigenesis. If additional Mcm2-7 loading, in G1, could be promoted it may prevent tumorigenesis.

Bibliography

Aggarwal, B. D., and Calvi, B. R. (2004). Chromatin regulates origin activity in *Drosophila* follicle cells. Nature *430*, 372-376.

Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. Nat Rev Mol Cell Biol *13*, 153-167.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). The Cell Cycle and Programmed Cell Death. In Molecular Biology of the Cell, eds. (NY,NY: Garland Science), pp. 983-1026.

Anderson, D. E., Losada, A., Erickson, H. P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. J Cell Biol *156*, 419-424.

Anglana, M., Apiou, F., Bensimon, A., and Debatisse, M. (2003). Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. Cell *114*, 385-394.

Aparicio, O. M., Stout, A. M., and Bell, S. P. (1999). Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. Proc Natl Acad Sci U S A *96*, 9130-9135.

Aparicio, O. M., Weinstein, D. M., and Bell, S. P. (1997). Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. Cell *91*, 59-69.

Aparicio, T., Guillou, E., Coloma, J., Montoya, G., and Mendez, J. (2009). The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication. Nucleic Acids Res *37*, 2087-2095.

Araki, M., Wharton, R. P., Tang, Z., Yu, H., and Asano, M. (2003). Degradation of origin recognition complex large subunit by the anaphase-promoting complex in *Drosophila*. EMBO J *22*, 6115-6126.

Arias, E. E., and Walter, J. C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. Nat Cell Biol *8*, 84-90.

Bell, S. P., and Kaguni, J. M. (2013). Helicase loading at chromosomal origins of replication. Cold Spring Harb Perspect Biol *5*:a010124.

Bell, S. P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128-134.

Besnard, E., Babled, A., Lapasset, L., Milhavet, O., Parrinello, H., Dantec, C., Marin, J. M., and Lemaitre, J. M. (2012). Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. Nat Struct Mol Biol *19*, 837-844.

Bester, A. C., Roniger, M., Oren, Y. S., Im, M. M., Sarni, D., Chaoat, M., Bensimon, A., Zamir, G., Shewach, D. S., and Kerem, B. (2011). Nucleotide deficiency promotes genomic instability in early stages of cancer development. Cell *145*, 435-446.

Bochman, M. L., and Schwacha, A. (2008). The Mcm2-7 complex has in *vitro* helicase activity. Mol Cell *31*, 287-293.

Bochman, M. L., and Schwacha, A. (2010). The Saccharomyces cerevisiae Mcm6/2 and Mcm5/3 ATPase active sites contribute to the function of the putative Mcm2-7 'gate'. Nucleic Acids Res *38*, 6078-6088.

Bowers, J. L., Randell, J. C., Chen, S., and Bell, S. P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell *16*, 967-978.

Bowman, G. D., O'Donnell, M., and Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. Nature *429*, 724-730.

Breier, A. M., Chatterji, S., and Cozzarelli, N. R. (2004). Prediction of Saccharomyces cerevisiae replication origins. Genome Biol *5*, R22.

Brewer, B. J., and Fangman, W. L. (1994). Initiation preference at a yeast origin of replication. Proc Natl Acad Sci U S A *91*, 3418-3422.

Broach, J. R., Li, Y. Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K. A., and Hicks, J. B. (1983). Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harb Symp Quant Biol *47 Pt 2*, 1165-1173.

Burkhart, R., Schulte, D., Hu, D., Musahl, C., Gohring, F., and Knippers, R. (1995). Interactions of human nuclear proteins P1Mcm3 and P1Cdc46. Eur J Biochem *228*, 431-438.

Cadoret, J. C., Meisch, F., Hassan-Zadeh, V., Luyten, I., Guillet, C., Duret, L., Quesneville, H., and Prioleau, M. N. (2008). Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proc Natl Acad Sci U S A *105*, 15837-15842.

Cayrou, C., Coulombe, P., Puy, A., Rialle, S., Kaplan, N., Segal, E., and Mechali, M. (2012). New insights into replication origin characteristics in metazoans. Cell Cycle *11*, 658-667.

Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R., and Mechali, M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. Genome Res *21*, 1438-1449.

Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E., and Campbell, J. L. (1984). Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. Mol Cell Biol *4*, 2455-2466.

Chen, S., and Bell, S. P. (2011). CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. Genes Dev *25*, 363-372.

Chen, S., de Vries, M. A., and Bell, S. P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. Genes Dev *21*, 2897-2907.

Chuang, C. H., Wallace, M. D., Abratte, C., Southard, T., and Schimenti, J. C. (2010). Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. PLoS Genet *6*, e1001110.

Chuang, L. C., Teixeira, L. K., Wohlschlegel, J. A., Henze, M., Yates, J. R., Mendez, J., and Reed, S. I. (2009). Phosphorylation of Mcm2 by Cdc7 promotes pre-replication complex assembly during cell-cycle re-entry. Mol Cell *35*, 206-216.

Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., and Nasmyth, K. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Mol Cell *5*, 243-254.

Clurman, B. E., Sheaff, R. J., Thress, K., Groudine, M., and Roberts, J. M. (1996). Turnover of cyclin E by the ubiquitin-proteasome pathway is regulated by cdk2 binding and cyclin phosphorylation. Genes Dev *10*, 1979-1990.

Cocker, J. H., Piatti, S., Santocanale, C., Nasmyth, K., and Diffley, J. F. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. Nature *379*, 180-182.

Craddock, N., Hurles, M. E., Cardin, N., Pearson, R. D., Plagnol, V., Robson, S., Vukcevic, D., Barnes, C., Conrad, D. F., Giannoulatou, E., Holmes, C., Marchini, J. L., Stirrups, K., Tobin, M. D., Wain, L. V., Yau, C., Aerts, J., Ahmad, T., Andrews, T. D., Arbury, H., Attwood, A., Auton, A., Ball, S. G., Balmforth, A. J., Barrett, J. C., Barroso, I., Barton, A., Bennett, A. J., Bhaskar, S., Blaszczyk, K., Bowes, J., Brand, O. J., Braund, P. S., Bredin, F., Breen, G., Brown, M. J., Bruce, I. N., Bull, J., Burren, O. S., Burton, J., Byrnes, J., Caesar, S., Clee, C. M., Coffey, A. J., Connell, J. M., Cooper, J. D., Dominiczak, A. F., Downes, K., Drummond, H. E., Dudakia, D., Dunham, A., Ebbs, B., Eccles, D., Edkins, S., Edwards, C., Elliot, A., Emery, P., Evans, D. M., Evans, G., Eyre, S., Farmer, A., Ferrier, I. N., Feuk, L., Fitzgerald, T., Flynn, E., Forbes, A., Forty, L., Franklyn, J. A., Freathy, R. M., Gibbs, P., Gilbert, P., Gokumen, O., Gordon-Smith, K., Gray, E., Green, E., Groves, C. J., Grozeva, D., Gwilliam, R., Hall, A., Hammond, N., Hardy, M., Harrison, P., Hassanali, N., Hebaishi, H., Hines, S., Hinks, A., Hitman, G. A., Hocking, L., Howard, E., Howard, P., Howson, J. M., Hughes, D., Hunt, S., Isaacs, J. D., Jain, M., Jewell, D. P., Johnson, T., Jolley, J. D., Jones, I. R., Jones, L. A., Kirov, G., Langford, C. F., Lango-Allen, H., Lathrop, G. M., Lee, J., Lee, K. L., Lees, C., Lewis, K., Lindgren, C. M., Maisuria-Armer, M., Maller, J., Mansfield, J., Martin, P., Massey, D. C., McArdle, W. L., McGuffin, P., McLay, K. E., Mentzer, A., Mimmack, M. L., Morgan, A. E., Morris, A. P., Mowat, C., Myers, S., Newman, W., Nimmo, E. R., O'Donovan, M. C., Onipinla, A., Onyiah, I., Ovington, N. R., Owen, M. J., Palin, K., Parnell, K., Pernet, D., Perry, J. R., Phillips, A., Pinto, D., Prescott, N. J., Prokopenko, I., Quail, M. A., Rafelt, S., Rayner, N. W., Redon, R., Reid, D. M., Renwick, Ring, S. M., Robertson, N., Russell, E., St Clair, D., Sambrook, J. G., Sanderson, J. D., Schuilenburg, H., Scott, C. E., Scott, R., Seal, S., Shaw-Hawkins, S., Shields, B. M., Simmonds, M. J., Smyth, D. J., Somaskantharajah, E., Spanova, K., Steer, S., Stephens, J., Stevens, H. E., Stone, M. A., Su, Z., Symmons, D. P., Thompson, J. R., Thomson, W., Travers, M. E., Turnbull, C., Valsesia, A., Walker, M., Walker, N. M., Wallace, C., Warren-Perry, M., Watkins, N. A., Webster, J., Weedon, M. N., Wilson, A. G., Woodburn, M., Wordsworth, B. P., Young, A. H., Zeggini, E., Carter, N. P., Frayling, T. M., Lee, C., McVean, G., Munroe, P. B., Palotie, A., Sawcer, S. J., Scherer, S. W., Strachan, D. P., Tyler-Smith, C., Brown, M. A., Burton, P. R., Caulfield, M. J., Compston, A., Farrall, M., Gough, S. C., Hall, A. S., Hattersley, A. T., Hill, A. V., Mathew, C. G., Pembrey, M., Satsangi, J., Stratton, M. R., Worthington, J., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W., Parkes, M., Rahman, N., Todd, J. A., Samani, N. J., and Donnelly, P. (2010). Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature *464*, 713-720.

Crevel, G., Hashimoto, R., Vass, S., Sherkow, J., Yamaguchi, M., Heck, M. M., and Cotterill, S. (2007). Differential requirements for MCM proteins in DNA replication in *Drosophila* S2 cells. PLoS One *2*, e833.

de Nooij, J. C., Letendre, M. A., and Hariharan, I. K. (1996). A cyclin-dependent kinase inhibitor, Dacapo, is necessary for timely exit from the cell cycle during *Drosophila* embryogenesis. Cell *87*, 1237-1247.

Dierssen, M., Herault, Y., and Estivill, X. (2009). Aneuploidy: from a physiological mechanism of variance to Down syndrome. Physiol Rev *89*, 887-920.

Dimitrova, D. S., Todorov, I. T., Melendy, T., and Gilbert, D. M. (1999). Mcm2, but not RPA, is a component of the mammalian early G1-phase prereplication complex. J Cell Biol *146*, 709-722.

Donnellan, R., and Chetty, R. (1999). Cyclin E in human cancers. FASEB J *13*, 773-780.

Donovan, S., Harwood, J., Drury, L. S., and Diffley, J. F. (1997). Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. Proc Natl Acad Sci U S A *94*, 5611-5616.

Drury, L. S., Perkins, G., and Diffley, J. F. (1997). The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. EMBO J *16*, 5966-5976.

Drury, L. S., Perkins, G., and Diffley, J. F. (2000). The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. Curr Biol *10*, 231-240.

Dulic, V., Lees, E., and Reed, S. I. (1992). Association of human cyclin E with a periodic G1-S phase protein kinase. Science *257*, 1958-1961.

Eaton, M. L., Galani, K., Kang, S., Bell, S. P., and MacAlpine, D. M. (2010). Conserved nucleosome positioning defines replication origins. Genes Dev *24*, 748-753.

Edwards, M. C., Tutter, A. V., Cvetic, C., Gilbert, C. H., Prokhorova, T. A., and Walter, J. C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. J Biol Chem *277*, 33049-33057.

Ekholm-Reed, S., Mendez, J., Tedesco, D., Zetterberg, A., Stillman, B., and Reed, S. I. (2004). Deregulation of cyclin E in human cells interferes with prereplication complex assembly. J Cell Biol *165*, 789-800.

Elsasser, S., Chi, Y., Yang, P., and Campbell, J. L. (1999). Phosphorylation controls timing of Cdc6p destruction: A biochemical analysis. Mol Biol Cell *10*, 3263-3277.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc Natl Acad Sci U S A *106*, 20240-20245.

Evrin, C., Fernandez-Cid, A., Zech, J., Herrera, M. C., Riera, A., Clarke, P., Brill, S., Lurz, R., and Speck, C. (2013). In the absence of ATPase activity, pre-RC formation is blocked prior to MCM2-7 hexamer dimerization. Nucleic Acids Res *41*, 3162-3172.

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

Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M., Aburatani, H., Jones, K. W., Tyler-Smith, C., Hurles, M. E., Carter, N. P., Scherer, S. W., and Lee, C. (2006). Copy number variation: new insights in genome diversity. Genome Res *16*, 949-961.

Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V., Brewer, B. J., and Fangman, W. L. (1996). Multiple determinants controlling activation of yeast replication origins late in S phase. Genes Dev *10*, 1595-1607.

Fu, Y. V., Yardimci, H., Long, D. T., Ho, T. V., Guainazzi, A., Bermudez, V. P., Hurwitz, J., van Oijen, A., Scharer, O. D., and Walter, J. C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell *146*, 931-941.

Fujita, M., Hori, Y., Shirahige, K., Tsurimoto, T., Yoshikawa, H., and Obuse, C. (1998).

Cell cycle dependent topological changes of chromosomal replication origins in Saccharomyces cerevisiae. Genes Cells *3*, 737-749.

Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol *8*, 358-366.

Gambus, A., Khoudoli, G. A., Jones, R. C., and Blow, J. J. (2011). MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. J Biol Chem *286*, 11855-11864.

Ge, X. Q., Jackson, D. A., and Blow, J. J. (2007). Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. Genes Dev *21*, 3331-3341.

Ge, X. Q., and Blow, J. J. (2010). Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. J Cell Biol *191*, 1285-1297.

Gelbart, M. E., and Kuroda, M. I. (2009). *Drosophila* dosage compensation: a complex voyage to the X chromosome. Development *136*, 1399-1410.

Geng, Y., Lee, Y. M., Welcker, M., Swanger, J., Zagozdzon, A., Winer, J. D., Roberts, J. M., Kaldis, P., Clurman, B. E., and Sicinski, P. (2007). Kinase-independent function of cyclin E. Mol Cell *25*, 127-139.

Geng, Y., Whoriskey, W., Park, M. Y., Bronson, R. T., Medema, R. H., Li, T., Weinberg, R. A., and Sicinski, P. (1999). Rescue of cyclin D1 deficiency by knockin cyclin E. Cell *97*, 767-777.

Gillespie, P. J., and Hirano, T. (2004). Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. Curr Biol *14*, 1598-1603.

Gillespie, P. J., Li, A., and Blow, J. J. (2001). Reconstitution of licensed replication origins on Xenopus sperm nuclei using purified proteins. BMC Biochem *2*, 15.

Goldman, M. A., Holmquist, G. P., Gray, M. C., Caston, L. A., and Nag, A. (1984). Replication timing of genes and middle repetitive sequences. Science *224*, 686-692.

Goren, A., Tabib, A., Hecht, M., and Cedar, H. (2008). DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin. Genes Dev *22*, 1319-1324.

Graveley, B. R., Brooks, A. N., Carlson, J. W., Duff, M. O., Landolin, J. M., Yang, L., Artieri, C. G., van Baren, M. J., Boley, N., Booth, B. W., Brown, J. B., Cherbas, L., Davis, C. A., Dobin, A., Li, R., Lin, W., Malone, J. H., Mattiuzzo, N. R., Miller, D., Sturgill, D., Tuch, B. B., Zaleski, C., Zhang, D., Blanchette, M., Dudoit, S., Eads, B., Green, R. E., Hammonds, A., Jiang, L., Kapranov, P., Langton, L., Perrimon, N., Sandler, J. E., Wan, K. H., Willingham, A., Zhang, Y., Zou, Y., Andrews, J., Bickel,

P. J., Brenner, S. E., Brent, M. R., Cherbas, P., Gingeras, T. R., Hoskins, R. A., Kaufman, T. C., Oliver, B., and Celniker, S. E. (2011). The developmental transcriptome of *Drosophila* melanogaster. Nature *471*, 473-479.

Guacci, V., and Koshland, D. (2012). Cohesin-independent segregation of sister chromatids in budding yeast. Mol Biol Cell *23*, 729-739.

Guillou, E., Ibarra, A., Coulon, V., Casado-Vela, J., Rico, D., Casal, I., Schwob, E., Losada, A., and Mendez, J. (2010). Cohesin organizes chromatin loops at DNA replication factories. Genes Dev *24*, 2812-2822.

Haering, C. H., Lowe, J., Hochwagen, A., and Nasmyth, K. (2002). Molecular architecture of SMC proteins and the yeast cohesin complex. Mol Cell *9*, 773-788.

Hancock, R. (1974). Interphase chromosomal deoxyribonucleoprotein isolated as a discrete structure from cultured cells. J Mol Biol *86*, 649-663.

Harvey, K. J., and Newport, J. (2003). CpG Methylation of DNA Restricts Prereplication Complex Assembly in Xenopus Egg Extracts. Molecular and Cellular Biology *23*, 6769-6779.

Hedglin, M., Kumar, R., and Benkovic, S. J. (2013a). Replication clamps and clamp loaders. Cold Spring Harb Perspect Biol *5*, a010165.

Hedglin, M., Perumal, S. K., Hu, Z., and Benkovic, S. (2013b). Stepwise assembly of the human replicative polymerase holoenzyme. Elife *2*, e00278.

Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S., and Bell, S. P. (2011). Eukaryotic origin-dependent DNA replication in *vitro* reveals sequential action of DDK and S-CDK kinases. Cell *146*, 80-91.

Holland, A. J., and Cleveland, D. W. (2009). Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nat Rev Mol Cell Biol *10*, 478-487.

Huang, C. E., Milutinovich, M., and Koshland, D. (2005). Rings, bracelet or snaps: fashionable alternatives for Smc complexes. Philos Trans R Soc Lond B Biol Sci *360*, 537-542.

Huberman, J. A., and Riggs, A. D. (1968). On the mechanism of DNA replication in mammalian chromosomes. J Mol Biol *32*, 327-341.

Ibarra, A., Schwob, E., and Mendez, J. (2008). Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. Proc Natl Acad Sci U S A *105*, 8956-8961.

Ilves, I., Petojevic, T., Pesavento, J. J., and Botchan, M. R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Mol Cell *37*, 247-258.

Jackson, D. A., and Pombo, A. (1998). Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol *140*, 1285-1295.

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol *28*, 329-347.

Jares, P., and Blow, J. J. (2000). Xenopus cdc7 function is dependent on licensing but not on XORC, XCdc6, or CDK activity and is required for XCdc45 loading. Genes Dev *14*, 1528-1540.

Jin, J., Arias, E. E., Chen, J., Harper, J. W., and Walter, J. C. (2006). A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. Mol Cell *23*, 709-721.

Johansson, E., and Dixon, N. (2013). Replicative DNA polymerases. Cold Spring Harb Perspect Biol *5*: a012799.

Jones, R. M., Mortusewicz, O., Afzal, I., Lorvellec, M., Garcia, P., Helleday, T., and Petermann, E. (2012). Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. Oncogene 32, 3744-3753.

Kamimura, Y., Tak, Y. S., Sugino, A., and Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. EMBO J *20*, 2097-2107.

Kanemaki, M., and Labib, K. (2006). Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. EMBO J *25*, 1753-1763.

Kenna, M. A., and Skibbens, R. V. (2003). Mechanical link between cohesion establishment and DNA replication: Ctf7p/Eco1p, a cohesion establishment factor, associates with three different replication factor C complexes. Mol Cell Biol *23*, 2999-3007.

Kim, T. M., Xi, R., Luquette, L. J., Park, R. W., Johnson, M. D., and Park, P. J. (2013). Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. Genome Res *23*, 217-227.

Knoblich, J. A., Sauer, K., Jones, L., Richardson, H., Saint, R., and Lehner, C. F. (1994). Cyclin E controls S phase progression and its down-regulation during *Drosophila* embryogenesis is required for the arrest of cell proliferation. Cell *77*, 107-120.

Koch, B., Kueng, S., Ruckenbauer, C., Wendt, K. S., and Peters, J. M. (2008). The Suv39h-HP1 histone methylation pathway is dispensable for enrichment and protection of cohesin at centromeres in mammalian cells. Chromosoma *117*, 199-210.

Koepp, D. M., Schaefer, L. K., Ye, X., Keyomarsi, K., Chu, C., Harper, J. W., and Elledge, S. J. (2001). Phosphorylation-dependent ubiquitination of cyclin E by the SCFFbw7 ubiquitin ligase. Science *294*, 173-177.

Krasinska, L., Besnard, E., Cot, E., Dohet, C., Mechali, M., Lemaitre, J. M., and Fisher, D. (2008). Cdk1 and Cdk2 activity levels determine the efficiency of replication origin firing in Xenopus. EMBO J *27*, 758-769.

Kreitz, S., Ritzi, M., Baack, M., and Knippers, R. (2001). The human origin recognition complex protein 1 dissociates from chromatin during S phase in HeLa cells. J Biol Chem *276*, 6337-6342.

Krishna, T. S., Fenyo, D., Kong, X. P., Gary, S., Chait, B. T., Burgers, P., and Kuriyan, J. (1994a). Crystallization of proliferating cell nuclear antigen (PCNA) from Saccharomyces cerevisiae. J Mol Biol *241*, 265-268.

Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994b). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell *79*, 1233-1243.

Krude, T., Musahl, C., Laskey, R. A., and Knippers, R. (1996). Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. J Cell Sci *109*, 309-318.

Krysan, P. J., and Calos, M. P. (1991). Replication initiates at multiple locations on an autonomously replicating plasmid in human cells. Mol Cell Biol *11*, 1464-1472.

Kumar, R., Nashine, V. C., Mishra, P. P., Benkovic, S. J., and Lee, T. H. (2010). Stepwise loading of yeast clamp revealed by ensemble and single-molecule studies. Proc Natl Acad Sci U S A *107*, 19736-19741.

Labib, K., Diffley, J. F., and Kearsey, S. E. (1999). G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. Nat Cell Biol *1*, 415-422.

Labib, K. (2000). Uninterrupted MCM2-7 Function Required for DNA Replication Fork Progression. Science *288*, 1643-1647.

Lane, M. E., Sauer, K., Wallace, K., Jan, Y. N., Lehner, C. F., and Vaessin, H. (1996). Dacapo, a cyclin-dependent kinase inhibitor, stops cell proliferation during *Drosophila* development. Cell *87*, 1225-1235.

Latt, S. A. (1973). Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. Proc Natl Acad Sci U S A *70*, 3395-3399.

Lee, S. H., Eki, T., and Hurwitz, J. (1989). Synthesis of DNA containing the simian virus 40 origin of replication by the combined action of DNA polymerases alpha and delta. Proc Natl Acad Sci U S A *86*, 7361-7365.

Lee, S. H., Kwong, A. D., Pan, Z. Q., and Hurwitz, J. (1991). Studies on the activator 1 protein complex, an accessory factor for proliferating cell nuclear antigen-dependent DNA polymerase delta. J Biol Chem *266*, 594-602.

Lees, E., Faha, B., Dulic, V., Reed, S. I., and Harlow, E. (1992). Cyclin E/cdk2 and cyclin A/cdk2 kinases associate with p107 and E2F in a temporally distinct manner. Genes & Development *6*, 1874-1885.

Lei, M., Kawasaki, Y., and Tye, B. K. (1996). Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae. Mol Cell Biol *16*, 5081-5090.

Leonard, A. C., and Mechali, M. (2013). DNA replication origins. Cold Spring Harb Perspect Biol *5*, a010116.

Li, C. J., and DePamphilis, M. L. (2002). Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition in the cell division cycle. Mol Cell Biol *22*, 105-116.

Li, J. J., and Kelly, T. J. (1984). Simian virus 40 DNA replication in *vitro*. Proc Natl Acad Sci U S A *81*, 6973-6977.

Li, J. J., and Kelly, T. J. (1985). Simian virus 40 DNA replication in *vitro*: specificity of initiation and evidence for bidirectional replication. Mol Cell Biol *5*, 1238-1246.

Li, M. Z., and Elledge, S. J. (2012). SLIC: a method for sequence- and ligation-independent cloning. Methods Mol Biol *852*, 51-59.

Liku, M. E., Nguyen, V. Q., Rosales, A. W., Irie, K., and Li, J. J. (2005). 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.

Liu, E., Li, X., Yan, F., Zhao, Q., and Wu, X. (2004). Cyclin-dependent kinases phosphorylate human Cdt1 and induce its degradation. J Biol Chem *279*, 17283-17288.

Loupart, M. L., Krause, S. A., and Heck, M. S. (2000). Aberrant replication timing induces defective chromosome condensation in *Drosophila* ORC2 mutants. Curr Biol *10*, 1547-1556.

Lubelsky, Y., DeNapoli, L., Prinz, J., and MacAlpine, D. M. (In press). Genome Research

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

Ma, H., Samarabandu, J., Devdhar, R. S., Acharya, R., Cheng, P. C., Meng, C., and Berezney, R. (1998). Spatial and temporal dynamics of DNA replication sites in mammalian cells. J Cell Biol *143*, 1415-1425.

MacAlpine, H. K., Gordan, R., Powell, S. K., Hartemink, A. J., and MacAlpine, D. M. (2010). *Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading. Genome Res *20*, 201-211.

Madine, M. A., Khoo, C. Y., Mills, A. D., Musahl, C., and Laskey, R. A. (1995). The nuclear envelope prevents reinitiation of replication by regulating the binding of MCM3 to chromatin in Xenopus egg extracts. Curr Biol *5*, 1270-1279.

Mahbubani, H. M., Chong, J. P., Chevalier, S., Thommes, P., and Blow, J. J. (1997). Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor. J Cell Biol *136*, 125-135.

Mailand, N., and Diffley, J. F. (2005). CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis. Cell *122*, 915-926.

Maiorano, D., Moreau, J., and Mechali, M. (2000). XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. Nature *404*, 622-625.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N., and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? Annu Rev Biochem *79*, 89-130.

Masai, H., Taniyama, C., Ogino, K., Matsui, E., Kakusho, N., Matsumoto, S., Kim, J. M., Ishii, A., Tanaka, T., Kobayashi, T., Tamai, K., Ohtani, K., and Arai, K. (2006). Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. J Biol Chem *281*, 39249-39261.

Masumoto, H., Muramatsu, S., Kamimura, Y., and Araki, H. (2002). S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. Nature *415*, 651-655.

Mayer, M. L., Gygi, S. P., Aebersold, R., and Hieter, P. (2001). Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in *S. cerevisiae*. Mol Cell *7*, 959-970.

McGarry, T. J., and Kirschner, M. W. (1998). Geminin, an inhibitor of DNA replication, is degraded during mitosis. Cell *93*, 1043-1053.

Mendez, J., Zou-Yang, X. H., Kim, S. Y., Hidaka, M., Tansey, W. P., and Stillman, B. (2002). Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. Mol Cell *9*, 481-491.

Mesner, L. D., Crawford, E. L., and Hamlin, J. L. (2006). Isolating apparently pure libraries of replication origins from complex genomes. Mol Cell *21*, 719-726.

Mimura, S., Seki, T., Tanaka, S., and Diffley, J. F. (2004). Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. Nature *431*, 1118-1123.

Moberg, K. H., Bell, D. W., Wahrer, D. C., Haber, D. A., and Hariharan, I. K. (2001). Archipelago regulates Cyclin E levels in *Drosophila* and is mutated in human cancer cell lines. Nature *413*, 311-316.

Moyer, S. E., Lewis, P. W., and Botchan, M. R. (2006). 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.

Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y., and Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. Genes Dev *24*, 602-612.

Murray, A. W. (2004). Recycling the cell cycle: cyclins revisited. Cell *116*, 221-234.

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

Nasmyth, K. (2011). Cohesin: a catenase with separate entry and exit gates? Nat Cell Biol *13*, 1170-1177.

Nguyen, V. Q., Co, C., Irie, K., and Li, J. J. (2000). Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. Curr Biol *10*, 195-205.

Nguyen, V. Q., Co, C., and Li, J. J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. Nature *411*, 1068-1073.

Nieduszynski, C. A., Knox, Y., and Donaldson, A. D. (2006). Genome-wide identification of replication origins in yeast by comparative genomics. Genes Dev *20*, 1874-1879.

Nishitani, H., Sugimoto, N., Roukos, V., Nakanishi, Y., Saijo, M., Obuse, C., Tsurimoto, T., Nakayama, K. I., Nakayama, K., Fujita, M., Lygerou, Z., and Nishimoto, T. (2006). Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. EMBO J *25*, 1126-1136.

Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S. I., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nat Cell Biol *4*, 89-93.

Novak, B., Toth, A., Csikasz-Nagy, A., Gyorffy, B., Tyson, J. J., and Nasmyth, K. (1999). Finishing the cell cycle. J Theor Biol *199*, 223-233.

O'Donnell, M., Langston, L., and Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb Perspect Biol *5*, Ohtsubo, M., and Roberts, J. M. (1993). Cyclin-dependent regulation of G1 in mammalian fibroblasts. Science *259*, 1908-1912.

Ohtsubo, M., Theodoras, A. M., Schumacher, J., Roberts, J. M., and Pagano, M. (1995). Human cyclin E, a nuclear protein essential for the G1-to-S phase transition. Mol Cell Biol *15*, 2612-2624.

Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H., and Walter, J. C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. Mol Cell *21*, 581-587.

Pacek, M., and Walter, J. C. (2004). A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. EMBO J *23*, 3667-3676.

Pak, D. T., Pflumm, M., Chesnokov, I., Huang, D. W., Kellum, R., Marr, J., Romanowski, P., and Botchan, M. R. (1997). Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. Cell *91*, 311-323.

Peters, J. M., and Nishiyama, T. (2012). Sister chromatid cohesion. Cold Spring Harb Perspect Biol *4*: a011130.

Pflumm, M. F., and Botchan, M. R. (2001). Orc mutants arrest in metaphase with abnormally condensed chromosomes. Development *128*, 1697-1707.

Polyak, K., Kato, J. Y., Solomon, M. J., Sherr, C. J., Massague, J., Roberts, J. M., and Koff, A. (1994). p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. Genes & Development *8*, 9-22.

Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L., and Stillman, B. (2004). Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. EMBO J *23*, 2651-2663.

Prasanth, S. G., Shen, Z., Prasanth, K. V., and Stillman, B. (2010). Human origin recognition complex is essential for HP1 binding to chromatin and heterochromatin organization. Proc Natl Acad Sci U S A *107*, 15093-15098.

Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B., and Stillman, B. (1987). The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in *vitro*. Nature *326*, 471-475.

Prelich, G., and Stillman, B. (1988). Coordinated leading and lagging strand synthesis during SV40 DNA replication in *vitro* requires PCNA. Cell *53*, 117-126.

Pruitt, S. C., Bailey, K. J., and Freeland, A. (2007). Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. Stem Cells *25*, 3121-3132.

Team, R. C. (2012). R: A Language and Environment for Statistical Computing. (Vienna:

R Foundation for Statistical Computing).

Randell, J. C., Bowers, J. L., Rodriguez, H. K., and Bell, S. P. (2006). Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. Mol Cell *21*, 29-39.

Randell, J. C., Fan, A., Chan, C., Francis, L. I., Heller, R. C., Galani, K., and Bell, S. P. (2010). Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. Mol Cell *40*, 353-363.

Rao, H., and Stillman, B. (1995). The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proc Natl Acad Sci U S A *92*, 2224-2228.

Remus, D., Beall, E. L., and Botchan, M. R. (2004). DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC-DNA binding. EMBO J *23*, 897-907.

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., and Diffley, J. F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719-730.

Remus, D., Blanchette, M., Rio, D. C., and Botchan, M. R. (2005). CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the *Drosophila* origin recognition complex. J Biol Chem *280*, 39740-39751.

Resnitzky, D., Gossen, M., Bujard, H., and Reed, S. I. (1994). Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. Mol Cell Biol *14*, 1669-1679.

Rhind, N., and Gilbert, D. M. (2013). DNA replication timing. Cold Spring Harb Perspect Biol *5*, a010132.

Rialland, M., Sola, F., and Santocanale, C. (2002). Essential role of human CDT1 in DNA replication and chromatin licensing. J Cell Sci *115*, 1435-1440.

Ritzi, M., Baack, M., Musahl, C., Romanowski, P., Laskey, R. A., and Knippers, R. (1998). Human minichromosome maintenance proteins and human origin recognition complex 2 protein on chromatin. J Biol Chem *273*, 24543-24549.

Rollins, R. A., Korom, M., Aulner, N., Martens, A., and Dorsett, D. (2004). *Drosophila* nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. Mol Cell Biol *24*, 3100-3111.

Romanowski, P., Madine, M. A., and Laskey, R. A. (1996a). XMCM7, a novel member of the Xenopus MCM family, interacts with XMCM3 and colocalizes with it throughout replication. Proc Natl Acad Sci U S A *93*, 10189-10194.

Romanowski, P., Madine, M. A., Rowles, A., Blow, J. J., and Laskey, R. A. (1996b). The

Xenopus origin recognition complex is essential for DNA replication and MCM binding to chromatin. Curr Biol *6*, 1416-1425.

Roth, Y. F. (1987). Eucaryotic primase. Eur J Biochem *165*, 473-481.

Rowles, A., Tada, S., and Blow, J. J. (1999). Changes in association of the Xenopus origin recognition complex with chromatin on licensing of replication origins. J Cell Sci *112*, 2011-2018.

Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., Lin, M. F., Washietl, S., Arshinoff, B. I., Ay, F., Meyer, P. E., Robine, N., Washington, N. L., Di Stefano, L., Berezikov, E., Brown, C. D., Candeias, R., Carlson, J. W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M. Y., Will, S., Alekseyenko, A. A., Artieri, C., Booth, B. W., Brooks, A. N., Dai, Q., Davis, C. A., Duff, M. O., Feng, X., Gorchakov, A. A., Gu, T., Henikoff, J. G., Kapranov, P., Li, R., MacAlpine, H. K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S. K., Riddle, N. C., Sakai, A., Samsonova, A., Sandler, J. E., Schwartz, Y. B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K. H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S. E., Brent, M. R., Cherbas, L., Elgin, S. C., Gingeras, T. R., Grossman, R., Hoskins, R. A., Kaufman, T. C., Kent, W., Kuroda, M. I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J. W., Ren, B., Russell, S., Cherbas, P., Graveley, B. R., Lewis, S., Micklem, G., Oliver, B., Park, P. J., Celniker, S. E., Henikoff, S., Karpen, G. H., Lai, E. C., MacAlpine, D. M., Stein, L. D., White, K. P., and Kellis, M. (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. Science *330*, 1787-1797.

Rudra, S., and Skibbens, R. V. (2013). Cohesin codes - interpreting chromatin architecture and the many facets of cohesin function. J Cell Sci *126*, 31-41.

Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M., and Zink, D. (1999). Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. J Cell Biol *146*, 1211-1226.

Santocanale, C., Sharma, K., and Diffley, J. F. (1999). Activation of dormant origins of DNA replication in budding yeast. Genes Dev *13*, 2360-2364.

Schraml, P., Bucher, C., Bissig, H., Nocito, A., Haas, P., Wilber, K., Seelig, S., Kononen, J., Mihatsch, M. J., Dirnhofer, S., and Sauter, G. (2003). Cyclin E overexpression and amplification in human tumours. J Pathol *200*, 375-382.

Schubeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J., and Groudine, M. (2002). Genome-wide DNA replication profile for *Drosophila* melanogaster: a link between transcription and replication timing. Nat Genet *32*, 438-442.

Schvartzman, J. M., Sotillo, R., and Benezra, R. (2010). Mitotic chromosomal instability and cancer: mouse modelling of the human disease. Nat Rev Cancer *10*, 102-115.

Sekedat, M. D., Fenyo, D., Rogers, R. S., Tackett, A. J., Aitchison, J. D., and Chait, B. T. (2010). GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. Mol Syst Biol *6*, 353.

Senga, T., Sivaprasad, U., Zhu, W., Park, J. H., Arias, E. E., Walter, J. C., and Dutta, A. (2006). PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. J Biol Chem *281*, 6246-6252.

Sequeira-Mendes, J., Diaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N., and Gomez, M. (2009). Transcription initiation activity sets replication origin efficiency in mammalian cells. PLoS Genet *5*, e1000446.

Shima, N., Alcaraz, A., Liachko, I., Buske, T. R., Andrews, C. A., Munroe, R. J., Hartford, S. A., Tye, B. K., and Schimenti, J. C. (2007). A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nat Genet *39*, 93-98.

Shimada, K., and Gasser, S. M. (2007). The origin recognition complex functions in sister-chromatid cohesion in Saccharomyces cerevisiae. Cell *128*, 85-99.

Siddiqui, K., On, K. F., and Diffley, J. F. X. (2013). Regulating DNA Replication in Eukarya. Cold Spring Harbor Perspectives in Biology *5*, a012930-a012930.

Skibbens, R. V., Corson, L. B., Koshland, D., and Hieter, P. (1999). Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. Genes Dev *13*, 307-319.

Smyth, G. K. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor, ed. (New York: Springer), pp. 397-420.

Speck, C., Chen, Z., Li, H., and Stillman, B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. Nat Struct Mol Biol *12*, 965-971.

Spencer, S. L., Cappell, S. D., Tsai, F. C., Overton, K. W., Wang, C. L., and Meyer, T. (2013). The Proliferation-Quiescence Decision Is Controlled by a Bifurcation in CDK2 Activity at Mitotic Exit. Cell *155*, 369-383.

Spruck, C. H., Won, K. A., and Reed, S. I. (1999). Deregulated cyclin E induces chromosome instability. Nature *401*, 297-300.

Stillman, B. W., and Gluzman, Y. (1985). Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. Mol Cell Biol *5*, 2051-2060.

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

Storchova, Z., and Kuffer, C. (2008). The consequences of tetraploidy and aneuploidy. J Cell Sci *121*, 3859-3866.

Strohmaier, H., Spruck, C. H., Kaiser, P., Won, K. A., Sangfelt, O., and Reed, S. I. (2001). Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. Nature *413*, 316-322.

Su, J. Y., Rempel, R. E., Erikson, E., and Maller, J. L. (1995). Cloning and characterization of the Xenopus cyclin-dependent kinase inhibitor p27XIC1. Proc Natl Acad Sci U S A *92*, 10187-10191.

Svitin, A., and Chesnokov, I. (2010). Study of DNA replication in *Drosophila* using cell free in *vitro* system. Cell Cycle *9*, 815-819.

Symeonidou, I. E., Kotsantis, P., Roukos, V., Rapsomaniki, M. A., Grecco, H. E., Bastiaens, P., Taraviras, S., and Lygerou, Z. (2013). Multi-step Loading of Human Minichromosome Maintenance Proteins in Live Human Cells. J Biol Chem *288*, 35852-35867.

Szuplewski, S., Sandmann, T., Hietakangas, V., and Cohen, S. M. (2009). *Drosophila* Minus is required for cell proliferation and influences Cyclin E turnover. Genes Dev *23*, 1998-2003.

Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y., and Araki, H. (2006). A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. EMBO J *25*, 1987-1996.

Takahashi, T. S., Basu, A., Bermudez, V., Hurwitz, J., and Walter, J. C. (2008). Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in Xenopus egg extracts. Genes Dev *22*, 1894-1905.

Takahashi, T. S., and Walter, J. C. (2005). Cdc7-Drf1 is a developmentally regulated protein kinase required for the initiation of vertebrate DNA replication. Genes Dev *19*, 2295-2300.

Takahashi, T. S., Wigley, D. B., and Walter, J. C. (2005). Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. Trends Biochem Sci *30*, 437-444.

Takahashi, T. S., Yiu, P., Chou, M. F., Gygi, S., and Walter, J. C. (2004). Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. Nat Cell Biol *6*, 991-996.

Tanaka, S., and Araki, H. (2013). Helicase Activation and Establishment of Replication Forks at Chromosomal Origins of Replication. Cold Spring Harb Perspect Biol

Tanaka, S., and Diffley, J. F. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. Nat Cell Biol *4*, 198-207.

Tanaka, S., Nakato, R., Katou, Y., Shirahige, K., and Araki, H. (2011). Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. Curr Biol *21*, 2055-2063.

Tanaka, S., Tak, Y. S., and Araki, H. (2007). The role of CDK in the initiation step of DNA replication in eukaryotes. Cell Div *2*, 16.

Tanaka, T., Fuchs, J., Loidl, J., and Nasmyth, K. (2000). Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. Nat Cell Biol *2*, 492-499.

Tanaka, T., and Nasmyth, K. (1998). Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. EMBO J *17*, 5182-5191.

Taylor, J. H. (1960). Asynchronous duplication of chromosomes in cultured cells of Chinese hamster. J Biophys Biochem Cytol *7*, 455-464.

Terret, M. E., Sherwood, R., Rahman, S., Qin, J., and Jallepalli, P. V. (2009). Cohesin acetylation speeds the replication fork. Nature *462*, 231-234.

Thomson, A. M., Gillespie, P. J., and Blow, J. J. (2010). Replication factory activation can be decoupled from the replication timing program by modulating Cdk levels. J Cell Biol *188*, 209-221.

Torres, E. M., Williams, B. R., and Amon, A. (2008). Aneuploidy: cells losing their balance. Genetics *179*, 737-746.

Toth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. (1999). Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. Genes Dev *13*, 320-333.

Toyoshima, H., and Hunter, T. (1994). p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. Cell *78*, 67-74.

Tsakraklides, V., and Bell, S. P. (2010). Dynamics of pre-replicative complex assembly. J Biol Chem *285*, 9437-9443.

Tsuji, T., Ficarro, S. B., and Jiang, W. (2006). Essential role of phosphorylation of MCM2 by Cdc7/Dbf4 in the initiation of DNA replication in mammalian cells. Mol Biol Cell *17*, 4459-4472.

Tsurimoto, T., and Stillman, B. (1990). Functions of replication factor C and proliferating-cell nuclear antigen: functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. Proc Natl Acad Sci U S A *87*, 1023-1027.

Tsurimoto, T., and Stillman, B. (1991). Replication factors required for SV40 DNA replication in *vitro*. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. J Biol Chem *266*, 1950-1960.

Uhlmann, F., and Nasmyth, K. (1998). Cohesion between sister chromatids must be established during DNA replication. Curr Biol *8*, 1095-1101.

Valton, A. L., Hassan-Zadeh, V., Lema, I., Boggetto, N., Alberti, P., Saintome, C., Riou, J. F., and Prioleau, M. N. (2014). G4 motifs affect origin positioning and efficiency in two vertebrate replicators. EMBO J 33, 732-746.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T. J., and Walter, J. C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes Dev *17*, 1894-1908.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. Mol Cell *10*, 1223-1233.

Whittaker, A. J., Royzman, I., and Orr-Weaver, T. L. (2000). *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.

Wilmes, G. M., Archambault, V., Austin, R. J., Jacobson, M. D., Bell, S. P., and Cross, F. R. (2004). Interaction of the S-phase cyclin Clb5 with an "RXL" docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. Genes Dev *18*, 981-991.

Wimmel, A., Lucibello, F. C., Sewing, A., Adolph, S., and Muller, R. (1994). Inducible acceleration of G1 progression through tetracycline-regulated expression of human cyclin E. Oncogene *9*, 995-997.

Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., and Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science *290*, 2309-2312.

Wold, M. S., and Kelly, T. (1988). Purification and characterization of replication protein A, a cellular protein required for in *vitro* replication of simian virus 40 DNA. Proc Natl Acad Sci U S A *85*, 2523-2527.

Won, K. A., and Reed, S. I. (1996). Activation of cyclin E/CDK2 is coupled to site-specific autophosphorylation and ubiquitin-dependent degradation of cyclin E. EMBO J *15*, 4182-4193.

Woodward, A. M., Gohler, T., Luciani, M. G., Oehlmann, M., Ge, X., Gartner, A., Jackson, D. A., and Blow, J. J. (2006). Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. J Cell Biol *173*, 673-683.

Wyrick, J. J., Aparicio, J. G., Chen, T., Barnett, J. D., Jennings, E. G., Young, R. A., Bell, S. P., and Aparicio, O. M. (2001). Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. Science *294*, 2357-2360.

Yang, S. C., Rhind, N., and Bechhoefer, J. (2010). Modeling genome-wide replication kinetics reveals a mechanism for regulation of replication timing. Mol Syst Biol *6*, 404.

Yoshizawa-Sugata, N., Ishii, A., Taniyama, C., Matsui, E., Arai, K., and Masai, H. (2005). A second human Dbf4/ASK-related protein, Drf1/ASKL1, is required for efficient progression of S and M phases. J Biol Chem *280*, 13062-13070.

Zegerman, P., and Diffley, J. F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281-285.

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet *10*, 451-481.

Zhang, N., Kuznetsov, S. G., Sharan, S. K., Li, K., Rao, P. H., and Pati, D. (2008). A handcuff model for the cohesin complex. J Cell Biol *183*, 1019-1031.

Zhang, Y., Malone, J. H., Powell, S. K., Periwal, V., Spana, E., Macalpine, D. M., and Oliver, B. (2010). Expression in aneuploid *Drosophila* S2 cells. PLoS Biol *8*, e1000320.

Zink, D., Bornfleth, H., Visser, A., Cremer, C., and Cremer, T. (1999). Organization of early and late replicating DNA in human chromosome territories. Exp Cell Res *247*, 176-188.

Zou, L. (1998). Formation of a Preinitiation Complex by S-phase Cyclin CDK-Dependent Loading of Cdc45p onto Chromatin. Science *280*, 593-596.

Biography

Sara Kathleen Powell was born in Abilene, Kansas, USA on May 15th. She attended the University of Illinois at Chicago and received her B.S. with a major in Biochemistry in 2004. She got her Doctor of Philosophy at Duke University in 2014. Her publications include:

Powell, S.K, J. Prinz, H.K. MacAlpine, and D.M. MacAlpine. “The dynamic loading and redistribution of the Mcm2-7 complex throughout the cell cycle.” Under Revision for EMBO.

Lee, H. et al. “DNA copy number evolution in *Drosophila* modENCODE cell lines.” Accepted Genome Biology.

Roy, S., J. Ernst, et al. (2010). "Identification of functional elements and regulatory circuits by *Drosophila* modENCODE." Science 330(6012): 1787-1797.

Zhang, Y., J. H. Malone, **S.K. Powell**, V. Periwal, E. Spana, D.M. MacAlpine and B. Oliver. (2010). "Expression in aneuploid *Drosophila* S2 cells." PLoS Biol 8(2): e1000320.

MacAlpine, H. K., R. Gordan, **S.K. Powell**, A.J. Hartemink and D.M. MacAlpine (2010). "*Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading." Genome Res 20(2): 201-211.

Plummer, R. S., C. R. Shea, M. Nelson, **S.K. Powell**, D.M. Freeman, C.P. Dan and D. Lang (2008). "PAX3 expression in primary melanomas and nevi." Mod Pathol 21(5): 525-530.

Lang, D., J. B. Mascarenhas, **S.K. Powell**, J. Halegoua, M. Nelson and B.A. Ruggeri (2008). "PAX6 is expressed in pancreatic adenocarcinoma and is downregulated during induction of terminal differentiation." Mol Carcinog 47(2): 148-156.

Lang, D., **S.K. Powell**, R.S. Plummer, K.P. Young, and B.A. Ruggeri (2007). "PAX genes: roles in development, pathophysiology, and cancer." Biochem Pharmacol 73(1): 1-14.