TUMOR SUPPRESSION BY REGULATION OF MINI-CHROMOSOME MAINTENANCE EXPRESSION IN RESPONSE TO DNA REPLICATION STRESS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Gongshi Bai

May 2016

MECHANISTIC INSIGHT IN MCM EXPRESION REGULATION DURING CELLULAR RESPONSE TO DNA REPLICATION STRESS TO PREVENT NEOPLASIA

Gongshi Bai

Cornell University 2016

DNA replication is an essential process during cell proliferation, when genomic information is completely and precisely duplicated. DNA replication machinery (replisome) is responsible for the generation and maintenance of replication structure (replication fork), where genomic DNA is replicated semi-conservatively. Conditions that can inhibit replisome progression are recognized as DNA replication stress (RS). RS originates from many endogenous sources, thus proliferating cells inevitably experience RS during DNA replication throughout their replicative lifespan. DNA damage response (DDR) is responsible for the detection and removal of DNA lesions associated with aberrant DNA replication. However, some of these DNA damages may escape the detection from DDR and accumulate in proliferating cell lineages, which pose great threat to genome integrity maintenance. Furthermore, genome alterations of oncogenic natures may contribute to malignant transformation and eventual neoplasia. Little is known about how proliferating cells record the overall experience of accumulative RS during their prolonged replicative lifespan to prevent cancerous transformation.

Minichromosome Maintenance 2-7 (MCM2-7) factors are essential for replication origin licensing and which also constitute catalytic core of replicative helicase. MCM2-7 proteins also partner with DDR, especially in the presence of RS, to faithfully replicate the genome. One of

such actions is through sufficient origin licensing. MCM2-7 proteins are excessively expressed in proliferating cells to sufficiently license origins, including dormant origins, which are only activated in response to RS to rescue the nearby stalled replisomes. In contrast to the idea that cells requires high MCM2-7 expression to ensure proper response to RS, I found that chronic RS actually suppresses MCM2-7 expression. This regulation is mediated by the DDR mechanism, and depends on the central tumor suppressor *Trp53* in mice, which is partly accomplished through microRNA mediated gene silencing. Moderate MCM2-7 downregulation does not affect normal DNA replication and cell proliferation, however reduced MCM expression sensitized cells to additional RS by inducing terminal cell cycle arrest such as senescence to prevent cellular transformation. Thus, MCM2-7 expression level serves as a molecular memory of accumulative RS experienced by the proliferating cells, when normal high expression level ensures sufficient RS tolerance during their replicative lifespan.

BIOGRAPHICAL SKETCH

Gongshi Bai was born on January 7th, 1987 in Tianjin, China. A month later, he moved back to Beijing where his family was and grew up there. He attended every school, including kindergarten that was affiliated to Peking University. After graduation from Peking University affiliated High School, he made the obvious choice to enter Peking University for college education, majoring in Biological Sciences. During the four years in college, he was exposed to basic ideas such as evolution and molecular biology. He started to participate in active scientific researches under the guidance of Dr. Long Hong in Dr. Lijia Qu's laboratory to study the molecular function of eIF6 during abiotic stress response in *Arabidopsis*. He strived to acquire many molecular biology research skills under the mentoring of many excellent researchers in the laboratory, and became very capable experimentally by the time he graduated from Peking University with a B.S. degree in 2009. In the same year, he came to the U.S. and was admitted into the graduate program of Genetics and Development (currently Genetics, Genomics and Development), in the Department of Molecular Biology and Genetics, Cornell University. He became interested in genome maintenance mechanisms during DNA replication and cancer biology. Inspired by the experience of a previous graduate student, he wanted to combine the elegant and powerful yeast genetics with animal research to study cancer genetics by joining Dr. Bik-Kwoon Tye and Dr. John Schimenti's laboratories. Unfortunately, the plan did not work out as Dr. Bik-Kwoon Tye moved her laboratory to Hong Kong, China. In 2011, Gongshi joined Dr. John Schimenti's laboratory to study the regulation of mini-chromosome maintenance (MCM) genes during cellular response to DNA replication stress and its impact on genome integrity maintenance and prevention of malignant transformation. His detailed work helped to change the interpretation of a cancer model, which had been studied by the lab and others for many years.

He also discovered a new mechanism on how cells normally repress MCM expression in response to replication stress, which contributes to the prevention of replication stress induced cellular transformation. After graduation from Cornell University, Gongshi plans to remain in academia and continues scientific researches in the area of genome integrity maintenance and cancer biology.

谨以此毕业论文献给在地球另一端日夜牵挂我的父母。

Everything good in me starts from you.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. John Schimenti for his gracious supports over the years. It has been a great pleasure working with John, and a tremendous learning experience for myself so to become a scientist as Dr. Schimenti some day. I heartily appreciate that John provided great amount of research freedom so I could pursue the direction as the research directed me to. Thank you, John.

Second, I would like to thank Dr. Bik-Kwoon Tye for her kind supports when I first came to Cornell. Although the original plan to work with Dr. Tye had to be altered, I enjoyed guidance and suggestions later over the years from Bik. I would also like to thank my committee members Dr. Robert Weiss and Dr. Natasza Kurpios for their discussion and valuable input into my research.

Third, I would like to thank previous and present members of the Schimenti Lab. Dr. Chuang Chen Hua was the initiator of my thesis research project, who also laid down great amount of ground work that we are still working on in the lab. Dr. Adrian McNairn provided great amount of technical support, who also helped with scientific discussions.

Finally, I would like to thank my family on the side of the globe, who are always there for me spiritually. Without your support, I could not have been here. And Willa Wang, for the lovely time we had together.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

LITERATURE REVIEW

1.1. Eukaryotic DNA replication

1.1.1. Overview

Cell proliferation is the process when cell number increases as a consequence of cell growth and division. Unicellular organisms depend on cell proliferation to reproduce. In metazoan, a small group of cells (i.e. stem/progenitor cells) are devoted to cell proliferation to replenish cell population and produce new cells that can differentiate into other cell types to perform specialized biological functions. Key to the cell proliferation process is DNA replication, when the genomic information is duplicated once in the mother cell and equally distributed into the two daughter cells.

The massive and highly complex eukaryotic genome locates on spatially separated chromosomes. Genomic information residing on these chromosomes must be duplicated coordinately within the defined S-phase of the eukaryotic cell cycle. In accordance with the replication program in eukaryotic cells, the DNA replication machinery ("replisome") must be regulated in a spatial and temporal manner. First, replisomes assemble on all the chromosomes and physically move through every part of the genome to duplicate the parental DNA semi-conservatively. The completeness of DNA replication is thus ensured. However, reassembly of the replisome on newly synthesized DNA is strictly prohibited to prevent over-replication(1, 2). Secondly, functional assembly and activation of the replisome can only occur during S-phase of the cell cycle, despite the preparation for replisome formation (replication licensing) is already initiated during late M phase of the previous cycle (3). Eukaryotic DNA replication is normally tightly regulated yet can display high level of plasticity. For instance, the same *Drosophila*

melanogaster genome can be replicated more than 10 times faster in embryo than in somatic cells (4). In all, the feature that replication machinery can be controlled temporally and spatially is key to the faithful and complete duplication of the eukaryotic genomic information.

1.1.2. Two stage model of eukaryotic DNA replication

DNA replication process can be generally divided into two stages in every system studied (Figure 1.1). During the first stage called initiation, components of the replisome are recruited to the parental DNA. Later during elongation, replisome assembly is complete and is further activated to extend DNA synthesis semi-conservatively.

The first step of replication initiation is the binding of initiator proteins to the defined DNA sequence within the genome known as DNA replication origins. In eukaryotic cells, Origin Recognition Complex (ORC) proteins are the initiators. The six ORC subunits (Orc1-6) form a heterohexamer complex to perform origin binding, and are highly conserved throughout eukaryotes. The binding of ORCs to origin provides the first touch between the replication proteins and their DNA template, which serves as a landing pad for further replication protein recruitment. Next, ORCs bound at origins recruit Mini-Chromosome Maintenance 2-7 (MCM2-7) complex and form the pre-Replication Complex (pre-RC), with the assistant of other pre-RC factors such as Cdc6 and Cdt1. Pre-RC formation is also known as replication licensing. MCM2-7 loading defines the licensing event, which makes the genomic DNA replication-competent. MCM2-7 proteins are not only essential for replication licensing; they are also the core components of the replication helicase complex that unwinds double-strand DNA (dsDNA) template during replication elongation. The last step of initiation involves incorporating MCM2-7 into functional helicase complex, which then is activated to melt dsDNA at replication origins. This reaction provides room for the formation of the rest of replisome, which is responsible for

replication elongation. As mentioned before, eukaryotic DNA replication is a highly dynamic yet tightly controlled process. These regulations ensure complete and faithful duplication of the genomic information in a timely manner. Most of the regulations are imposed on the initiation stage of replication, while elongation can also be controlled [reviewed in (5-7)]. The detailed processes of the two stages eukaryotic DNA replication will be summarized in the following sections.

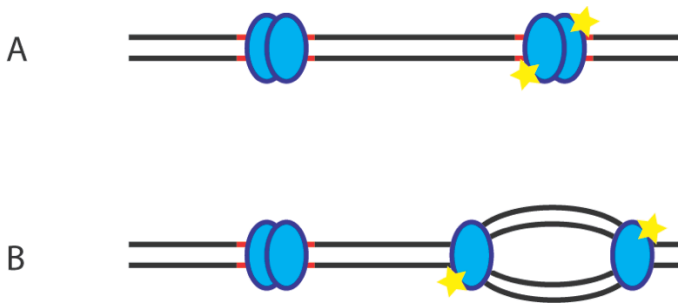


Figure 1.1 Two stage model of DNA replication. (A) Replication initiation. Replication proteins (replisome, blue circles) start to associate with DNA (black line) at selected regions (red line, origins). Some replisomes will be activated (blue circle with yellow star) and initiate **(B)** bidirectional elongation of the DNA synthesis semi-conservatively.

1.1.3. Selective DNA replication licensing and initiation

1.1.3.1. Sequence characteristic of DNA replication origins

The first step of DNA replication initiation is the binding of initiator protein to DNA replication origins. In *Escherichia coli*, the 4Mbp circular genome has only one replication origin, the *oriC*. The ~250bp fragment contains multiple repeats of a clearly defined consensus sequence, which can be bound by a single initiator protein DnaA with high affinity (8). Unlike the prokaryotes, the massive eukaryotic genome resides on multiple spatially separated liner chromosomes. This complex arrangement of the genomic information requires the presence of numerous origins on each chromosome, so replisomes can assemble on all of the chromosomes to completely replicate genomic information.

In *Saccharomyces cerevisiae* (budding yeast), the presence of replication origins were first implied using the Autonomous Replicating Sequence (ARS) assay, when cloned genomic fragment can autonomously initiate replication and thus the maintenance of a extrachromosomal plasmid in yeast cells (9). The cloned ARSs were later shown to be *bona fide* replication origins in budding yeast (10, 11). Various genome-wide techniques were also used to identify all of the ARSs in the budding yeast genome, thus the common characters of ARS sequences were identified (12, 13). In budding yeast, the ARS is about 100bp long. It contains an 17bp A-domain with a 11bp A/T rich ACS [ARS consensus sequence, (A/T)TTTA(T/C)(A/G)TTT(A/T)], and B-domains that further determines the specificity and strength of ORCs binding. However, budding yeast is the only eukaryotic organism that has a defined replication origin sequence. ARS in *Schizosaccharomyces pombe* (fission yeast) usually spans ~1kb long that is highly A/T rich (~70%), but clearly lacks consensus sequence (14). Efforts to identify defined ARS in

metazoans were not successful. Except for a few genomic loci that serve as replication origins, most of the replication events are initiated from a broad "initiation zone" (15).

1.1.3.2. Relaxed binding of ORCs defines replication origins in eukaryotes

In the budding yeast genome, an estimated of over 12,000 potential ACS sites can be identified based on the consensus sequence signature (16). However, only ~500 of them are bound by ORCs-MCM during licensing (13, 17); ~300 of these licensed origins actually fire to initiate DNA replication during S phase, while the rest remains dormant (12). These observations suggest 1) ORCs binding is necessary in defining replication origins. 2) ACS is important but not necessary for ORCs binding. Other genomic signatures of the replication origins define licensing events and 3) ORCs-MCM binding does not guarantee origin usage. Only a subset of licensed origins is further selected to initiate DNA replication.

What is the genomic DNA signature that determines ORC binding and thus replication licensing sites in the genome? In budding yeast, ORC binding specificity and affinity is usually determined by the B-domains of the ARS in addition to the highly conserved A-domain (5). In fission yeast, the A/T hook domain of the Orc4 subunit preferentially bindings to A/T rich stretch of the DNA, which contributes to the high A/T content selection of replication origins in this species (18, 19). In metazoans, replication origins lack sequence specificity. However, replication is still initiated from certain genomic loci that are mitotically stable (20). In human and fruit fly, ORC binding colocalize with transcription regulatory elements (21, 22), however transcription *per se* does not contribute to selective ORCs binding. In general, unlike prokaryote and budding yeast, relaxed binding of ORCs protein to replication origins without apparent sequence specificity is a conserved feature of the eukaryotes. Other genomic features such as

epigenetic chromatin modifications and chromosome arrangement within the nucleus, or local chromatin accessibility in general determine selective ORCs binding.

What factors contribute to the preferential replication initiation from a subset of the licensed replication origins, while the rest (dormant origins) are not used during normal replication? Now it's known that chromatin structure rather than the DNA sequence determines selective firing. In budding yeast, deletion of the Sir2 histone deacetylase, which establishes and maintains repressive chromatin structure, can rescue replication initiation defects (23, 24). A combination of histone deacetylase deletion and overexpression of the replication initiation factors can activate dormant origins (25). Higher hierarchy chromatin structure such as chromatin loop also regulates origin firing. Modulation of chromatin cohesion changes the patterns of chromatin loop formation, which in turn affects replication initiation frequency (26). In general, local chromatin accessibility decided by epigenetic chromosomal modifications and higher order chromosome structure, determines the binding affinity of lowly expressed replication initiation factors to licensed origins, thus selectively activate a subset of origins over the others such as the dormant origins. The function of dormant origins will be discussed later.

1.1.4. DNA replication timing

Besides the selective origins usage based on their physical location within the genome, different origins may fire during different period of the S-phase. This temporal regulation on replication initiation is known as DNA replication timing (RT). The mechanisms of physical and temporal selection of licensed origins are highly interconnected. For example, the preferentially selected origins usually fire early during S phase, while the genomic regions containing dormant origins are passively replicated by converging replication forks.

The RT program was first implied and visualized using 5-bromo-2'-deoxyuridine (BrdU) incorporation and fluorescent microscopy [reviewed in (27)]. In the presence of nucleotide analogue BrdU, clusters of replicating DNA form BrdU positive foci. The number and localization of these foci within the nucleus changes during S phase progression (28), however mitotically stable for a specific cell type (29). Each foci is recognized as a replication domain, which contains 5~10 replication forks. Replication origins within each domain fire coordinately, while the RT program regulates different domains globally during S phase.

Many factors determine the RT program. Early genome-wide studies established a strong correlation between high gene transcription activity and early replication in higher eukaryotes (30-32). Thus it was implied that transcription regulation dictates replication timing. However, many examples contradicted this hypothesis, as altered transcription *per se* failed to change replication timing. Histone modification was shown to have great impact on RT decision (33). In metazoans, development status and cell differentiation can also change RT. The X chromosome inactivation in mice accompanies a shift from early to late replication timing transition during early embryonic development (34). The human β globin locus changes from late to early replication during differentiation into erythroid cells, in which the gene is also highly expressed (35). A genome-wide study demonstrated global change of replication domains after mouse embryonic stem cell (ESC) differentiation into neural precursor cells (NPCs), when independent ESC lines and induced pluripotent stem cells (iPS) have identical RT profile (36).

The establishment of the RT program happens early in G1 phase during the hypothesized Timing Decision Point (TDP). Cells isolated after TDP retains the normal temporal replication control, suggesting the establishment of RT during a certain cell cycle stage (37). Given that this hypothetical TDP happens soon after nuclear envelope formation and coincides with the

chromosome reorganization event in the nucleus, it was thus proposed that the 3-dimensional chromosome positioning in the nucleus determines the replication domains and thus the RT. Later evidences further support this hypothesis (38). In fact, the aforementioned chromosome modification and development program might be primarily determined by this high order chromosome architecture, thus producing distinguished RT profile for different cell types (36, 39, 40). This hypothesis also account for the distinct BrdU staining patterns during different stages of S phase, whereas early replicated foci are retained within the central nucleus while the mid/late replicated domains are located at the nuclear peripheral, where the chromosomes are anchored or stabilized by sub-nuclear structures.

1.1.5. DNA replication licensing

After binding to replication origins, ORCs further recruits MCM2-7 heterohexamer complex onto DNA. This step is known as licensing, which defines the true commitment to DNA replication. Licensing also distinguishes unreplicated vs. newly synthesized DNA and thus the latter will not be replicated again. MCM2-7 is loaded at ORCs bound origins by Cdc6 and Cdt1, two other components of the pre-replication complex (pre-RC), and is the only protein complex that later remains in the actual replisome, making MCM2-7 good targets for effective DNA replication control.

Replication licensing is tightly regulated in accordance with cell cycle progression. All of the licensing events must be restricted to late M to early G1 phase and are completely prohibited during S phase to prevent relicensing. The details about DNA replication process will be reviewed in the following sections.

1.1.5.1. Biochemistry of pre-replication complex (pre-RC) formation

The goal of pre-RC formation is to load MCM2-7 complex onto chromatin. Biochemically, the loading process can be divided into two phases based on the affinity of MCM2-7 association with chromatin DNA. During the "association" phase, MCM2-7/chromatin interaction is relatively weak and is sensitive to high-salt washes *in vitro*. Once MCM2-7 is stably "lock and loaded" onto chromatin, it becomes irremovable from chromatin by high salt wash. The transition from unstable to stable MCM2-7/chromatin interaction depends on ATP hydrolysis by ORCs and Cdc6 [reviewed in (1, 5, 6)].

All but ORC6 of the ORC1-6 proteins are AAA+ ATPases (ATPases Associated with various cellular Activities). ORCs binding to DNA depends on ATP, however ATP hydrolysis is not required at this step (41). CDC6 is also an AAA+ ATPase, which binds to ORCs at origin DNA in an ATP bound form. The ATPase activity of CDC6 helps to define ORCs binding specificity to replication origins (42, 43). Cdc6, in association with Orc1-6 forms a ring-shaped structure that contains six (Orc1-5 & Cdc6) AAA+ ATPase subunits. The overall dimension of this heptamer is similar to MCM2-7, which may help the loading of the latter complex onto DNA (42).

CDT1 serves as a bridge between ORC-CDC6 and MCM2-7 complex. CDT1 interacts with MCM2-7 (44) and is recruited to origin DNA by ORCs-Cdc6. This intermediate pre-RC structure (salt-sensitive phase) contains ORCs-Cdc6 in the ATP bound form, and CDT1 in association with a single MCM2-7 hexamer (45, 46). ATP hydrolysis by ORCs and Cdc6 opens up MCM2-7 ring so it can encircle the double stranded DNA within its central channel once Cdc6 and Cdt1 dissociate (salt-insensitive phase). Two copies of MCM2-7 hexamers are loaded onto chromatin DNA in a head-to-head conformation during each round of ORCs-Cdc6 action,

which is later important for bi-directional DNA replication initiation (46, 47). However, the exact mechanism governing the loading of the second MCM2-7 hexamer is unclear, probably through a secondary interaction site between Orc6 and Cdt1 (48). The MCM2-7 hexamer pairs circling the origin DNA can then slide off to make room for additional MCM2-7 loading (47). Thus, MCM2-7 loading onto chromatin is complete and makes it replication-competent (Figure 1.2). The entire replication licensing reaction had been recapitulated *in vitro* using purified pre-RC proteins (47).

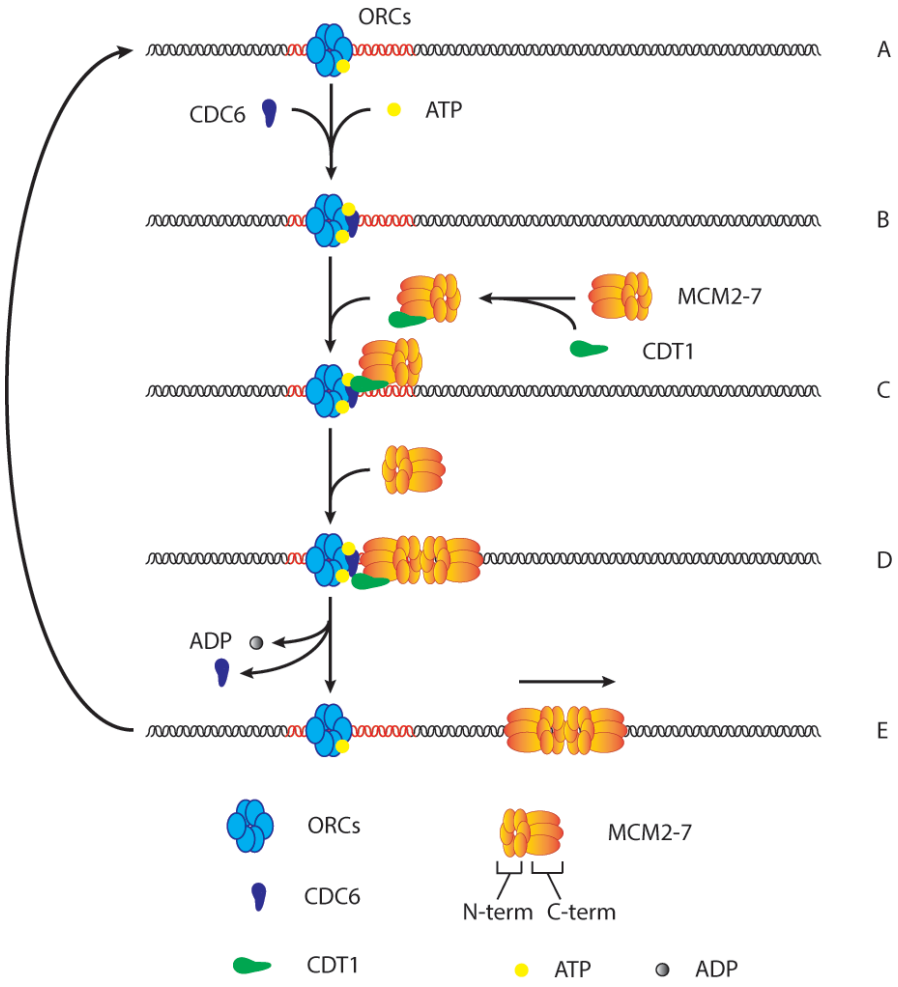


Figure 1.2 DNA replication licensing. (A) Licensing starts by the binding of ORCs to the replication origins. ATP binding but not hydrolysis is important for ORCs-DNA binding. **(B)** ORCs further recruit Cdc6 in the ATP bound form and form a ring-shaped structure on chromatin, presumably facilitates further MCM2-7 loading. **(C)** Cdt1 recruits MCM2-7 hexamer to ORCs-Cdc6 bound origins. At this step, MCM2-7 are only associated with replication origins (salt-sensitive). **(D)** MCM2-7 hexamer recruits another MCM2-7 hexamer through their interaction between the N-terminal of the MCMs. ATP hydrolysis opens up MCM2-7 ring to encircle the double stranded DNA through its central channel (salt-insensitive). Two MCM2-7 hexamers are loaded in a head-to-head conformation. **(E)** Loaded MCM2-7 hexamers can slide off and make room for additional MCM2-7 loading.

1.1.5.2. Inhibition of re-replication

Licensing on newly synthesized DNA must be prohibited once replication is initiated to prevent re-replication of any region of the genome. Thus, licensing is strictly limited to late M to early G1 phase, and no additional licensing is allowed during S phase. This temporal licensing regulation is achieved through spatially separation of licensing factors from replicating DNA into a different cellular compartment during cell cycle progression. Given that licensing factors ORCs, Cdc6 & Cdt1 are only needed for loading but not the maintenance of MCM2-7 on DNA, these components of the pre-RC are primary targets of re-licensing prevention.

In yeast, *de novo* pre-RC formation is primarily prohibited by Cyclin dependent kinases (CDKs) through various mechanisms. CDKs-dependent phosphorylation of CDC6 targets the protein for rapid degradation during S phase (49, 50). CDKs activity also excludes free-form MCM2-7 from nucleus (51). In metazoan, Cdt1 availability is mainly targeted to prevent re-licensing. CDK-dependent Cdt1 degradation during G1/S transition prevents pre-RC formation during S phase (52). Geminin, a protein that is only present in metazoan cells stably binds to CDT1 during S and G2 phase to prevent its interaction with MCM2-7 thus re-licensing (53). Proteolysis of geminin during mitosis renders CDT1 available for licensing in the next cell cycle (54). In general, re-replication is normally prevented by redundant mechanisms in eukaryotic cells, otherwise leading to genomic instability and cancer (1).

1.1.6. Helicase activation

By the end of G1 phase, the majority of ORCs bound genomic regions are also loaded with multiple copies of MCM2-7 hexamers. These MCM2-7 hexamers, however, is in a helicase inactive form, consistent with the *in vitro* observations that MCM2-7 complex alone does not possess helicase activity. It also suggests additional factors are needed for MCM2-7 helicase

activation. In eukaryotes, these cofactors include highly conserved Cdc45 and GINS complex (Sld5, Psf1, Psf2, and Psf3). Together with MCM2-7, these proteins form the stable CMG complex (Cdc45, MCM2-7 & GINS) that has *in vitro* helicase activity and functions as replication helicase *in vivo* (55). Helicase activation is the process when Cdc45 and GINS are recruited to MCM2-7 at licensed origins to assemble into the CMG helicase complex. CMG formation melts origin DNA and provides DNA template for further replisome assembly. This process requires two types of kinases activity: CDKs and Dbf4-dependent kinase (DDK), which also connect physical activation of replication machinery to the temporal cell cycle control. The targets of CDKs and DDK during helicase activation diverged during eukaryotic evolution, however the general mechanisms remain similar.

DDK contains two subunits: Cdc7 as the catalytic protein, whose activity is regulated by the partner subunit Dbf4. In budding yeast, DDK mainly phosphorylate the N-terminal domain of MCMs during helicase activation (56). These post-translational modifications of MCMs are likely to cause conformational changes in the MCM2-7 complex, which facilitates assembly of CMG helicase complex (57). Early evidences came from the study of *mcm5-bob1* mutant which carries a point mutation in MCM5 protein (58). This mutation causes a conformational change to MCM5 and thus the entire MCM2-7 complex, which bypasses the DDK-dependent replication initiation in budding yeast. MCM2-7 phosphorylation by DDK also enhances the association of Cdc45 to licensed origins (59). As mentioned before, two MCM2-7 hexamers are loaded in a head-to-head conformation at replication origins, with N-terminal domains of MCM2-7 from each complex facing the other. DDK phosphorylation may also alter this N-terminal interface, thus facilitates the separation of two MCM2-7 complexes during helicase activation, which leads to bidirectional replication fork assembly (57). S phase CDKs (S-CDKs) phosphorylate Sld2 and

Sld3, which increases their association with Dpb11 scaffold protein. Together, they help to recruit GINS complex to MCM2-7 at licensed origins (57, 59). The entire process of helicase activation has also been recapitulated using purified proteins *in vitro* (60). Helicase activation in metazoans is slightly different from that in budding yeast, with many changes to the initiation proteins.

How does the MCM2-7 hexamers loaded at replication origins remain in the helicase inactive form? First, MCM2-7 encircles dsDNA during replication licensing, however it only displays helicase activity on DNA template that contains ssDNA region, suggesting the exclusion of one strand of the DNA duplex while maintaining the other within the MCM2-7 channel is how this helicase unwinds dsDNA. Thus the MCM2-7 ring must be opened once more after licensing during helicase activation. Second, structurally analysis on MCM2-7 complex suggests a gapped-ring shaped hexamer, with MCM2 & 5 residing on either side of the opening (57, 61). This discontinuity in the hexamer probably disrupts the helicase activity exerted by the coordinated interactions between MCM2-7. Binding of Cdc45 and GINS sealed the gap between MCM2 & 5. Additional nucleotide association further caused a conformational change to the entire helicase complex, constricting the big channel formed between MCM2-7 and Cdc45 / GINS into two smaller ones. DNA duplex may go through these two conduits while being separated by the MCM2-7 helicase during replication (61). The requirement of helicase cofactors thus functionally separates MCM2-7 complex from licensing factors to core helicase components.

1.1.7. DNA replication elongation

Helicase activation at each replication origins results in the assembly of two replisomes, moving bidirectional in the form of two “replication forks”. At each replication fork, replication

helicase unwinds dsDNA, which is followed by the rest of the replisome that synthesize DNA semi-conservatively. DNA metabolism is achieved by DNA polymerases. Three DNA polymerases are used during regular eukaryotic genomic DNA replication, each of which is a multiunit complex: DNA polymerase α, ε and δ. Due to the antiparallel nature of the DNA duplex and that polymerase can only synthesis DNA in the 5' to 3' direction, DNA replication on the two separated strands can be classified into "leading" and "lagging" strand DNA synthesis.

Once replication helicase is activated, it unwinds dsDNA and generates ssDNA regions at replication origins (also known as "origin melting"). The ssDNA is then coated by replication protein A (RPA), a heterotrimeric ssDNA binding complex essential for all forms of DNA metabolism. RPA binding recruits DNA polymerase α (Pol α), which in turn inhibits excessive helicase unwinding at origins (62). Pol α holoenzyme contains four different subunits: Pol1, Pol12, Pri1 and Pri2. Pol1 performs the DNA polymerase function while Pri1 catalyzes RNA primer formation during both leading and lagging strand DNA synthesis (63). Among all the DNA polymerases, Pol α has the unique ability to initiate *de novo* DNA synthesis owing to its simultaneous polymerase and primase activity. Pol α first synthesizes a short RNA primer (~10nt) and further extends it with a short DNA fragment (~20nt) (64). However, Pol α has very limited processivity, DNA replication on leading and lagging strand is then taken over by other polymerases (65).

The RNA and short DNA synthesized by Pol α provides short dsDNA on the unwound ssDNA. This primer-template (dsDNA-ssDNA) junction is recognized by replication factor C (RFC), a heteropentameric AAA+ ATPase that serves as a DNA clamp loader (66, 67). Proliferating cell nuclear antigen (PCNA) is loaded at this junction by RFC and forms a homotrimeric complex, which tethers DNA polymerase to the site of DNA synthesis. This

interaction also greatly enhance polymerase activity and processivity (68). On the leading strand, Pol ε continuously synthesize DNA following primer sequence synthesized by Pol α. However, Pol α functions repeatedly on the lagging strand to produce short RNA-DNA hybrid primers, which are then extended by Pol δ to generate a series of discontinued DNA fragments. A recent genetic analysis suggested that Pol δ is the major polymerase that is responsible for DNA synthesis on both the leading and lagging strands, contradictory to the conventional views (69). The products (100~200nt) containing both the 5' RNA and 3' DNA sequence during lagging strand synthesis are known as Okazaki fragments (70). The 5' RNA and a few immediate DNA nucleotides following it are displaced by the incoming Pol δ synthesis, creating a "flap" structure. Pol δ coordinates with Flap endonuclease 1 (Fen1) to remove the flap (71), followed by DNA ligase I sealing two adjacent Okazaki fragments (72). Disrupted Pol δ / Fen1 coordination leads to production of longer flap structure, which would also be inclined to form secondary structures. Dna2 has both helicase and nuclease activity, is needed to process long and structured flaps, generating short substrates for optimal Fen1 action (73). Polymerases and helicase function must be coordinated as well to faithfully replicate DNA on both the leading and lagging strand.

1.1.8. Summary

DNA replication in eukaryotic cells can be generally divided into two stages: initiation and elongation. Initiation is the period when replisome components selectively bind to genomic DNA regions known as origins, when the licensed genome becomes DNA replication competent. The remaining components of the replisome are further recruited to the licensed origins and assemble into fully functional replication forks, which elongate to finish replication semi-conservatively. In contrast to DNA replication in prokaryotes, eukaryotic DNA replication is more complex and requires the coordination of replication events through space and time.

DNA replication in eukaryotes is tightly regulated. Most of the replication regulation is at the initiation step, which happens at multiple levels. First, genomic DNA regions (origins) are selectively bound by licensing proteins. Origins must be distributed throughout the massive eukaryotic genome, which reside on physically separated chromosomes. Second, a small subset of the licensed origins will initiate DNA replication, while the rest will be passively replicated by these primary firing events, or fire later, or not being used at all. The factors that activate licensed origins are limited and tightly controlled, so some licensed origins are physically selected over the others. Origin usage is also controlled by the replication-timing program, which is tightly connected to this physical selection. In higher eukaryotes, origin selection is likely to be governed by the local accessibility of the replication proteins to the chromosome. This chromatin accessibility is determined by local epigenetic modifications, which is primarily decided by the high-order chromosome organization in the nucleus such as chromosome tethering and chromosome loop formation. These characters are mitotically stable and thus a subset of origins is selected for usage in a single cell type with minor alterations during each generation. It can be changed once cell alters identity, i.e. through changes during development programming such as differentiation or reprograming (iPS induction). The unicellular budding yeast is an exception in eukaryotes in terms of defined origin locations and selection, probably due to streamlined selection of this organism through its interaction with human activity over the history. Finally, replication licensing and the activation of the licensed origins must have absolute no overlap. Origins activation must be inhibited before whole genome licensing is accomplished; licensing does not occur during the elongation phase so newly synthesized genome regions are not replicated more than once.

Besides the local regulation of replication events in eukaryotic cells, all of the replisomes scattered across the entire genome must follow temporally coordinated regulation. This is achieved by the diffusible components of the replisome and their regulatory partners. The physical regulation of replication is thus connected to the temporal regulation, as cell cycle related regulatory factors determine the affinity of these diffusible replisome components to the rests that are already assembled on the chromatin, and when or where replication will start.

Eukaryotic DNA replication is also regulated during the elongation phase, especially under sub-optimal conditions that inhibit replication. Any defects sensed during replication elongation can also regulate initiation events. The detailed cellular response to DNA damage arouse through replication will be discussed in later sections.

1.2. Mini-Chromosome maintenance proteins during eukaryotic DNA replication

Mini-chromosome maintenance (MCM) proteins are essential eukaryotic DNA replication proteins. They were primarily identified in a genetic screening performed in *Saccharomyces cerevisiae*, in which the MCM mutants failed to maintain an artificial mini-chromosome that is not essential for cell survival. MCMs discovered in this way were thus named after the assay. Other members of the MCM family were isolated in other genetic screenings. Since they all share similar biochemical characters and have related biological function, they were later named as MCMs.

The actual biological functions of MCMs were debated since their first discovery, which were all related to DNA replication. As described in the previous section, MCM proteins serves as replication licensing factors and the core of replication helicase complex. Thus, they function throughout the entire eukaryotic DNA replication process. Indeed, MCM2-7, the six major members of the MCM family, are the only proteins present during replication licensing and

further participate in replication elongation, making them good targets for effective DNA replication control. Besides their well-studied role in DNA replication, recent studies also indicate their function during gene transcription, DNA damage response and repair, chromosome remodeling and the maintenance of genome integrity. In this section, I will summarize the current understanding on the MCM proteins, from their biochemical character to the related functions.

1.2.1. The discovery of MCM proteins

The first genetic screenings for MCMs were performed in budding yeast *Saccharomyces cerevisiae* using mini-chromosome maintenance assay. A circular artificial plasmid containing a defined centromere sequence, a single ARS and two selective markers was transfected into yeast cells (74). Yeast mutants that failed to maintain this plasmid under different culturing conditions either had DNA replication or centromere related defects, which leads to the loss of these artificial mini-chromosomes. Later, one of these identified MCM mutants: Mcm2 was further studied. It was confirmed that Mcm2 mutants caused loss of mini-chromosome independent of centromere sequence, regardless of the ARS sequence being analyzed. This Mcm2 mutant had DNA replication rather than centromere defects, and was the first evidence suggesting a DNA replication role of MCM gene (75). Similar screening performed in budding and fission yeast followed guidelines generated in the original studies and identified additional member of MCM family that are essential genes involved in DNA replication (76, 77).

1.2.2. Mcm2-7 are highly conserved and closely related paralogues

MCM mutants discovered in the original genetic studies were divided into six complementation groups, each represented by a single Mcm2-7 gene. Sequence analysis of these six genes showed they share high level of sequence similarity. Each of the Mcm2-7 gene encodes

an AAA+ ATPase, that contains Walker A and B domains for ATP binding and hydrolysis, respectively (77). This ATPase domain within in the MCM genes is also termed "MCM box", which distinguishes Mcm2-7 from other MCM genes that lacks this domain but were identified in the same MCM screening. Each MCM2-7 protein also contains a zinc-finger domain between the N-terminal and MCM box. Mutations in the zinc-finger domain can disrupt MCM2-7 complex formation (78, 79) and cause cell survival defects (80, 81). Eukaryotic MCM genes are also very similar to the archaeal MCM. In two of archaeal genomes (*Methanobacterium thermoautotrophicum* and *Archaeoglobus fulgidus*), only one MCM gene was identified in each species (82, 83), suggesting the eukaryotic Mcm2-7 probably arise through gene duplication events. The most conserved eukaryotic and archaeal MCM domains are the MCM box and zinc-finger, suggesting the functional evolution of MCM predates the emergence of eukaryotes. Due to this similarity, archaeal MCM was used as a model to study eukaryotic MCMs. MCM2-7 are also highly conserved throughout eukaryotes. About 50-70% of the amino acids are conserved between yeast and human MCM2-7 proteins, with MCM box and zinc-finger being the most conserved regions (84).

1.2.3. MCM2-7 complex formation

Early genetic evidences suggested that MCM genes perform DNA replication related functions by interacting with other MCMs (85). Later, physical interactions between MCM subunits were identified using various biochemical methods in multiple eukaryotic systems (86-89). MCM2-7 proteins often co-precipitate with each other during biochemical purifications, implying that MCM2-7 forms a large heterohexameric complex containing each of the six proteins at equal stoichiometry. Smaller complexes were also observed, suggesting MCM subunits may preferentially form sub-complexes before assembling into the large complex (86).

It is known that MCM4, 6 & 7 form a stable trimeric complex, which also has helicase activity *in vitro* (90). MCM3 & 5 associate with each other to form a dimer (91). MCM2 preferentially interact with MCM4, 6 & 7 trimeric complex, while interacting with MCM3 & 5 dimer to eventually assemble into the hexameric MCM2-7 complex. The order of MCM2-7 subunits within the hexamer complex is thus determined (77).

The structure of MCM2-7 complex was first studied using electron microscopy (EM) on MCM proteins purified from fission yeast (87). The six subunits together adopt either a planer or globular structure with a central channel wide enough to accommodate dsDNA. The planer or ring-shaped structure was later confirmed using other methods with higher resolution (61), suggesting the original reported globular structure might be due to purification artifact, or was composed of MCM sub-complexes. The most recent report using cryo-electron microscopy (cryo-EM) revealed structural detail of the MCM2-7 complex at near-atomic resolution (92). MCM2-7 proteins form a heterohexameric complex with equal stoichiometry. As mentioned before, two hexamers are loaded at replication origins in a head-to-head conformation, the structure of which was also demonstrated in this recent report. Two hexamers are tilted when interacting with each other, with a 14° wedge between the two N-terminal planers. The axes of the two hexamers are also misaligned due to a twisted interaction between the two complexes. dsDNA goes through the central channel of both hexamers. However due to the tilt-twisted conformation, the two channels together form a kinked central channel, with a narrowed passage at the junction tightly fitting around dsDNA. MCM2 & 5 from both hexamers flank on either side of the dsDNA at this double-hexamer interface (Figure 1.3). This observation supports the hypothesis that MCM2 & 5 forms a gate on the MCM2-7 hexamer ring, where one strand of the DNA duplex is extruded when Cdc45 and GINS complex close this gap and form the CMG

holohelicase during helicase activation (61). The kinked channel formed by the double hexamers also bends and deforms dsDNA, which may provide a structural basis for origin DNA melting during helicase activation.

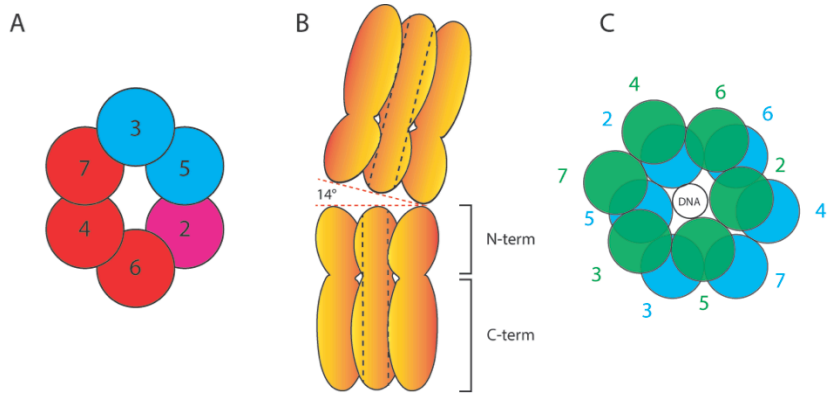


Figure 1.3 MCM2-7 complex formation. (A) The order of MCM2-7 subunits in MCM2-7 heterohexamer. Subunits with strong interaction have the same color. MCM2 have stronger interaction to MCM4,6,7 trimer than to MCM3,5 dimer. **(B)** Tilted MCM2-7 double hexamer interaction at licensed origins. A 14° wedge is between the N-term planers of the two hexamers. Dashed lines mark the central channel within each MCM2-7 hexamer, which encircles dsDNA. The tilted axes bend DNA and deform it, which is important for origin DNA melting during helicase activation. **(C)** Top view of twisted two hexamers formation around dsDNA. Tilt and twist causes an offset of the axes of the two hexamers. MCM2 & 5 from each hexamer constrict the bigger channel in each hexamer into a smaller one, tightly flanking dsDNA.

1.2.4. Function of MCM2-7 complex during eukaryotic DNA replication

1.2.4.1. Replication licensing

Chromosome was believed to have different DNA replication competency states based on the early cell fusion study. Chromatin DNA isolated from G1, but not G2 phase cells could be replicated when fused into S phase cells, suggesting chromatin bound factors determines chromatin DNA replication competency in a cell cycle regulated manner (93). The first concept of "replication licensing" came from the study on *Xenopus* eggs, in which nuclear envelope non-permeable licensing factors were found to support replication on chromosomes (94). Upon the discovery of the MCM mutants using the mini-chromosome maintenance assay, their role in DNA replication was implied. Later MCM3 became the first among MCM2-7 proteins being discovered as a *bona fide* licensing protein in *Xenopus* eggs and human cells (95), suggesting replication licensing related roles for other MCM genes as well. In conjunction with the genetics studies performed in the yeast where MCM mutants had replication initiation defects, MCMs as important licensing proteins was confirmed. Now it is clear that licensing is a major step towards DNA replication initiation, and the biochemical nature of licensing reaction has been determined as aforementioned.

1.2.4.2. MCM2-7 as DNA replication helicase

MCM as initiation factors was intuitively determined by the nature of mini-chromosome maintenance assay, in which replication initiation from a selected set of known replication origins was evaluated. A few of the MCM mutants, however, could progress into S phase and commit to substantial DNA synthesis, suggesting additional role of MCM proteins during replication elongation (81). Chromatin immunoprecipitation (ChIP) experiments performed in buddying yeast suggested that MCM proteins had similar genomic footprints as the replication

polymerases during S phase, providing physical evidence that MCM proteins travel with replication fork (96). Given that MCM proteins are AAA+ ATPases, and the biochemical architecture of the MCM2-7 heterohexamer complex resembles many known DNA helicases, MCM2-7 function as the replication helicase was proposed.

If MCM2-7 is the replication helicase, mutations in these genes should lead to immediate replication fork arrest as seen in the prokaryotes. For example, mutation in *dnaB*, the only replication helicase gene in *Escherichia coli* often resulted in immediate cessation of DNA replication (97). Eukaryotic cells, however, appeared to tolerate MCM2-7 mutations, as many MCM mutants can still support DNA synthesis, which contradicted their hypothesized role as replication helicase. This discrepancy is probably due to the fact that replication helicase in eukaryotic cell contains many different subunits. Mutation in a single one may only cause very minor change to the structure and function of the entire protein complex. The construction of MCM temperature sensitive degron (ts-degron) mutant provided the long-awaited genetic evidence to support MCM2-7 as components of the replication helicase complex. These MCM mutants can initiate replication normally, however the protein will be degraded once shifted to a non-permissive temperature. When allowed to enter S phase, each of the MCM2-7 ts-degron mutants displayed immediate replication arrest once the culture was shifted to non-permissive temperature (98). It also suggested indistinguishable biological significance of each MCM2-7 protein during replication elongation.

Contrary to the *in vivo* evidences, *in vitro* studies on helicase activity of the isolated MCM2-7 complex failed to support the helicase hypothesis. Purified MCM2-7 heterohexamer had minimal ATPase activity and failed to display any helicase activity (87). Another hexameric complex was later isolated, which contains two copies of the MCM4, 6 & 7 trimers. This

complex required DNA template that contains ssDNA and also displayed 3' to 5' ATP-dependent helicase activity (99). However, the helicase activity is very weak and non-processive. These observations were quite surprising when compared to observations made in archaeal. The single MCM homologue forms homohexamer complex, which displays high level of ATPase and helicase activity with good processivity (100). The puzzle was not solved until the discovery of Cdc45 and GINS complex as helicase complex components. Reconstitution of the entire CMG complex using recombinant methods enhanced ATPase activity and processivity of the MCM2-7 complex (55). Thus, it is clear that MCM2-7 in association with other auxiliary subunits form the functional helicase complex *in vitro* and *in vivo*. The additional requirement for these subunits in eukaryotes must happened later during evolution, since they are not required by archaeal MCM helicase.

1.2.5. Other MCM genes

1.2.5.1. Mcm8 & 9

Mcm8 & 9 were not identified in the original mini-chromosome maintenance assay performed in yeast, as these two genes are absent in lower eukaryotes. Mcm8 exists in higher eukaryotes except yeast and nematodes, while only vertebrates have Mcm9. Mcm9 shares more sequence homology with Mcm8 than with Mcm2-7, suggesting it may arose from a gene duplication of Mcm8. Both Mcm8 & 9 contain the signature "MCM box" that defines the Mcm2-7 genes, yet they are only distantly related to the latters. A search for additional MCM-like genes in the human genome based on sequence homology suggest that Mcm9 may be the last "true" MCM (101).

Based on sequence similarity, Mcm8 was hypothesized to have similar roles during DNA replication as Mcm2-7, which is probably only required in higher eukaryotes. Indeed,

recombinant MCM8 protein has ATPase and helicase activity *in vitro*. However, it is probably involved in other aspects of DNA replication, as MCM8 deficiency did not affect replication licensing and elongation *in vitro* (102, 103). Consistent with this, Mcm8 deletion in mice is not lethal (104), confirming that it is dispensable for normal DNA replication. Similarly, Mcm9 is also not essential for replication and cell survival *ex vivo* and *in vivo* (104-106).

Mcm8 & 9 might participate in homologous recombination (HR) mediated DNA damage repair. *Drosophila* Mcm8 is required for meiotic crossing over (107, 108). In vertebrates, MCM8 & 9 form a dimeric complex not essential for replication initiation (103). However, Mcm8 & 9 deletion causes hypersensitivity to DNA crosslinking reagent (105). The MCM8-9 complex helps to recruit Rad51 to the site of DNA damage (109) and promote MRN (MRE11-RAD50-NBS1) complex mediated DNA resection to facilitate DNA double-strand breaks (DSBs) repair by HR (110). Consistent with these observations, MCM8 & 9 mutations led to meiotic defects and genomic instability in mice (104, 106) and human (111, 112). In general, Mcm8 & 9 probably evolved from the major Mcm2-7 members in higher eukaryotes to perform other biological functions, which still depends on their ATPase and/or helicase activity.

1.2.5.2. Mcm10

Mcm10 was identified in the same mini-chromosome maintenance experiment that discovered many members of the Mcm2-7 genes. Despite Mcm10 mutants share several similar phenotypes as Mcm2-7 mutants, they are not related genetically or biochemically as MCM10 lacks the "MCM box" that defines the major MCM genes (113). The exact function of MCM10 during eukaryotic DNA replication is still under debate, however, it is generally accepted that Mcm10 is essential for DNA replication, which is involved in both initiation and elongation steps.

MCM10 extensively interacts with many of the replication proteins, including members of the MCM2-7, components of replication polymerases, and CDC45 [reviewed in (114)]. It is general consensus that MCM10 is not required for replication licensing. MCM10 is only loaded onto chromatin during helicase activation and travels with replication fork. However, the exact moment when MCM10 first interacts with other replication proteins is under debate. Some reports suggested it facilitates CMG assembly by recruiting CDC45 (115), while others showed CMG formation is dispensable of MCM10 (116). It is clearer that MCM10 is required for the recruitment of lagging strand polymerases Pol α & δ (116). Together with CTF4, MCM10 helps to recruit Pol α to the site of DNA synthesis (117). Since Pol α is needed for its unique primase activity to initiated *de novo* DNA synthesis on both the leading and lagging strand, MCM10 may thus contribute to origin firing. And also, Pol α activity is needed repeatedly for Okazaki fragment synthesis on the lagging strand, thus MCM10 is essential for replication elongation. Consistent with these suggested functions, Mcm10 mutations often led to elevated DNA replication stress (114, 118). In general, MCM10 connects helicase and polymerases during both replication initiation and elongation, thus helps maintaining the integrity of replication fork.

1.2.6. "MCM paradox" and dormant origins

1.2.6.1. The observations of "MCM paradox"

MCM2-7 proteins are abundantly expressed in proliferating cells. Early study in budding yeast suggest that only ~5% of the MCM proteins produced are chromatin bound. Since MCM2-7 form a multiunit complex to perform its biological function, excessive MCM2-7 production was believed to ensure sufficient amount of the whole complex can form (84). Despite this, it was estimated that ~5-25 MCM2-7 molecules are loaded at each replication origin on average, which still vastly exceeds the number of complexes needed for origin activation (119). Similar

observations were made using *Xenopus* eggs and mammalian cells (120, 121). Paradoxically, only one copy of ORCs is present at each origin, suggesting multiple copies of MCMs are loaded and distributed in a large genomic region near the ORCs bound origins (122). However, not all of the MCM licensed origins will initiate DNA replication during S phase.

As mentioned before, MCM2-7 are core components of the replication helicase and travels with DNA replication fork. A rational deduction based on this fact is that MCM2-7 proteins should always colocalize with replicating foci in nucleus. However, immunofluorescence microscopy evidences suggested that unlike PCNA and RPA, which co-localize with newly synthesized DNA, MCM proteins bind uniformly to chromatin during S phase (123, 124). Thus, "MCM paradox" describes an overall cellular condition that: 1) MCM2-7 are excessively expressed and multiple copies of the complex are loaded onto each replication origin. 2) MCM2-7 binding sites on the chromosome cannot be used to predict genomic regions that are actively engaged in DNA replication.

1.2.6.2. Dormant origin licensing and replication stress

The observation that MCM binding does not always colocalize with actively replicating genomic regions suggests that not all of the licensed origins participate in active DNA replication. These licensed origins, which are not being used during normal DNA replication, are termed "dormant origins" and first studied in *Xenopus* eggs (125). The density of the origins scattered over the large genomic regions was evaluated using DNA fiber analysis. Nucleotide labeled newly synthesized genomic regions can be visualized using fluorescent microscopy once the chromosome is stretched into a linear structure and attached to glass slide. The average physical distance between two nucleotide labeled regions, or replication origins was measured and used as an indication of origin density. Short inter-origin distance indicates high origin

density, or *vice versa*. Origin usage was thus studied and compared between "maximal licensed" condition where normal licensing is achieved, and "minimal licensed" when licensing is inhibited by overexpression of Cdt1 inhibitor geminin. Additional replication stress led to reduced inter-origin distance comparing to normal replication condition when the genome was "maximal licensed", suggesting additional firing from licensed origins that were not usually used. The presence of these "backup" origins that are not usually used was thus proposed, which is termed "dormant origins". However, under "minimal licensed" condition, despite that the origin density was not changed comparing to "maximal licensed" condition during normal DNA replication, additional replication stress failed to reduce inter-origin distance. It suggested that sufficient licensing ensures usage of "dormant origins" under stressful DNA replication conditions (125). Similar observations were later made using cultured human cancer cells (126, 127). However, instead of inhibiting licensing reaction *per se*, expression of MCM proteins was reduced severely by RNA interference (RNAi). Consistent with the *Xenopus* study, dormant origin licensing inhibition by reduced MCM expression did not affect normal DNA replication dynamic or cell survival. However, administration of reagents that slows or stalls replication fork progression when dormant origins are insufficiently licensed caused cell death (126, 127). In addition to these *ex vivo* observations, sub-lethal dosage of replication stress treatment on *Caenorhabditis elegans* with reduced MCM expression became lethal *in vivo* (125). *Mus musculus* (mice) with severely reduced MCM2 expression developed stem/progenitor cell deficiency, genomic instability and high incidence of cancer predisposition (128). Mutations in pre-RC components can lead to licensing defects and cause Meier-Gorlin syndrome in human, which is defined by developmental defects and stem cell deficiency in certain tissues (129). However, no apparent increase in cancer predisposition is reported so far. Thus, sufficient DNA

replication licensing of dormant origins by excessively expressed MCM proteins is important for normal cellular response to DNA replication stress. It explains the second part of the "MCM paradox" where uniformed chromatin bound MCM licenses dormant origins that do not usually participate in normal DNA replication.

As mentioned before, only a small subset of the licensed origins will be selected to initiate DNA replication under normal conditions. The limiting factor within the replisome that determines origin usage appears to be Cdc45, especially in somatic proliferating mammalian cells that have an intrinsically low origin density. Cdc45 protein is expressed at low level comparing to other replication licensing factors, yet it is highly stable throughout the entire S phase, so it can be shuffled between replisomes for recycled usage. The chromatin bound MCM proteins at licensed origins compete for CDC45 association to assemble into helicase complex and initiate DNA replication (121). Since multiple copies of MCM2-7 double hexamers are loaded by repeated ORCs-CDC6 action near each replication origin (47), it is conceivable that MCM2-7 dosage can also affect their distribution over licensed origins. Origins loaded with more copies of MCM2-7 complexes may become more attractive to CDC45 association, thus will be preferentially activated to initiate replication. Consistent with this idea, early observation in budding yeast suggested that altered MCM expression would affect replication initiation from selected ARSs (119). Proper MCM dosage may have big impact on genomic integrity maintenance in higher eukaryotes by counteracting replication stress induced neoplasia. MCM2 deficiency in mice causes initiation defects in the gene-rich regions of the genome, which also coincide with the focal points of the copy number variation (CNVs) events that frequent in tumors developed in this mutant (130).

1.2.7. MCM and disease

Since Mcm2-7 are essential DNA replication genes and they are needed for replication licensing that precedes actual replication, normal Mcm2-7 expression was implied to predict proliferation potential of the cell (131). Besides using MCM2-7 expression to identify the proliferative cell population in the normal tissues, ectopic MCM expression is used as histological marker to predict the prognosis of certain diseases, especially malignant transformation and cancer with even higher specificity and sensitivity than conventional cell proliferation markers [reviewed in (132)]. Individual or multiple MCM2-7 proteins were found to be highly expressed in pre-malignant tissues. A list of ectopic MCM expression used as cancer diagnostic markers can be found in Table 1.1. Increased MCM expression may be due to the amplification of the genomic loci harboring MCM genes, or altered tumor suppressor gene expression (133).

Besides increased MCM expression, mutations in the MCM genes can also lead to diseases. Point mutations in Mcm4 cause replication helicase defects and thus genomic instability and cancer predisposition in mice (134, 135). MCM4 mutation due to partial truncation of the protein was found in human, which causes growth retardation and genomic instability (136). An examination of the TCGA (The Cancer Genome Atlas) dataset revealed that mutations in the MCM2-7 genes could be found in various cancers. The frequency of these mutations depends on tumor type, however most of them are predicted to have deleterious effect on protein function. Finally, a great proportion of the cancers also contains heterozygous loss of MCM2-7 expression, which may also contribute to tumorigenesis (130).

Table 1.1 Ectopic MCM expression as cancer diagnostic markers

Gene name	Tumor type	Reference
Mcm2	Colon cancer, meningiomas, breast carcinoma, medulloblastoma, gastric adenocarcinoma, cervical cancer, oligodendrioma, renal cell carcinoma, esophageal squamous cell carcinoma, laryngeal carcinoma, breast cancer, large B cell lymphoma, oral cancer, ovarian cancer, and gastric cancer	(137-143) (144-151)
Mcm3	Meningiomas, medulloblastoma, papillary thyroid carcinoma	(138); (140, 152)
Mcm4	Meningiomas, nonsmall cell lung cancer	(138, 153)
Mcm5	Colon cancer, meningiomas, gastric adenocarcinoma, ovarian cancer, prostate cancer	(137, 138); (141, 150, 154)
Mcm6	Meningiomas, hepatocellular carcinoma	(138, 155)
Mcm7	Prostate cancer, meningiomas, medulloblastoma, esophageal squamous cell carcinoma, colorectal cancer, small lung adenocarcinoma, oral squamous cell carcinoma	(138, 156); (140, 157-160)

1.3. DNA damage responses during DNA replication

1.3.1. Overview

The integrity of the genomic information must be maintained throughout the cell cycle. The nature of DNA replication process requires opening up the DNA duplex, which poses great threat to the genome, as separated and exposed ssDNA is DNA damage prone. Many environmental factors originated from both the exogenous and endogenous sources can attack this vulnerable intermediate DNA structure, resulting in DNA replication errors. Besides, all the replisomes scattered in the vast eukaryotic genome must function coordinately so the replication related resources (i.e. dNTPs, components of the replisome, etc.) could be distributed rationally to ensure faithful replication in every part of the genome.

Replisome is a whole genome traveller. During its trip around the entire genome, it encounters many chromosome structures, though expected, yet still have the potential to inhibit replisome progression. Such structures include genomic regions that contain nucleotide repeats, DNA lesions, protein-DNA compounds, etc. Although DNA replication process defines S phase, it is not the only biological process happening during this period of the cell cycle. DNA replication machinery also encounters many other protein complexes engaged in diversified biological functions, such as transcription and chromatin remodeling.

Eukaryotic cell evolved multileveled sensor-effector cascade to detect and resolve DNA damage, including those arise during DNA replication. These factors are members of the DNA damage response (DDR) mechanism (Figure 1.4). DDR is separately headed by three distinct protein kinases, Ataxia-telangiectasia Mutated (ATM), ATM and Rad3-related (ATR), and DNA-dependent protein kinase (DNA-PK). All of these kinases are members of the phosphoinositide-3-like kinase kinase (PIKK) family, which preferentially phosphorylate serines

and threonines followed by a glutamine (S/TQ) in hundreds of protein substrates. ATM, ATR and DNA-PK recognize different types of DNA damage structures: DSBs or DNA ends usually activate ATM and/or DNA-PK, while ATR mainly responses to ssDNA coated with RPA. A simplified view of DDR pathways is shown in Figure 1.4. Since these different DNA substrates are interchangeable due to DNA damage processing and DDR kinases sometimes phosphorylate same substrates, there is plenty of crosstalk between the DDR pathways. The collective efforts of DDR during DNA replication composed a specialized mechanism termed “intra S-phase checkpoint” or “DNA replication checkpoint”. DNA replication checkpoint is responsible for detecting any DNA lesions associated with DNA replication process, or sub-optimal cellular conditions that potentially hinder replication. A “licensing checkpoint” was also proposed, which evaluates whether licensing is sufficient for replication initiation during G1 phase.

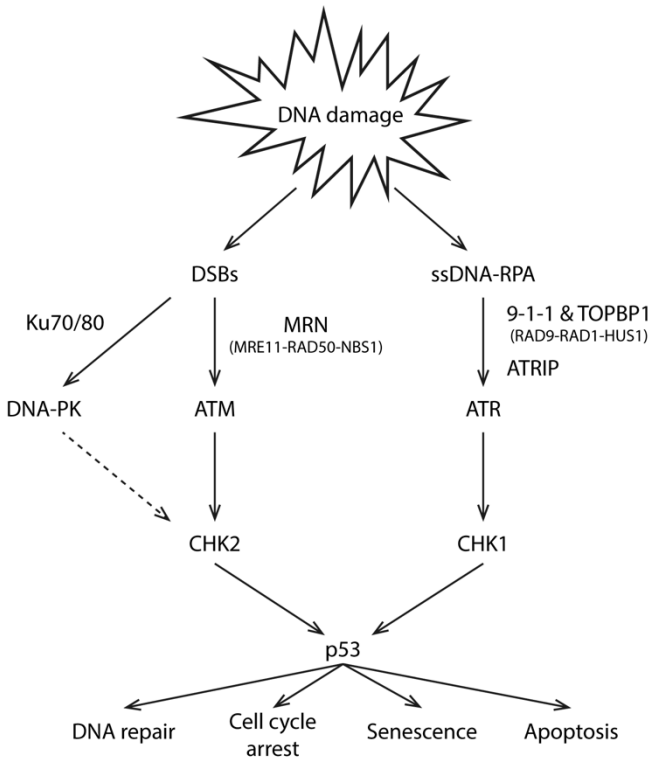


Figure 1.4 Simplified DDR pathways. DNA damage in different forms can be separately recognized by DDR kinases ATM, ATR and DNA-PK. ATM and DNA-PK are mainly responsible for DSBs recognition. DNA replication stress (RS) stalls replication fork progression and leads to excessive ssDNA exposure at stalled forks, which is bound by RPA. ATR is mainly responsible for ssDNA-RPA detection. Activated DDR kinases phosphorylate hundreds of downstream substrates. CHK1 and CHK2 are major downstream substrate/effector kinases following ATR and ATM activation respectively, which further convert on central tumor suppressor p53 to activate its function. DDR activation also targets many other substrates, together with p53 to initiate downstream effects such as transient or terminal (senescence) cell cycle arrest, DNA damage repair and apoptosis.

1.3.2. Sources of DNA replication stress (RS)

DNA replication stress (RS) is defined as any cellular conditions that inhibit the replication process, which may lead to replication fork stalling and collapse. RS originates from many endogenous and exogenous sources.

1.3.2.1. DNA lesions and aberrant DNA structure

Various forms of nucleotide repeats and GC-rich DNA sequences can cause RS. Due to the unparalleled nature of double strand DNA replication, ssDNA generated by helicase unwinding may accumulate, especially on the lagging strand before it can be duplicated promptly, which tends to adopt a more stable secondary structure. These structures may directly block, or generate topological tension to stall replication polymerases thus causing RS (161).

DNA lesions can arise from endogenous and exogenous sources that cause RS. Common exogenous environmental factors such as UV light and ionizing radiation (IR) can change DNA bases or damage DNA, which make them not suitable for replication. Reactive oxygen species (ROS) generated by normal or irregular cellular metabolism can also alter nucleotide property that leads to DNA lesions (162). Reactive aldehyde produced through alcohol metabolism and histone demethylation is another endogenous metabolic byproduct that can cause DNA lesions by inducing DNA inter-strand crosslinks (ICLs) that inhibit replication helicase unwinding thus RS (163). Ribonucleotides (rNTPs) can be misincorporated into genomic DNA during previous round of DNA replication. If not detected and replaced properly, these rNTPs can also stall replication forks (163).

1.3.2.2. DNA replication/transcription conflict

The conflict between DNA replication and transcriptional machinery is another major sources of RS. In addition to the common fragile sites (CFSs), which often replicate late during S

phase (discussed later), early replication fragile sites (ERFSs) were identified in B lymphocytes (164). ERFSs coincide with the gene clusters that are highly expressed and replicated early during S phase, which is consistent with the common observations that high transcriptional activity correlates with early replication timing. ERFSs are enriched for repetitive elements and CpG dinucleotides sequence, which have a tendency to break spontaneously and produce DSBs even during normal DNA replication. Additional RS or S-phase checkpoint inhibition accelerates ERFSs formation, which contributes to genomic instability and mutations having tumorigenesis potential. RS resulted from replication/transcription conflict is not limited to ERFSs in B lymphocytes, but is a common feature in all the replicating cells (163). This type of RS may not only arise when replication and transcriptional machinery collide physically, but can also happen due to increased topological constraints generated by the two partners (161).

Transcription can also produce R-loop that inhibits replication. RNA transcript that failed to detach from the DNA template can form a RNA:DNA hybrid, or R-loop. This three-stranded structure must be resolved by specialized topoisomerases and helicases before the arrival of replisome, otherwise leading to replication fork stalling and genome fragility (165).

1.3.2.3. Common fragile sites (CFSs)

Common fragile sites (CFSs) are cytogenetically distinguishable genomic regions that are susceptible to DNA breakages in the presence of exogenous RS such as DNA polymerase inhibitor aphidicolin (APH) or ribonucleotide reductase (RNR) inhibitor hydroxyurea (HU). Detection of CFSs in cells usually relies on fluorescent *in-situ* hybridization (FISH). CFSs lacks DNA sequence signatures that are usually associated with genome fragility such as repetitive DNA sequences, but they are mitotically stable and cell type specific, suggesting they are determined by higher order chromosome structures (166). One common feature of the CFSs is

that they usually accommodate large genes (>300kb long), which become recurrent sites of deletions found in cancer cells (167). Whether the replication/transcription conflict within these large genes contributes to CFS fragility is still under debate (167, 168). However, it is clear that CFSs are often devoid of replication initiation events, thus requires long-travelling forks to complete replication (166). Given that CFSs are late-replicated genomic regions, lack of replication initiation or dormant origins within CFS regions may lead to incomplete replication, which is subject to DNA breakages and loss of genomic information.

1.3.2.4. Deregulation of replication resources

Resources needed for normal DNA replication must be distributed to all of the active replication sites. These recourses include deoxynucleotides (dNTPs), histone and histone chaperones and protein components of the replisome (163).

Sufficient dNTP is needed for DNA synthesis. In fact, nucleotide deficiency might promote genomic instability during early stage of malignant transformation (169). Depletion of dNTP pool can result from S-phase checkpoint defects that lead to unregulated origins firing (170). Ectopic expression of oncogenes such as c-Myc, Ras and Cyclin E can also abnormally accelerate S phase progression by increasing replication initiation rate, leading to overconsumption of nucleotides. Oncogene induced RS can also result from abruptly increased transcriptional activity in the cell, which causes more replication/transcription conflicts.

1.3.3. Direct detection of stalled replication forks by ATR

Among ATM, ATR and DNA-PK, ATR is mainly responsible for recognizing ssDNA-RPA structures that are usually found at stalled replication forks, thus serving as the major intra S-phase checkpoint kinase. Many types of DNA damage such as UV and DNA intra-strand crosslinks, or dNTPs deficiency can stall replication polymerases but not helicase, which causes

uncoupled helicase and polymerase progression that leads to excessive helicase unwinding generated ssDNA (171). RPA binds and stabilizes this ssDNA stretch, which further recruits ATR to the stalled replication forks through its obligatory activation partner ATRIP (ATR interaction protein) (172-174). ATR recruitment by ATRIP, however, is not sufficient to activate ATR. Pol α mediated primer synthesis is further required at stalled replication forks (171). The RNA: DNA hybrid primer synthesized by Pol α creates a primer: template junction, which recruits the RAD17-RFC2-5 complex serving as a clamp loader to assemble the 9-1-1 (RAD9-RAD1- HUS1) complex onto DNA at the site of stalled replication fork. The 9-1-1 clamp interacts with TOPBP1 (DNA topoisomerase 2-binding protein 1), a critical activator of ATR. Thus the independent recruitment of ATR-ATRIP and 9-1-1-TOPBP1 to the stalled replication forks follows the "two-men rule" to activate ATR, which avoids aberrant checkpoint activation in the absence of replication fork stalling (172-174). Since most of the structures found at stalled replication forks also exist at normal forks, such as ssDNA-RPA and Pol α produced primer: template junction, the prolonged exposure of these structures due to polymerase stalling may decisively distinguish stalled vs. normal forks, thus results in ATR activation. However, helicase-polymerase uncoupling independent ssDNA-RPA formation can also activate ATR. For example, DNA ICLs stalls helicase progression. Fanconi Anemia (FA) pathway mediated ICLs processing produces ssDNA-RPA, which can also mark the site of stalled replication fork (161).

1.3.4. Function of ATM pathway during DNA replication

DDR cascade headed by ATM is also involved in recognition of DNA damage during DNA replication. ATM function is very likely secondary when stalled replication forks eventually collapse and form DSBs, which can be repaired by two different pathways: the error-prone non-homologous end joining (NHEJ) and the error-free homologous recombination (HR).

DSBs can be recognized and bound by Ku70/80 heterodimer, which further recruits DNA-PK to promote NHEJ (175). DSBs formed during S or G2 phase have the opportunity to use replicated sister chromatids as template to conduct HR-based repair. In this case, DSBs can also be recognized by MRN (MRE11-RAD50-NBS1) complex, which further recruits ATM to the site of DNA damage. ATM autophosphorylation and monomerization activates its kinase activity, which phosphorylate the histone H2A variant H2AX (forms phosphor-H2AX or γH2AX). γH2AX serves as a marker of DSBs, which can spread a few kilobases away from the actual DNA damage site. MDC1, which binds to both γH2AX and ATM helps to further recruit ATM and amplify the γH2AX signal (176). γH2AX formation can induce chromosome remodeling to increase DNA accessibility, or directly promote the recruitment of repair proteins to the chromosome and thus facilitate DNA damage repair. A recent method developed to study protein dynamics at active and stalled replication forks demonstrated γH2AX formation adjacent to the replication forks shortly after stalling. This H2AX phosphorylation appeared to depend on ATR, which happens before fork collapse and DSBs formation (177). Thus ATR may also phosphorylate H2AX to recruit DNA repair proteins to the stalled replication forks immediately.

1.3.5. Consequences of DNA replication checkpoint activation

The major affects of intra S-phase checkpoint activation include stalled replication fork integrity maintenance to facilitate fork restart through various mechanisms, regulation of origin firing and cell cycle progression. It also facilitate DNA damage repair to resolve DNA lesions at or behind replication forks. Besides these transient and localized reactions, persistent checkpoint activation diffuses through the nucleus and contributes to cell fate decisions such as senescence and apoptosis.

1.3.5.1. Stabilize stalled replication forks

It is believed that proper checkpoint function during DNA replication is important for stalled replication fork maintenance in the presence of RS (178). Checkpoint mutation in yeast caused collapse of stalled replication fork thus DSBs. Despite other reports suggested that replisome proteins maintain their interactions with each other and DNA at stalled replication forks in yeast (179) (174), prolonged exposure to replication inhibition led to dissociation of replication proteins in mammalian cells (177). ATR phosphorylates many components of the replisome, which may contribute to stalled replication fork stabilization, although how it is achieved biochemically is unclear (172). Stabilized stalled replication forks can be restarted later through many different mechanisms such as fork remodeling and Holliday junction processing (180).

1.3.5.2. Regulate origin firing

The function of intra S-phase checkpoint during RS induced origin firing is well studied. More recent studies suggest that it both activates and inhibits origin initiation paradoxically, depending on their relative physical relationship to the stalled replication forks.

Early studies in yeast suggest that DNA replication checkpoint mutation led to shortened S phase in the presence of RS due to increased initiation frequency (181). The discovery of dormant origins explains the source of these additional firing events under RS (125-127). ATR activation phosphorylates secondary kinase CHK1, one of the major downstream substrate of ATR to prohibit additional origin firing by modulating CDK and DDK activity needed for helicase activation and replication initiation globally (182). On the contrary, ATR can also promote origin firing locally. Regional increased ATR activity due to replication fork stalling phosphorylates MCM2, which is required for PLK1 (Polo-like kinase 1) loading onto chromatin.

PLK1 association helps to recruit CDC45 and promotes helicase activation and origin initiation (183). This is consistent with the observation that ATR-CHK1 activity inhibit origin firing in new replication domains globally, while local ATR activity cancels out CHK1 mediated global origin regulation to promote initiation events in the established replication domains (184).

Why inhibits origin firing globally while promoting it locally in the presence of RS? Actively replicated regions in eukaryotic cells usually arrange into replication domains, where the resources needed for replication are concentrated. Local ATR activity facilitates additional replication initiation from dormant origins in the close vicinity to the stalled replication forks in the established replication domains to complete replication. It minimizes the efforts to stabilize a stalled replication fork for long period of time, which often requires additional error-prone mechanisms to restart. Global replication is inhibited by diffused ATR-CHK1 activity to prevent additional assembly of new replication domains, which will compete for resources with the existing ones experiencing RS. In general, localized ATR activity promotes replication initiation to "power through" RS in the close vicinity of stalled forks, while strategically manage global replication resource distribution to counteract RS.

1.3.5.3. Delay cell cycle progression

Any DNA lesions detected by DDR during replication must be repaired during S and G2 phase before proceeding into M phase. Activated intra S-phase checkpoint can negatively regulate cell cycle progression allowing cells to have more time resolving the problems. One of the major mechanisms is to delay S phase progression, mainly by inhibiting replication initiation globally as mentioned before (185). Another major mechanism is by CHK1 inhibition of CDC25 phosphatases (186). CHK1 phosphorylates CDC25 to inhibit its phosphatase activity or its

interaction with the substrates, which is important for removing the inhibitory phosphorylation on CDKs that promotes G2/M transition.

1.3.5.4. Modulate DNA damage repair

Activated ATR phosphorylates many of the proteins involved in recombination mediated DNA damage repair. However, little is known on how these modifications would affect their activity. Many of these recombination pathways may facilitate the rescue of stalled replication forks by restarting it (180). However, irregular recombination can be mutagenic thus should be prohibited (172, 174).

1.3.6. Persistent DNA replication related damage induce senescence

RS usually induces localized DDR to resolve DNA replication defects, and transiently delays cell cycle progression. However, cultured mammalian cells terminally arrest cell cycle progression and cease to proliferate after many rounds of DNA replication *ex vivo*. This phenotype is termed senescence, which is also observed *in vivo*.

Cellular senescence was first recognized as "Hayflick limit", whereas Hayflick observed that normal human primary cells could divide a finite times before assuming an terminal cell cycle arrest state. Similar observations were made using cultured primary mammalian cells. Despite that senescent cells are viable, the senescence state is irreversible as changing of culturing conditions such as removal of contact inhibition or supplement with nutrient would not revive the cells to resume proliferation *ex vivo* (187).

Development of senescence is intrinsically related to DNA replication process. In cultured human cells, repetitive DNA replication cycles shorten telomeres below a crucial level that induces senescence (188). In contrast to human, mice have very long telomeres and constitutively express telomere maintenance genes. Disruption of telomere maintenance in mice

caused critical telomere attrition in later generations, which renders cell growth defects as observed in cultured primary human cells (189). Normal murine cells, however, also displays replicative senescence phenotype in culture independent of telomere shortening (190). Ectopic expression of oncogene induces transient hyper cellular proliferation followed by senescence (191, 192).

Senescence induction is a DNA replication related DDR. In replicative senescent human cells, chromosome ends are bound by DDR proteins, which also activate DDR pathways. Inhibition of ATR-CHK1 & ATM-CHK2 axes leads to S-phase reentry and rescues senescence-like phenotypes (193). Telomeres are incompletely replicated thus reduce in length with each replication cycle. Severe telomere attrition leads to loss of telomere-bound "sheltering" proteins that inhibit checkpoint activation (194, 195). TRF2 (telomeric repeat-binding factor 2) or POT1 (protection of telomeres 1) binds to either double stranded or single stranded telomere end and thus protect it from activating ATM, ATR respectively. Removal of telomere capping proteins exposed telomere ends and triggers DDR (195). Senescence in cultured mouse cells is independent of telomere erosion. However, oxidative stress under normal culturing conditions causes accumulated DNA damage that drives senescence (190). Oncogene expression induces hyper proliferation followed by DDR activation, which precedes senescence (187). Elevated level of RS was observed in oncogene induced hyper proliferative cells, which generated DNA damage that contribute to DDR activation and senescence (196).

Senescence serves as a barrier to malignant transformation. As aforementioned, senescence is tightly connected to DDR, which evolves in eukaryotes to prevent neoplastic malignancy. Cellular senescence markers could be detected in neoplastic tissues, or tumors after DNA damaging chemotherapy *in vivo*, suggesting senescence is induced to limit the expansion

of cells having tumorigenesis potential (197). Mouse with critically short telomere was senescence-prone, and was also resistant to skin tumorigenesis (198). On the other hand, senescence is connected to ageing and age-related defects. Senescence markers can be detected in many naturally aged but not young tissues (187, 197). Mice with genetically engineered telomere defects showed signs of stem cell deficiency and premature ageing (199, 200). Senescence can be induced in proliferative stem/progenitor cell populations in animals due to their prolonged exposure to various stresses over the long replicative lifespan. Aberrant senescence induction challenges stem/progenitor cell maintenance and leads to stem/progenitor cell depletion related developmental syndromes (201). After all, senescence is a type of DDR regulated stress response that is usually associated with DNA replication process, which terminally arrests cell cycle.

1.3.7. Licensing checkpoint

As mentioned before, DNA replication licensing is a critical step in preparation for S phase DNA replication. It mainly takes place during G1 phase of the cell cycle, when licensing proteins are expressed and loaded onto chromatin. Before committing to proliferation, eukaryotic cells pass G1 checkpoint to make the transition from G0 or quiescent state to entering the cell cycle, which is also known as start point in yeast and restriction point in higher eukaryotes. Licensing checkpoint also takes place during G1 phase, however, should not be confused with the G1 checkpoint. Licensing checkpoint applies to cells that are already committed to DNA replication. The key determinant of licensing checkpoint is sufficiency of DNA replication licensing, instead of external stimuli that promote cell fate determination.

Direct inhibition of replication licensing through geminin overexpression in cultured human primary cells led to G1 arrest, which became the first experimental evidence to support

the “licensing checkpoint” (202). Additional studies performed later managed to inhibit licensing through other methods, such as depletion of licensing proteins including ORC2, CDC6, CDT1 and MCMs (203-205). All of the reports showed similar observations that acute licensing inhibition led to G1/S transition delay. Cyclins levels and their corresponding CDKs activity responsible for G1/S transition were inhibited when insufficient licensing occurs, which indicates the connection between licensing status and cell cycle control (203-205). Tumor suppressor gene p53 and its downstream effectors appeared to function during licensing checkpoint (205), as p53 deletion in primary cells also abolished licensing inhibition induced cell cycle arrest (202). Interestingly, transformed or cancer cell lines used in these studies were insensitive to licensing inhibition and enters S phase regardlessly (202, 205).

The potential function of licensing checkpoint is yet unknown, however it may be connected to the function of dormant origins. Before committing to full on DNA replication in S phase, it appears to be reasonable that cells self-evaluate their capability of dealing with potential stressful replication conditions. Sufficient replication licensing ensures extensive MCM2-7 loading at large number of origins including dormant origins, while the latters are specially activated under RS to rescue stalled replications. Limited licensing in normal primary cells potentially compromises their ability to counteract RS during S phase, thus these cells choose to avoid premature entry into S phase before sufficient licensing can be achieved. However, transformed cells may have dysfunctional or less restrictive licensing checkpoint due to mutations in key cell cycle regulatory and/or DDR genes, which allows them to proceed into S phase with minimally licensed chromosomes. Moreover, licensing proteins are usually highly expressed in cancer cells. Origins might still be sufficiently licensed to pass licensing checkpoint under moderate to acute degree of licensing inhibition. Interestingly, the function of dormant

origins might not be fully understood yet, if the initial observations were not made using cancer cell lines (125-127).

1.4. MicroRNA mediated gene silencing

MicroRNA (miRNA) is one of the three classes of small (20-30nt) non-coding RNAs that has specialized biological functions found in eukaryotes. The others are small interference RNAs (siRNAs) and piwi-interacting RNAs (piRNAs). These separated classes differ in small RNA origins, sequence structures, biogenesis processes, protein partners involved in biogenesis and downstream effects and biological roles. The one common feature of these small non-coding RNAs is that they guide specialized protein complexes to target sequences to perform certain function. Small RNA mediated target recognition follows Watson-Crick rule of base pairing, which further determines the exact location within the target sequence. In general, small non-coding RNAs are involved in, but not limited to, gene expression regulation, chromosome remodeling, and genome defense.

1.4.1. Modes of small RNA related regulatory mechanisms

Three classes of small RNA guided regulatory mechanisms have been identified in eukaryotes. All of these mechanisms rely on small RNA (~20-30 nt) to direct protein complexes to target based on sequence homology. The core protein component for all of these three mechanisms belongs to the Argonaute protein super family. siRNAs and miRNAs associate with the AGO subclass of Argonaute protein, while piRNAs related interaction depends on PIWI proteins. siRNAs, miRNAs and piRNAs also have different biogenesis processes (206).

siRNAs and miRNAs are produced from double stranded precursors. siRNA precursors are long, linear and perfectly base paired double stranded RNA molecules, which often come from exogenous sources, while miRNA precursors are endogenously expressed and adopt a

partially matched stem-loop structure (miRNA biogenesis and miRNA induced gene silencing will be discussed in detail below) (206). siRNA precursors are processed by RNase III type double stranded RNA endonuclease Dicer in the cytoplasm, which produces the ~22bp siRNA duplex that consists of the mature siRNA and its complementary strand (207). siRNA duplex is then loaded into siRNA-induced silencing complexes (siRISC) to search for target sequence based on homologous pairing, which has the Ago protein as the core. Dicer and accessory proteins are needed for siRNA loading into siRISC complex in animals. Once associated with Ago protein, one strand of the siRNA duplex also known as the passenger strand will be discarded while the mature siRNA (guide) strand will be retained. Guide vs. passenger selection is based on the thermodynamic stability of the two siRNA strands within the duplex: the strand with less stable-paired 5' end is retained, while the other strand is sliced and ejected from the RISC complex and later being digested (208). siRISC loaded with mature siRNA is mainly involved in post-transcriptional gene silencing. Target sequence perfectly complementary pair with the siRNA within the siRISC complex will be cut by Ago protein, dissociates and then subjects to exonuclease digestion and removal. siRISC complex can also induce post-transcriptional gene silencing to repress target gene expression.

Sources of siRNA precursor are usually believed to be exogenous, such as viral or transgenic RNA used in experimental manipulations. Recently, endogenously expressed siRNAs (endo-siRNA) have also been identified (209). Various genomic loci have been discovered in flies and mice that are devoted to producing the double stranded endo-siRNA precursors. Transposable elements (TEs) are a major source of endo-siRNA precursors. siRNAs produced from such loci are important for restricting the highly mutagenic TEs in flies and mice germline, especially in female mouse oocyte. Other sources include *Cis*-natural antisense transcripts (*cis*-

NAT), which produce RNA transcripts using both DNA strands as template; pseudogene pairs that produce complementary paired non-protein coding transcripts and hairpin RNA (hpRNA) that produces long hairpin transcripts with multiple inverted repeats.

The other class of small RNA found in eukaryotes is the piRNA, which slightly diverges from siRNA and miRNA in terms of biogenesis and functions. Mature piRNAs are single stranded about 24~32nt long, which are generated from single rather than double stranded precursors as for siRNA and miRNA. The primary transcripts of piRNA precursor are produced from genomic loci known as piRNA cluster, which is enriched for repetitive elements and often coincide with the location of TEs. Although the piRNA biogenesis process remains elusive so far, it is clear that RNase III type double stranded RNA endonuclease is not required. Mature piRNAs associate with PIWI class of the Argonaute protein to repress the activity of TEs mainly in germline. piRNA functions outside of the germline have also been identified (210).

1.4.2. miRNA biogenesis

Mature miRNA of single stranded small RNA nature is produced from double stranded stem-loop precursor. The process of miRNA biogenesis slightly diverged between animals and plants. Different proteins are involved as well even among different animals. Here I will focus on miRNA biogenesis in animals, with an emphasis on that in mammals.

The first miRNA identified was *lin-4* in *Caenorhabditis elegans*, which is important for post-embryonic development. A cloned genomic fragment was able to rescue *lin-4* deficiency, suggesting the genome basis for miRNA production (211). Combined efforts of positional cloning, next-generation sequencing, with computational analysis identified the potential genomic loci where the miRNA encoding sequences inhabit. The observation that the primary transcript containing miRNA sequence is much larger than the mature miRNA suggested the

presence of miRNA precursors, and subsequent obligatory processing to produce the mature shorter RNA.

Many types of RNA transcript can give rise to miRNA precursors. Some miRNA transcription behaves like normal protein coding genes, whereas a single miRNA is produced from each transcript. Moreover, about 50% of the mammalian miRNAs reside in close genomic clusters, and thus will be co-transcribed and processed. 40% of the miRNAs are positioned within introns of protein coding genes (212). Expression patterns of intronic miRNAs match that of their host genes, suggesting they are produced from the same transcript (213). This arrangement eliminates the needs to provide separated genomic sequences such as promoter and/or enhancer to control miRNA production. A special subtype of intronic miRNAs is known as the "mirtrons", which resides in short introns. Mirtron processing is slightly different from that of the canonical miRNAs, which will be discussed later. RNA polymerase II is responsible for most of the miRNA transcription (206).

The primary miRNA-containing transcript is quite long, extending from both 5' and 3' end of the mature miRNA sequence. This transcript is known as primary or pri-miRNA. Pri-miRNA forms a structured stem-loop, with the miRNA and its complementary strand (miRNA* or passenger strand) paired to form the major part of the double stranded stem. The entire stem is about 33bp long, separating the end loop from the unpaired flanking single stranded sequences (214). In animals, the first step of miRNA processing involves the cleavage of pri-miRNA to produce precursor miRNA (pre-miRNA). It takes place in the nucleus, which is performed by the RNase III type endonuclease Drosha. However, Drosha by itself is not sufficient to bind to the double stranded stem-loop pri-miRNA, which further requires the help from DGCR8 (215). DGCR8 recognizes the ssRNA-dsRNA junction at the feet of the stem-loop, which further

directs the precise cut of pri-miRNA exactly one helical turn (~11bp) away from the junction and into the stem. Thus, DGCR8 serves as both an anchor and ruler in the pri-miRNA processing action. This first cut by Drosha leaves a short 3' overhang (2~3nt), which also defines one end of the mature miRNA. Processing of mirtron omit the requirement of Drosha during pri-miRNA cleavage as the RNA splicing machinery can produce an exact mimic of the pre-miRNA structure from miRNAs transcripts within short introns (216). However, mitrons are very rare and only a few cases have been identified in each of the genome studied. Before the pre-miRNA can be further processed, the stem-loop (~70nt) is exported into cytoplasm by Exportin-5 in a Ran-GTP dependent manner (217).

The second cleavage performed by Dicer defines the other end of miRNA and thus produces mature miRNA: miRNA* duplex. Dicer is an important RNase III type dsRNA endonuclease, which is also needed for siRNA biogenesis (206). Dicer recognize the 3' overhang on the pre-miRNA stem-loop, and cuts 2 helical turns (~22bp) away from the end into the stem, thus producing the mature miRNA that is about 22nt long (212). Like Drosha, Dicer cleavage also produces a short (2~3nt) 3'overhang. Other accessory proteins associate with Dicer during this process, which might be needed for Dicer activity *per se* or the loading of miRNA: miRNA* into miRISC (miRNA induced silencing complex).

1.4.3. Target recognition by miRNA

Target mRNA recognition by miRNA is critical to miRNA mediated gene expression regulation and thus its downstream biological consequences. In contrast to plants, miRNA target recognition in animals does not require near-perfect complementary base pairing along the entire miRNA length. Instead, the 2~7 nucleotides on the 5' end of miRNA is crucial for target selection, which is known as the "seed" sequence. Seed sequence not only defines miRNA:

target specificity, it also helps to predict the effectiveness of miRNA regulated downstream events. miRNAs sharing the same seed sequence also belong to the same miRNA family.

Watson-Crick base pairing within and closely around the seed sequence is the primary determinant of targeting specificity and efficacy (218). The order of efficacy is 8mer > > 7mer-m8 > 7mer-A1 > > 6mer > no site (Figure 1.5). The dependency of seed sequence within the full-length miRNA for pairing purpose is probably due to how miRNA is bound and presented by the AGO Argonaute effector protein. Biochemical analysis of AGO structure suggested that certain domains of this protein interact with the miRNA backbone, which expose the Watson-Crick faces of the seed sequence to facilitate its target selection and base pairing (206, 214).

Hundreds of potential mRNA targets can be predicted *in silico* for each miRNA based on seed sequence base pairing. To identify the legitimate miRNA: target interaction that has true biological significance, one practical way is to select for interaction between conserved miRNA and conserved targeting sites (219). The rationale is that such miRNA: target interactions must be under evolutionary selective pressure to maintain its biological function in many living organisms. However, it is also possible that some of the miRNA: target interactions only occur under certain circumstances in a defined cellular context, or only arises late during evolution and being specifically selected in a certain species. However, even taking into account of the conservation factors, each miRNA is believed to have many targets within the genome. Over 60% of the human protein coding genes bear 3'UTR sequences that are evolutionarily selected as miRNAs targets based on computational analysis (220), which is further supported by experimental observations that ectopic expression of a single miRNA led to widespread changes to the gene expression profile due to direct miRNA: mRNA interactions (221).

Besides seed sequence, other miRNA: mRNA interaction can also contribute to binding affinity and specificity. Additional bindings can happen within the 3' end of the miRNA and mRNA directly opposing it, which can be either "supplementary" or "compensatory". In the first scenario, an additional continuously uninterrupted base pairing between the 13~16 nucleotides of miRNA and the target mRNA significantly increase miRNA targeting efficacy that already has 5' seed pairing (218). In the second scenario, additional 3' pairing compensates for the imperfect pairing within the seed sequence. However, more base pairs are needed for effective 3' compensatory than supplementary base pairing (222). Despite of the few experimental data, comparative studies suggested 3' supplementary pairing was not under evolutionary selection to contribute to miRNA targeting (219), and the 3' compensatory pairing was also very rare (220). Thus, neither type of 3' miRNA pairing has true predictive value.

The early discovery suggested that sequence within 3'UTR of the target mRNA is responsible for miRNA-mediated gene silencing (211), which was confirmed later by numerous studies on other miRNA: mRNA pairs. However, other regions of the mRNAs including 5'UTR and open reading frame (ORF) can also be targeted by miRNA based on seed pairing. Indeed, artificially positioned miRNA targeting sites within 5'UTR and ORF were still functional (223). *In silico* scanning of miRNA seed pairing in ORF yielded substantial potential targeting sites, yet far less than that found in 3'UTR. However, 5'UTRs were barely targeted according to the same scanning (224). One possible explanation for why 3'UTR is selected is that targeting events in the 5'UTR and ORF are probably displaced by the translational machinery, making them obsolete and thus not evolutionarily selected (219). However, targeting within the ORF can additionally contribute to miRNA mediated translational inhibition. After all, 3'UTRs of the mRNA are the major targets of miRNAs.

Location within 3'UTR affects miRNA targeting. Effective targeting often occurs close to both ends of 3'UTR, but not so close to the stop codon based on genome wide miRNA targeting study (218). Middle region of 3'UTR are not favored, especially for long 3'UTRs (>1300nt). Sequence context of miRNA targeting sites is even more important than their location in 3'UTR. AU-rich regions enhance miRNA targeting. All of these observations seem to describe different nature of the miRNA: target recognition characteristics, however, a common feature is that it makes the miRNAs and their cooperating effector proteins more accessible to the target mRNA, thus making miRNA repression more effective (218, 219).

Multiple miRNA targeting within a single 3'UTR by the same or different miRNAs often functions independently or cooperatively to enhance the repression, depends on the distance between the distinct binding sites. Targeting events happen not so close to each other independently repress target expression, leading to a multiplicative effect that can substantially repress target gene expression. For example, a very subtle 25% repression due to a single targeting event can be greatly increased to ~60% ($1-0.75^3 = 0.578$) when 2 additional targeting events happen independently. Two targeting sites positioned close to each other (less than 40 but greater than 8 nt to avoid interference) cooperatively elicit a greater repression than expected for any single targeting (218, 225). This may have great implication on miRNA-mediated expression regulation under diverse cellular context. For instance, a very subtle change to the miRNA expression profile might be enough to greatly impact target gene expression due to cooperative miRNA repression under different cellular circumstances.

In general, miRNA targeting sites are evolutionarily selected to maximize gene repression efficacy under different circumstances. Most of the targeting happens within 3'UTR.

Combined action by multiple miRNAs can fine tune the expression of a single gene or shift the entire gene expression profile to comply with certain biological requirements.

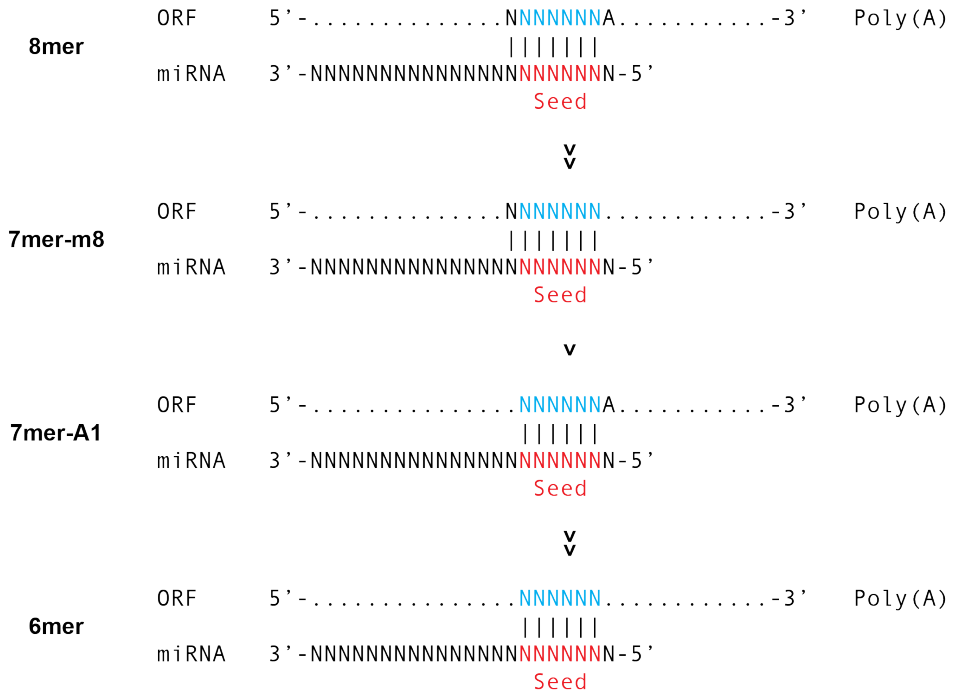


Figure 1.5 Ranking of miRNA seed sequence in targeting efficacy. The upper mRNA strand is paired with the lower miRNA strand. Seed sequence (Red) pair with sequence in the 3'UTR of the target mRNA. 8mer: perfect pairing between miRNA 2~8, and an A on the opposing mRNA strand facing position 1. 7mer-m8: perfect pairing between miRNA 2~8 and mRNA. 7mer-A1: perfect pairing between miRNA 2~7 (seed) and an A on the opposing mRNA strand facing position 1. 6mer: only perfect base pairing between seed and target mRNA.

1.4.4. miRNA mediated gene silencing

Dicer processed mature miRNA: miRNA* duplex associates with AGO protein in the miRISC (miRNA-induced silencing complex) to silencing target gene expression post-transcriptionally. Similar to siRNA, guide/passenger miRNA strand selection is based on the relative thermodynamic stability of the 5' end, whereas the strand with less stably paired 5' terminus will be retained as guide strand. However, unlike siRNA, the passenger strand will not be "sliced" by AGO protein in RISC nor later digested. Thus, a substantial amount of miRNA* strands can be detected based on small RNA next generation sequencing data. The siRNA-like removal of the passenger strand perhaps depends on perfect pairing within the siRNA: siRNA* duplex, which is a suitable substrate for the AGO slicing activity (226). miRNA: miRNA* pairs failed to comply with this criterion may retain the miRNA* strand outside of RISC.

There are different modes of miRISC induced target gene silencing, including miRNA induced mRNA destabilization and degradation, and translational inhibition. Unlike the canonical siRNA or miRNA induced gene silencing in plant, AGO induced slicing of the target mRNA and its consequent degradation is not the primary mode of suppression in animals (206). However, by default, the miRNA guided miRISC complex conduct translational inhibition on the target mRNA. Translation initiation in eukaryotic cells depends on the stabilization of mRNA transcript and its association with the translational machinery. The 40S ribosomal subunit components bind to the m^7GpppN cap at the 5' end of the mRNA, while polyadenylate-binding protein 1 (PABP1) binds to the 3' polyadenylated tail and interact with the 5' cap binding proteins to facilitate mRNA circulation that greatly enhances translational activity. Evidences of how miRISC represses translation have been found in many stages of the mRNA translation. In human, miRISC is proposed to compete against components of the 40S translation initiation

complex for m^7GpppN cap binding, thus prevents 40S pre-initiation complex formation (227). A structural domain in human AGO protein was identified that resembles the m^7GpppN cap-binding cavity in the 40S ribosomal subunit (228). Human AGO protein was also found to interact with eIF6, which is essential for the maturation of 60S ribosomal subunit and its subsequent association with the 40S subunit to form the functional ribosome at translation start codon (229). Translation can also be inhibited through deadenylation of the target mRNA. GW182 is an essential component of the miRISC complex, which interacts with the deadenylase to remove the 3' polyadenylated tail on miRNA targeted mRNA (230). Loss of polyadenylation represses PABP1 binding and thus the circulation of mRNA required for effective translation. mRNAs without sufficient length of 3' poly(A) tail are also subject to decay (231). miRISC can also repress translation elongation as it dissociates ribosome association with the mRNA (232).

miRNA can induce target mRNA degradation. Indeed, components involved in miRNA induced target mRNA destabilization through 5' decapping and/or 3' deadenylation, and the subsequent exonucleolytic degradation and decay have been identified (230, 233). However, unlike the canonical siRNA induced gene silencing that depends on the cleavage of the target sequence, the nature of only partially paired miRNA: mRNA inhibits the slicing activity of AGO protein in miRISC. Thus for the most of miRNAs: target pairs, translational inhibition is preferred. The observed miRNA-dependent mRNA degradation may be a secondary effect due to inhibited translation, during which deadenylation apparently plays pivotal role.

There is no consensus on which mechanism governs miRNA-induced gene silencing so far. For there are many proposed mechanisms, many more reports suggested the opposite under their own experimental conditions and setups. A combination of some or all of the proposed

mechanism and yet unknown ones may contribute to the overall gene silencing effect for each of the miRNA: mRNA pair.

miRNA and its repressed mRNA transcript, and protein components of the RISC are enriched in the microscopic foci such as P or GW-bodies, and stress granules (SGs), suggesting miRNA induced gene silencing takes place in specialized subcellular compartments. P-bodies and SGs are known cytoplasmic locations where untranslated mRNA transcripts are stored and degraded (234, 235). However, quantitative analysis suggested that only a small proportion of the miRNA and repressed mRNA targets localize to P-bodies, suggesting P-bodies formation is not necessary for miRNA induced gene silencing (236). Indeed, disruption of P-bodies formation *per se* has no effect on miRNA function (237). Only minority proportion of the expressed AGO protein is found in the P-bodies (236), while the most fractionates with ER and Golgi complexes (233). These observations suggest that miRNA mediated gene silencing is more likely to happen in submicroscopic cytoplasmic foci, which may accumulate in P-bodies. P-bodies formation is a consequence, rather than the cause of miRNA-mediated translational repression and mRNA degradation.

1.5. Summary

Mini-chromosome maintenance (MCM) genes are essential eukaryotic genes. Six major members of the MCM family, namely MCM2-7 physically interact with each other to assemble into MCM2-7 heterohexameric complex and perform DNA replication related functions. The major functions are twofold: first, by interacting with cofactors such as CDC45 and GINS complex, the entire CMG (CDC45/MCM2-7/GINS) complex possesses ATPase and helicase activity, which is responsible for the unwinding of double helical DNA during DNA replication. Previous observations have demonstrated that stable MCM2-7 complex formation and its

association with the replicating DNA is essential for normal helicase function and the maintenance of replication fork integrity. MCM2-7 complex should also be stabilized at stalled replication structure to enable fork restart mechanism. Secondly, MCM2-7 complex license the genome at discrete genomic loci known as replication origins. Replication initiation from the MCM2-7 complex licensed origins involves the incorporation of MCM2-7 complex into CMG complex, which melts double stranded DNA at replication origins to further provide room for replication fork assembly.

To ensure the completeness of eukaryotic DNA replication process, replication forks travel through every part of the eukaryotic genome. Conditions that can potentially inhibit replication fork progression pose great challenges to the maintenance of genome integrity, as it may lead to under-replication of the genomic information. Stalled replication forks can also collapse, which produces DNA double strand breaks (DSBs). When not properly repaired, DSBs can be the most detrimental form of DNA damage, which serves as a source for genome rearrangement and genomic instability (GIN). These conditions are recognized as replication stress (RS), which has been recognized as a primary source to tumorigenesis. Normally RS induced replication inhibition, and the DNA damage associated with it can activate DNA damage response (DDR), a collective cellular network that detects different forms of DNA damage while further transmit and amplify the signal to trigger the action of effector proteins that help to resolve the genomic defects. Part of the DDR related function of genome integrity maintenance during DNA replication is the regulation on dormant origin firing, which are licensed by excessively expressed MCM2-7 proteins while not being used during normal DNA replication. *De novo* replication initiated from dormant origins can converge on the stabilized stalled replication structure and complete replication process. Hence, excessive MCM2-7 production

and sufficient licensing of replication origins, including the normally dormant ones is recognized as one of the major cellular mechanisms that ensure the completeness of replication process, especially under conditions that are sub-optimal for normal DNA replication. Consequently, it is believed that any perturbation to the normally high MCM2-7 expression can prevent sufficient genome licensing, thus compromising normal cellular response to RS and leads to GIN and disease predisposition, especially cancer in mammals.

Based on our recent studies, it was implied that cells could tolerate moderate MCM2-7 expression loss without sabotaging normal DNA replication and cell proliferation, which leads to the impression that conventional explanation on excessive MCM2-7 expression is overrated. However, it does not rule out the possibility that MCM2-7 expression might be actively regulated in the cells, especially those experiencing elevated level of RS. This MCM2-7 regulation may reflect the collective RS and the associated DNA damage experienced by the proliferating cells, which further dictate their overall replicative lifespan. Hence, this dissertation aims to address the following questions: whether MCM2-7 expression can be regulated in the proliferating cells, especially under the influence of RS? How does MCM2-7 regulation occur? Since dormant origin usage depends on normal DDR function, which relies on excessive MCM2-7 expression in the first place, if DDR is involved in MCM2-7 expression regulation? If the MCM2-7 expression regulation does exist, then what is the potential biological function of it?

CHAPTER 2

CHRONIC DNA REPLICATION STRESS INDUCES MCM2-7 REDUCTION AND CELLULAR SENESCENCE[1]

Gongshi Bai[1, 2] and John C. Schimenti[1, 2, 3]

[1] Cornell University, Dept. of Molecular Biology & Genetics, Ithaca, NY 14853, USA

[2] Cornell University, Dept. of Biomedical Sciences, Ithaca, NY 14853, USA

[3] Cornell University Center for Vertebrate Genomics, Ithaca, NY 14853, USA

2.1. Abstract

Circumstances that compromise efficient DNA replication, such as disruptions to replication fork progression, cause a state known as DNA replication stress (RS). Whereas normally proliferating cells experience low levels of RS, excessive RS from intrinsic or extrinsic sources can trigger cell cycle arrest and senescence. Here, we report that a key driver of RS-induced senescence is active downregulation of the Minichromosome Maintenance 2-7 (MCM2-7) factors that are essential for replication origin licensing and which constitute the replicative helicase core. Proliferating cells produce high levels of MCM2-7 that enable formation of dormant origins that can be activated in response to acute, experimentally-induced RS. However, little is known about how physiological RS levels impact MCM2-7 regulation. We found that chronic exposure of primary mouse embryonic fibroblasts (MEFs) to either genetically-encoded or environmentally-induced RS triggered gradual MCM2-7 repression in a TRP53-dependent manner, which is followed by inhibition of replication and senescence that could be accelerated by MCM haploinsufficiency.

---

[1] Work presented in this chapter has been submitted for publication. Authors' contributions: G.B. & J.C.S conceived & designed the research; G.B. performed experiments and analyzed data; G.B. & J.C.S wrote the paper.

2.2. Introduction

In preparation for DNA replication, "licensing" of replication origins occurs during late M to early G1 phase (7, 94). These replication origins are selected and bound by the origin recognition complex (ORC) (238). ORCs further recruit CDC6 and CDT1 to eventually load the MCM2-7 heterohexameric complex onto replication origins, thus forming pre-replication complexes (pre-RCs) (5). Pre-RC formation is tightly regulated so origin licensing can only occur before, and not during, S phase to prevent re-replication of genomic regions (1). Chromatin becomes replication-competent after MCM2-7 loading. Later, during S phase, replication machinery assembly is initiated at selected licensed origins with the formation of Cdc45/MCM2-7/GINS (CMG) replicative helicase complex, of which MCM2-7 is the catalytic core (239, 240). Stable MCM2-7 chromatin association is required for uninterrupted replication fork progression and restart after stalling (98, 241, 242). MCM2-7 is the sole complex present in both the pre-RCs and the active replisome, making it a nexus of DNA replication control.

The genome is vulnerable to exogenous and endogenous genotoxic stresses during DNA replication, which can lead to replication fork stalling (163). Stalled replisomes must be stabilized to enable restart or displacement by converging replication forks to ensure complete and faithful DNA replication. Otherwise, mutations, genomic instability, and ultimately neoplasia can occur (243). Numerous mechanisms exist to promote error-free replication under stressful conditions (244). One of the mechanisms is utilization of dormant replication origins (163). Most growing cells produce abundant amounts of MCM2-7 proteins that license large numbers of replication origins, but only a small proportion of these are used and they are sufficient to accomplish whole genome replication. This role of dormant origins in responding to RS was revealed in experiments where licensing was severely inhibited in cultured cancer cells

via knockdown of MCM levels. While such cells can sustain limited proliferation under unchallenged conditions, the reduction of dormant origins renders them sensitive to additional RS (125-127, 245). Thus, abundant MCM production ensures adequate licensing of the dormant replication origins that serves as 'backups' and can be activated in response to stalled or collapsed replication forks and ensures completeness of DNA replication (246). Inhibition of licensing in primary cells causes cell cycle arrest in G1 phase, leading to the proposed existence of a "licensing checkpoint" that prevents DNA replication under sub-optimal licensing conditions (202, 247). Thus, the physiological relevance of severe experimental conditions in transformed cell lines is unclear, and more importantly, little is known about endogenous MCM2-7 regulation in response to RS.

Another major mechanism that protects the genome during replication is the DNA damage response (DDR), components of which detect replication-associated lesions or cellular conditions that impair DNA replication. In addition to directly interacting with MCM2-7 subunits to stabilizing stalled replisomes, the DDR regulates cell cycle progression in response to RS, such as inducing senescence (187, 197, 248). Central to this mechanism is the tumor suppressor gene *Trp53*. Genotoxic stress such as RS activates DDR that converges on TRP53 activation, which also plays pivotal role during senescence induction in mice and human.

As mentioned earlier, most studies of how RS impacts cell growth and the DDR involve treatment of cell lines with exogenous agents that hinder DNA replication. One model of intrinsic RS is the *Mcm4*^{Chaos3} mutation in mice. This allele encodes a single amino acid change (Phe345Ile) that causes high levels of genomic instability and cancer susceptibility (135). The "*Chaos3*" mutant cells also have a 40~60% decrease in MCM2-7 expression, leading to the conclusion that the associated phenotypes were primarily attributable to insufficient licensing of

dormant replication origins (135, 249-251). This view is supported by similar cancer predisposition phenoptypes in mice that are hypomorphic for *Mcm2*, and which also show premature aging and stem cell defects in certain cell lineages (128, 252).

Here, we compare the consequence of endogenous, low level RS in *Chaos3* cells to chemically-induced RS in WT and *Chaos3* cells. Our results indicate that RS induces TRP53-dependent MCM2-7 downregulation, which precedes loss of DNA replication potential in primary cells that eventually becomes senescent. Our observations reveal that active MCM2-7 regulation is a key aspect of senescence induction when cells are exposed to chronic low level RS, and is important for safeguarding organisms from cells that undergo potentially deleterious RS-induced genomic alterations.

2.3. Materials and methods

MEF preparation and proliferation assay

Primary MEFs were isolated from 12.5~14.5 dpc C3HeB/FeJ (C3H) strain embryos in which organs were removed and the remainder was lightly homogenized to make a cell suspension. Cells were cultured in DMEM with 10% FBS and 100 units/ml penicillin-streptomycin. The initial plating of embryonic cells is designated passage 0 (P0). For cell proliferation assays and general MEF maintenance, 1 x 10^6 cells were seeded in 100 mm tissue culture dishes and maintained under either standard conditions (37° C, 5% CO_2 and atmospheric O_2) or low oxygen culture (37° C, 5% CO_2 and 5% O_2 level) in parallel for 3~4 days between passages. Upon trypsinization for passage, cell numbers were counted using hemocytometer.

HU treatment on primary MEF

Early passaged primary WT or mutant MEFs were plated at a density equivalent to 1 x 10^6 cells/100 mm tissue culture dishes. After cell attachment, fresh media with or without HU

treatment was added to the control and treated cells in parallel, and cells are cultured under standard conditions. Fresh media containing HU treatment was changed every 24h of incubation.

Senescence-associated β-galactosidase (SA-β-gal) activity staining

SA-β-gal staining of cultured cells was performed as described (253). Control and treated cells fixed on coverslips are incubated in SA-β-gal staining solution in humidified 37° C chamber without CO_2 for at least 16h. To facilitate the counting of SA-β-gal positive cells, nuclei were counterstained with Hoechst 33342 before mounting the coverslips. The slides were examined using light and fluorescent microscopy.

Quantitative RT-PCR (qRT-PCR)

Total RNA was extracted from cultured cells using the EZNA total RNA kit (Omega). cDNA was synthesized from 1ug of total RNA using the iScript cDNA synthesis kit (Bio-Rad) and the supplied oligo-dT primer. qPCR reactions were performed as described (249). PCR amplification and real-time detection was performed with a Bio-Rad CFX96 Real-Time system and data analysis was performed with the Bio-Rad CFX Manager software (Bio-Rad). Relative gene expression was calculated using the ΔΔCq method with β-actin as endogenous control. A technical replicate was performed on each sample. List of qPCR primers can be found in Table 2.1.

Western blot analysis

Protein samples concentrations were determined with a BCA Protein Assay Kit (Thermo Scientific). Equal amounts of protein were loaded onto SDS-PAGE gels. Western blot analysis was performed as previously described (249, 250). Proteins were electrotransferred onto PVDF membranes (Millipore). Chemiluminescence was performed using the Luminata HRP substrate (Millipore). Bands were detected either by exposure of the probed membranes to X-ray film or

by scanning with a Bio-Rad Universal Hood II running Image Lab software (Bio-Rad). Western blot quantification was performed using ImageJ software.

Antibodies used were as follows: MCM2: ab108935 (Abcam); MCM3: 4012 (Cell Signaling); MCM6: sc-9843 (Santa Cruz Biotechnology); MCM7: ab2360 (Abcam); MCM7: 3735 (Cell Signaling); PCNA: P8825 (Sigma); total p53: 9282 (Cell Signaling); and β-actin: A1978 (Sigma).

EdU incorporation assay

Cells grown on coverslips were pulse labeled with 10μM EdU for 30min. Formaldehyde was added directly to the culture to a final concentration of 1% for 10min at room temperature (RT). After 3 washes with PBS, cells were permeablized on ice with 0.3% Triton X-100 in PBS for 15 min., followed by 3 washes in PBS containing 1% BSA. The "Click" reaction staining was performed by placing the cells in 10mM (+)-sodium-L-ascorbate, 0.1mM 6-Caboxyfluorescein-TEG azide and 2mM $CuSO_4$ cocktail for 30 min at RT. After PBS washes, nuclei were counterstained with Hoechst 33342. Coverslips were mounted using ProLong Gold antifade reagent (Invitrogen) before examination by fluorescence microscopy.

Flow cytometry and cell cycle profiling

Cells were trypsinized into a single cell suspension and fixed in 70% ice-cold ethanol overnight. They were stained for DNA content with 40μg/mL propidium iodide and 20μg/mL RNaseA for 30min at room temperature. Flow cytometry was performed on a BD Bioscience LSR II instrument. Stained cells were excited with a 488nm laser, and a 575/26 filter was applied for data collection. The percentages of cells in each cell cycle compartment were determined using ModFit LT software (Verity Software House).

Table 2.1 Primers used for genotyping and qRT-PCR

Name	Sequence (5' to 3')	Purpose
Chaos3typeL	CATTGATCAGCTCATCACCA	Chaos3 mutant genotyping
Chaos3typeR	CACATACCATTTGCTTGTCAG	
Mcm2-GT-F	CCCTCCTCCTGCAGGTGGAAAGCAC	Mcm2 gene-trap genotyping
Mcm2-GT-R1	GCAGTAGAGTTCCCAGGAGGAGCC	
Mcm2-GT-R2	GGTGGTGTAAGGAACAGATGGAC	
Mcm2 mRNA-F	TTCCCGCTTTGATGTCCTG	Mcm2 qRT-PCR
Mcm2 mRNA-R	ACCATTAGTCAACCCTTCATCC	
Mcm3 mRNA-F	AGGAAGACTCATGCCAAGGATGGA	Mcm3 qRT-PCR
Mcm3 mRNA-R	TGGGCTCACTGAGTTCCACTTTCT	
Mcm4 mRNA-F	ACAGGAATGAGTGCCACTTCTCGT	Mcm4 qRT-PCR
Mcm4 mRNA-R	AAAGCTCGCAGGGCTTCTTCAAAC	
Mcm5 mRNA-F	CTGGATGCTGCTTTGTCTGGCAAT	Mcm5 qRT-PCR
Mcm5 mRNA-R	TGTGTTCAGACACCTGAGAGCCAA	
Mcm6 mRNA-F	TCACCAAGTCCTCGTGGAGAATCA	Mcm6 qRT-PCR
Mcm6 mRNA-R	TTTAGGCTGAACCTCGTCACAGCA	
Mcm7 mRNA-F	CCCTGCCCAATTTGAACCTTTGGA	Mcm7 qRT-PCR
Mcm7 mRNA-R	TCTCCACATATGCTGCGGTGATGT	
Pcna mRNA-F	GGGTTGGTAGTTGTCGCTGT	Pcna qRT-PCR
Pcna mRNA-R	TCCAGCACCTTCTTCAGGAT	
Cdc6 mRNA-F	GCTGCCCTGGACTTTTTAAG	Cdc6 qRT-PCR
Cdc6 mRNA-R	GCTGCTTGACTCGGATATGA	
Cdkn2a-qF	ACATCAAGACATCGTGCGATATT	Cdkn2a qRT-PCR
Cdkn2a-qR	CGGTACACAAAGACCACCC	
mp16INK4a-qF	AATCTCCGCGAGGAAAGCGAACT	p16-Ink4a qRT-PCR
mp16INK4a-qR	GTGAACGTTGCCCATCATCATCACC	
mp19ARF-qF	CGTGAACATGTTGTTGAGGCTAGAGA	p19-ARF qRT-PCR
mp19ARF-qR	TCTGCACCGTAGTTGAGCAGAAGAG	
ActB mRNA-F	ACCTTCTACAATGAGCTGCG	beta-Actin qRT-PCR
ActB mRNA-R	CTGGATGGCTACGTACATGG	

2.4. Results

Cells with intrinsic RS senesce prematurely in culture

Mouse embryonic fibroblasts (MEFs) isolated from mutant that has elevated DNA replication stress (RS) senescence prematurely (254). *Mcm4*$^{Chaos3/Chaos3}$ ("*Chaos3*") MEFs also display signs of endogenous RS and showed growth defects (135, 251). To determine if it is a sign of premature senescence and whether it is linked to the increased RS in this mutant, we monitored the growth of freshly isolated wild-type (WT) and *Chaos3* MEFs for several passages, under typical culture conditions (~20% atmospheric O_2). The *Chaos3* cultures exhibited reduced growth and eventual arrest at earlier passages than WT cultures (Figure 2.1A). Furthermore, approximately twice as many cells in *Chaos3* MEF cultures were positive for senescence-associated β-galactosidase (SA-β-gal) expression than in WT cultures (Figure 2.1B).

Unlike normal human primary cells that senescence in culture is due to telomere erosion as a consequence of repetitive DNA replication cycles, oxygen sensitivity and DNA replication stress are the two major causes of natural senescence in cultured primary mouse cells (190). Low oxygen conditions (5% O_2) resulted in faster growth of both WT and *Chaos3* MEFs compared to standard conditions (20% O_2; Figure 2.1A). Whereas WT MEFs continued proliferating in 5% O_2, this condition only delayed the onset of senescence of *Chaos3* MEFs before they stopped growing at P6-P7 (Figure 2.1A). The growth defect of *Chaos3* MEFs under normal and low oxygen conditions is similar to that of MEFs defective in the non-homologous end joining (NHEJ) pathway of DNA double strand break (DSB) repair (190). In aggregate, these observations suggest that intrinsic RS caused by the *Chaos3* mutation, not hypersensitivity to oxidative stress, triggers premature senescence.

Endogenous RS induces MCM2-7 pan-reduction in association with senescence progression

Chaos3 MEFs were reported to have 40-60% less MCM2-7 protein compared to WT cells (249, 251). The difference is also reflected at the mRNA level and is largely specific to the MCMs, not other DNA replication and cell cycle related genes (250). Our observations that *Chaos3* cells senesce prematurely in culture prompted us to investigate the cause and effect relationships between RS, MCM2-7 regulation, and senescence.

To determine if lower MCM2-7 in *Chaos3* MEFs is constitutive or related to senescence, we measured their levels during the passaging of primary WT MEF cultures. The mRNA and protein levels of each MCM declined in WT primary MEFs as a function of time in culture (Figure 2.1C & D). Levels of PCNA, another essential DNA replication protein, also declined albeit less dramatically than MCMs (Figure 2.1C, D & E). Decreased MCM levels have also been observed in older mouse hematopoietic stem cells (HSCs), which have increased RS (255).

We further measured MCM2-7 mRNA and protein in primary *Chaos3* and WT MEFs at both early (P2) and later (P4) passages. Consistent with the aforementioned published reports (249, 251, 256), we observed 40% less MCM2-7 in P4 *Chaos3* MEFs. However, there was little reduction of MCM2-7 mRNA or protein in P2 *Chaos3* MEFs compared to WT littermate MEFs (Figure 2.1F & G). These results suggest that MCM2-7 levels decrease roughly in parallel with the progression of RS-associated cellular senescence, and is either a cause or consequence of senescence. Furthermore, the results indicate that MCM2-7 pan-reduction in *Chaos3* cells is not an incipient property, but like in WT MEFs, is acquired and likely a consequence of RS from culture conditions, which is exacerbated or accelerated by the defective MCM4^{Chaos3} protein.

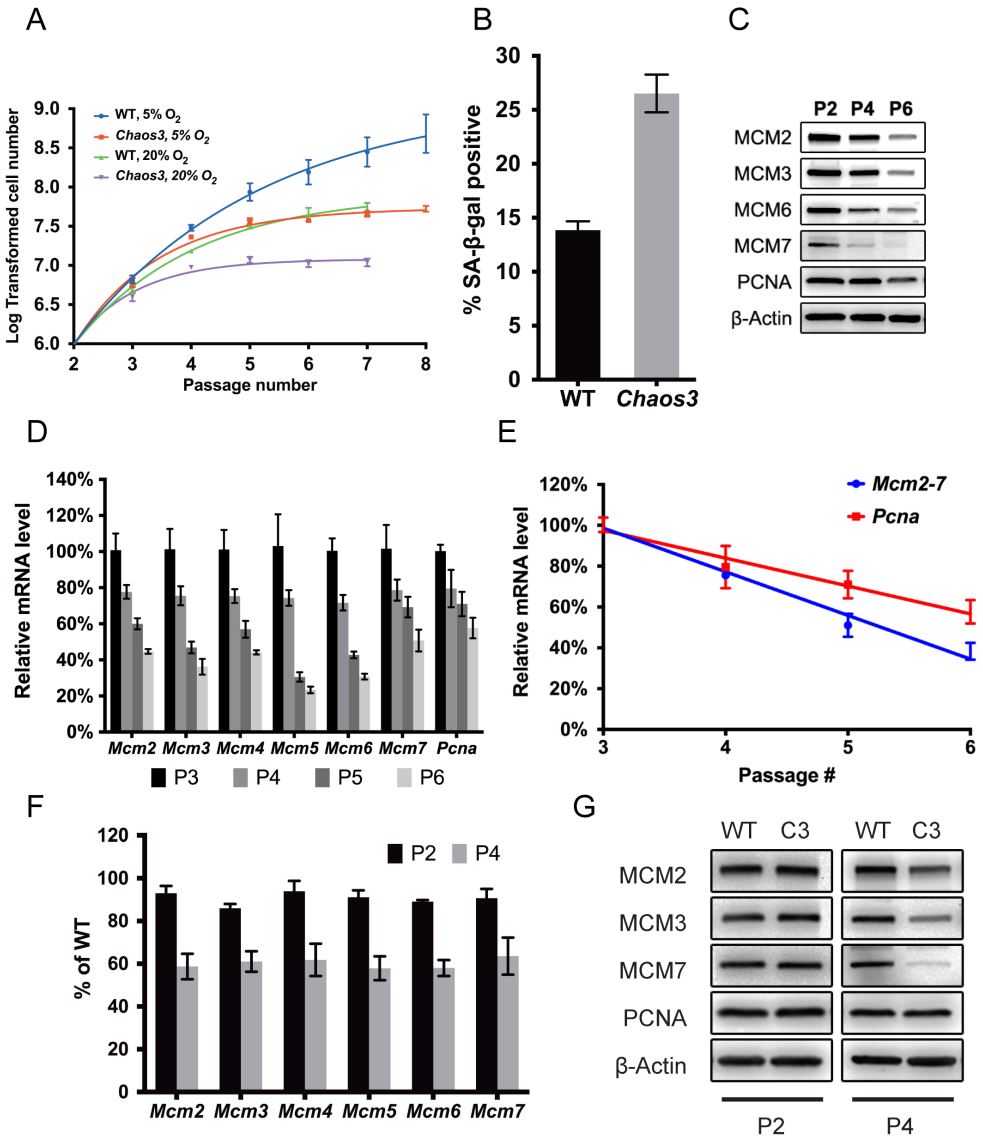


Figure 2.1 MCM2-7 pan-reduction accompanies senescence in primary MEFs, and both are accelerated by *Chaos3*. **(A)** Replicative lifespans of *Chaos3* and WT primary MEFs. Cells were maintained under the indicated O_2 levels. Error bar = SEM. **(B)** Senescence is accelerated in *Chaos3* MEFs. The % of cells positive for SA-β-gal (senescence-associated β-galactosidase) was assessed at passage 5. Error bar = SEM. **(C)** Western blot analysis of total cellular MCM protein during passage of primary MEFs. Immunoblots were probed with the indicated antibodies. **(D)** *Mcm2-7* and *Pcna* mRNA levels in WT MEFs decline precipitously in culture. The level of each gene was measured by qRT-PCR. mRNA level at P2 is considered to be 100%. At least 3 biological replicated were performed for each data point. Error bar = SEM. P = passage. **(E)** Same results from **(D)** is plotted to show accelerated *Mcm2-7* reduction than *Pcna*. **(F)** *Mcm2-7* mRNA and **(G)** protein levels decline more rapidly in primary *Chaos3* MEFs than WT MEFs during culture. Note that PCNA was unchanged between genotypes at both time points, while MCMs decreased between P2 and P4.

Chronic RS induces MCM2-7 reduction that coincides with loss of DNA replication potential

Given our data indicating that MCM2-7 downregulation accompanies senescence in *Chaos3* cells, we postulated that MCM2-7 pan-reduction is an authentic cellular response to RS in general, and that it contributes to the RS-induced cell cycle arrest.

To test this, we treated early passage (P2) WT MEFs with a low concentration of HU as a means of imposing chronic RS. Long term HU treatment inhibited cell growth (Figure 2.2A) and caused an accumulation of cells in S phase (2-4 fold increase over untreated after 24-72h; Figure 2.2B). DNA replication spiked within the first 24 hours of exposure, and declined compared to controls after 48 and 72 hours (Figure 2.2C). However, the HU treatment caused reduction in MCM2-7 but not *Pcna* or the licensing factor *Cdc6* (Figure 2.2D), a scenario similar to that in untreated *Chaos3* cells (250). The MCM2-7 pan-reduction was dependent upon HU dosage. Longer exposure (Figure 2.3A) or increased concentration of HU (Figure 2.3B & C) caused greater MCM2-7 reduction. Notably, MCM mRNA reduction in WT primary MEFs coincides with the decline in DNA replication (Figure 2.3A & C). These results suggested that moderate MCM2-7 pan-reduction is a legitimate cellular reaction to different levels of RS, whereas cells might further weigh on the degree of MCM2-7 repression and thus the severity of RS to make cell cycle related decisions.

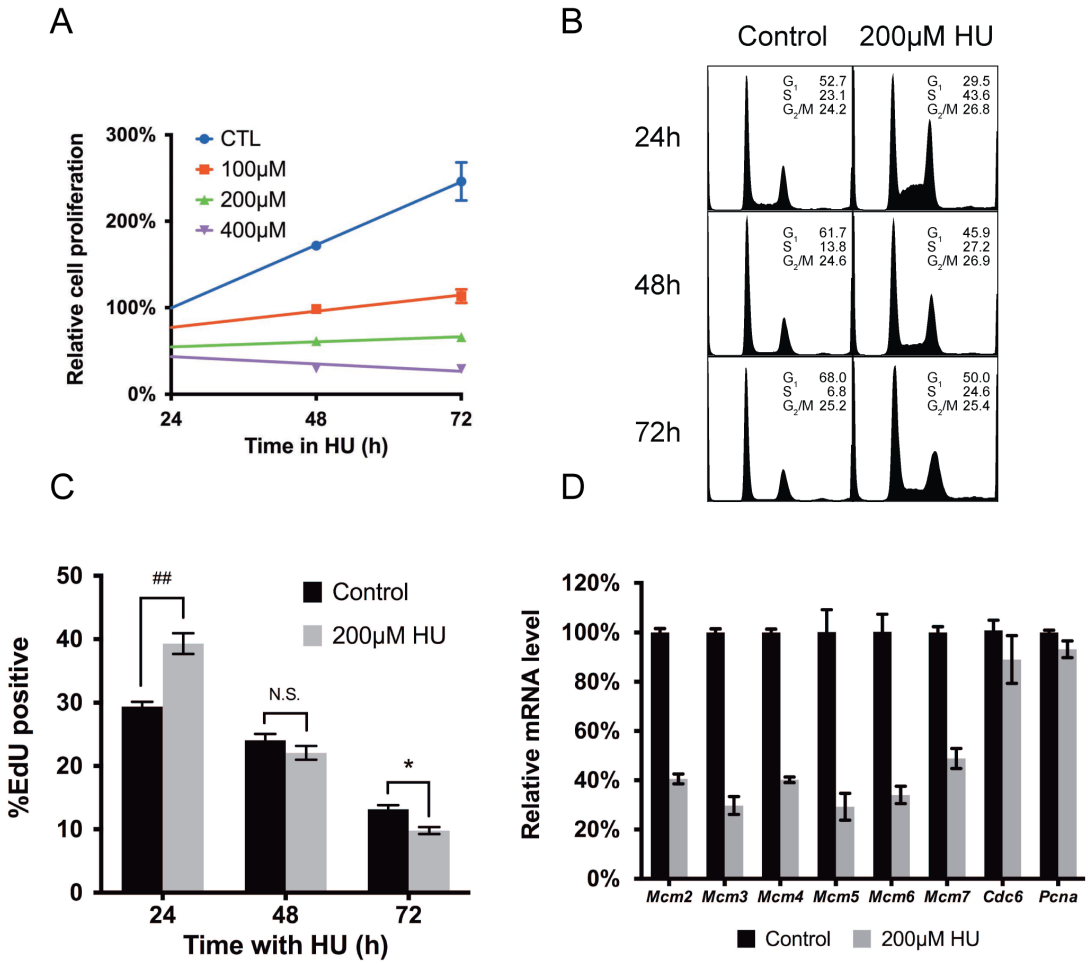


Figure 2.2 Chronic RS causes cell cycle arrest and *Mcm2-7* specific reduction. (A) Cell growth of WT primary MEFs treated with hydroxyurea (HU). Relative cell number is the percentage vs. the untreated group on day1 (considered to be 100%). Error bar = SEM. **(B)** Cell cycle profiles of WT primary MEFs under chronic RS. Plots are of DNA content as assessed by flow cytometry. The percentages of cells in each stage of the cell cycle are shown. Three technical repeats were performed. **(C)** Persistent low level RS induces progressive loss of DNA replication in WT primary MEFs. In the short term (24h), HU promotes EdU incorporation (**, $p \approx 5x10^{-5}$, two-sided t-test). However, long-term HU exposure (72h) eroded DNA replication potential significantly (*, $p \approx 0.001$, two-sided t-test). N.S. = not significant. **(D)** *Mcm2-7* specific repression after persistent RS. Plotted are mRNA levels as determined by qRT-PCR following 200μM HU treatment for 72h. Error bar = SEM.

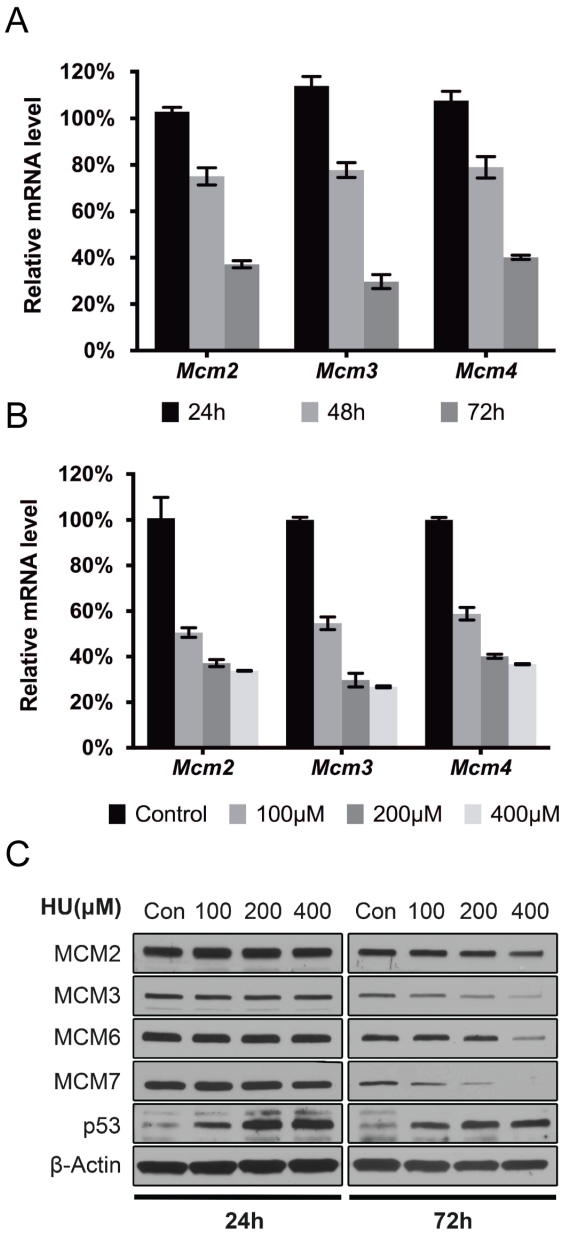


Figure 2.3 Chronic RS induced MCM2-7 reduction is time and dosage dependent. (A) Persistent RS induces MCM repression. *Mcm2-4* mRNA levels in WT primary MEFs were measured by qRT-PCR following 200μM HU treatment for the indicated periods of time. The values plotted are compared to untreated cells. Error bar = SEM. **(B)** Dosage-dependent *Mcm2-7* repression after persistent RS. *Mcm2-4* mRNA levels in WT primary MEFs were measured by qRT-PCR following 72h of culture in the indicated concentrations of HU. Error bar = SEM. **(C)** Induced RS reduces MCM protein expression in a time- and dose-dependent manner after HU treatment of WT primary MEFs. Shown are immunoblots of total protein extracted from primary MEFs exposed to the indicated concentrations of HU for 24 and 72h. Note TRP53 stabilization following HU treatment.

Reduced MCM dosage sensitizes cells to chronic RS induced senescence

The results presented thus far show that MCM2-7 levels are reduced in response to spontaneous or chemically-induced RS, leading to reduced DNA replication and cell cycle arrest. We also found that MCM2-7 expression level is closely connected to cell growth and the progression of senescence. These observations suggest an intrinsic connection between RS, MCM reduction and senescence induction.

We first evaluated if chronic RS treatment on primary WT MEFs can induce premature senescence. Two separate genes $p16^{Ink4a}$ and $p19^{ARF}$ are encoded from the same *Cdkn2a* genomic locus, which can be induced in response to persistent RS (254). They also function upstream of *Trp53* and retinoblastoma (*Rb*) respectively to promote senescence onset. Prolonged exposure (72h) to HU triggered a 2-3 fold increase of the $p16^{Ink4a}$ and $p19^{ARF}$ tumor suppressors (Figure 2.4A), and also an ~5-7 fold increase in the percentage SA-β-gal positive cells (Figure 2.4B) in early passaged WT primary MEFs. These results further confirmed the connection between chronic RS and senescence induction in cultured primary mouse cells.

Next we want to determine the role of MCM repression in RS induced senescence. Whereas RS induced MCM2-7 pan-reduction coincides with loss of DNA replication, we hypothesize that this intrinsic MCM downregulation should further sensitize the cells to RS. To test this, we utilize the primary MEFs heterozygous for *Mcm2* ("M2"), in which MCM2 expression is already reduced by at least 50% due to genetic intervention (249). Surprisingly, this level of reduced MCM2 dosage *per se* does not cause premature senescence in the primary MEFs isolated from this mutant, as they proliferated identically to WT littermate cultures under standard culture conditions (Figure 2.4C). This also argues against that the growth defect of primary *Chaos3* MEFs was due to reduced MCM2-7 expression (135), as the degrees of MCM

reduction between M2 and *Chaos3* mutants are comparable. However, HU treatment triggered more severe senescence-associated phenotypes in M2 than WT primary MEFs, both in terms of markedly higher $p16^{Ink4a}/p19^{ARF}$ induction (Figure 2.4A) and a ~2 fold increase in cells positive for SA-β-gal (Figure 2.4B). Interestingly, M2 primary MEFs expressed slightly higher basal levels of senescence markers than WT MEFs (Figure 2.4A & B). Genetic reduction of MCM2 alone also causes a moderate MCM2-7 pan-reduction (128, 249, 252), supporting the idea that MCM reduction itself sensitizes cells to senescence. Thus, we conclude that the although RS-induced MCM2-7 repression does not contribute to DNA inhibition and cell cycle arrest directly, but it further sensitize cells to additional RS by contributing to senescence induction.

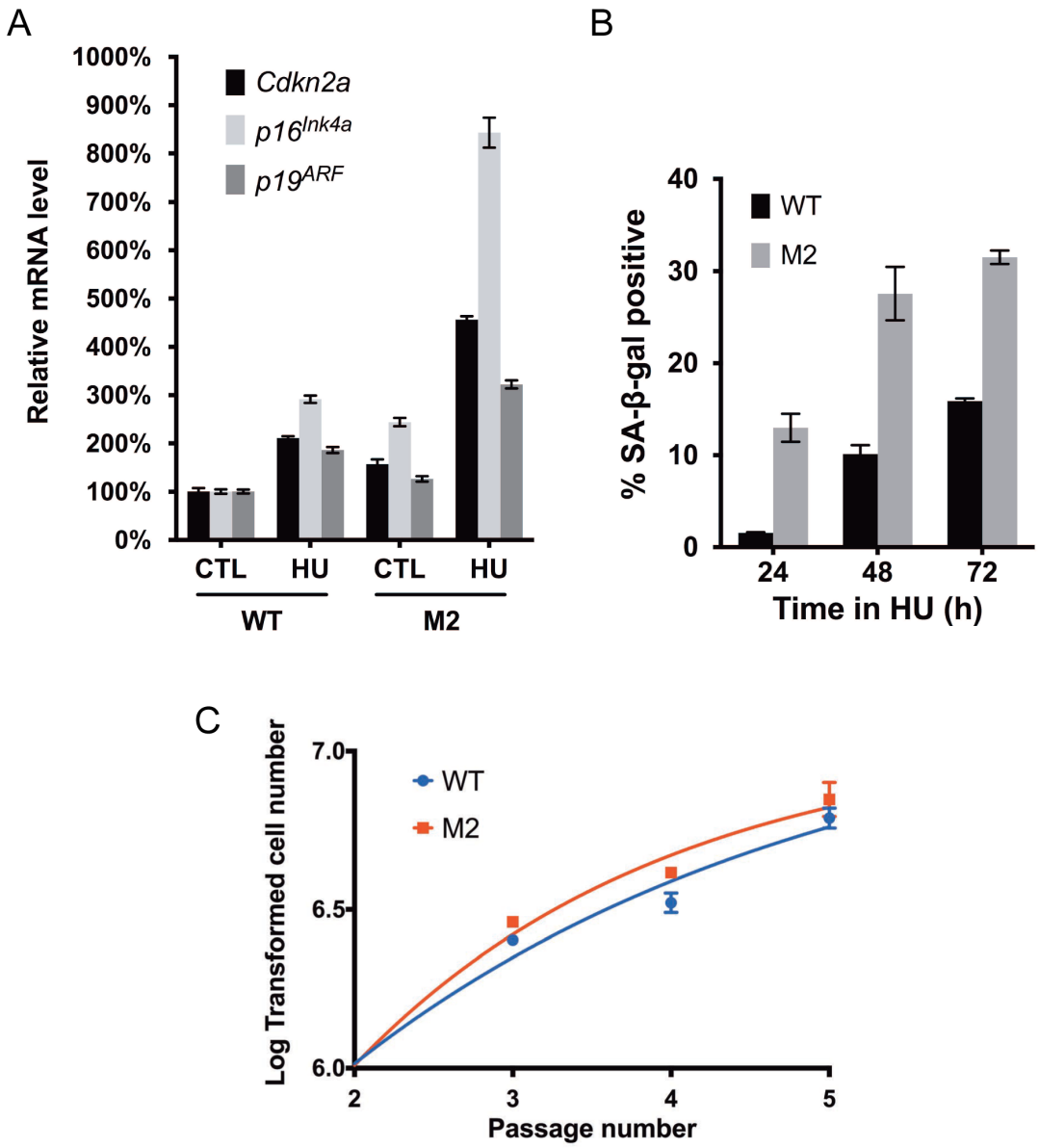


Figure 2.4 Chronic RS induces senescence in primary cells, accelerated by MCM haploinsufficiency. (A) Replication-stressed MEFs heterozygous for *Mcm2* exhibit dramatic upregulation of senescence-related transcripts. Cells of the indicated genotypes were exposed (or not) to 200μM HU for 72h. Primers specific to the $p16^{Ink4a}$ and $p19^{ARF}$ products of the *Cdkn2a* locus, as well as the primary transcript of *Cdkn2a* were quantified by qRT-PCR and normalized to β-actin levels. M2 = $Mcm2^{Gt/+}$, where "Gt" indicates that this is a gene trap allele that is functionally null (249). Error bar = SEM. **(B)** MEFs heterozygous for *Mcm2* are hypersensitive to RS-induced senescence. Cells were treated with 200μM HU for the indicated times. M2 = $Mcm2^{Gt/+}$. The percentage of cells positive for SA-β-gal staining is shown for each sample. Error bar = SEM. **(C)** Replicative lifespan of MCM2 gene-trap (M2) primary MEFs and their WT littermates. Error bar = SEM.

RS induced MCM2-7 repression and senescence depends on DDR and *Trp53*

Senescence is a DNA damage response (DDR), which is also tightly related to DNA replication process (187). Whereas human cells rely on both the *ARF-p53* and *Ink4a-Rb* pathways to induce senescence, mouse cells mainly depends on *ARF-Trp53* axis (257). Since MCM2-7 expressions are intrinsically downregulated reflecting the intensity and duration of RS, which in turn promotes terminal cell cycle arrest through senescence when RS persist, we wonder whether MCM2-7 repression is also regulated by DDR.

The ATM-CHK2 axis of the DDR is probably involved in responding to endogenous RS in the *Chaos3* mutant, as a genetic analysis revealed that impaired ATM-CHK2, but not ATR-related DDR pathway aggravated cancer susceptibility phenotypes in *Chaos3* mutant mice (258). In addition to these previous observations, we found that *Chk2* deletion in the *Chaos3* mutant partially restored double mutant MEFs proliferation over the *Chaos3* single mutant (Figure 2.5A). *Chk2* deletion also partially yet significantly rescued exogenous RS induced MCM2-7 pan-reduction (Figure 2.5C), and their expressions in the primary *Chaos3* mutant MEFs (Figure 2.5D). However, it failed to rescue the genomic instability phenotype in the *Chaos3* mutant as measure by micronucleus (MN) formation (Figure 2.5B).

Activated DDR pathways also converge on *Trp53* activation to promote DNA damage repair, transiently delay cell cycle progression, or terminally eliminate cells from cell cycle through apoptosis or senescence. *Trp53* function was implied in *Chaos3* mutant, in which *Trp53* deletion rescues MCM2-7 expression and restores proliferation of *Chaos3* cells (256). We further determined that *Trp53* is responsible for RS induced MCM2-7 repression and senescence induction. Similar to WT MEFs, the same HU treatment stimulated $p19^{ARF}$ expression in *Trp53*-null primary MEFs (Figure 2.6A). However, as ARF functions upstream of TRP53 to induce

senescence in primary mouse cells (248), the terminal senescence phenotype is bypassed in *Trp53*-null MEFs, as indicated by the barely detectable SA-β-gal expression after RS induction (Figure 2.6B). Surprisingly, whereas HU treatment caused a decrease in MCM mRNA in WT MEFs (Figure 2.2D), the identical treatment caused a dose-dependent 50-100% increase of *Mcm2*, *3* and *4* mRNA in *Trp53*-deficient cells treated with increasing concentrations of HU (Figure 2.6C). These increases were also reflected at the protein level (Figure 2.6D), unlike WT. Interestingly, HU-treated *Trp53*-null MEFs exhibited a drastic increase in H2AX Ser139 phosphorylation (γH2AX) compared to WT cells receiving the same treatment (Figure 2.6D), indicating greatly elevated DNA damage and/or replication fork errors, and likely reflecting a failure of cells that have accumulated such defects to undergo senescence / apoptosis. These results confirmed our hypothesis that MCM2-7 regulation in response to RS is connected to DDR through *Trp53* function, while *Chk2* and its upstream DDR components might be responsible for detecting and signaling the DNA damage caused by chronic RS.

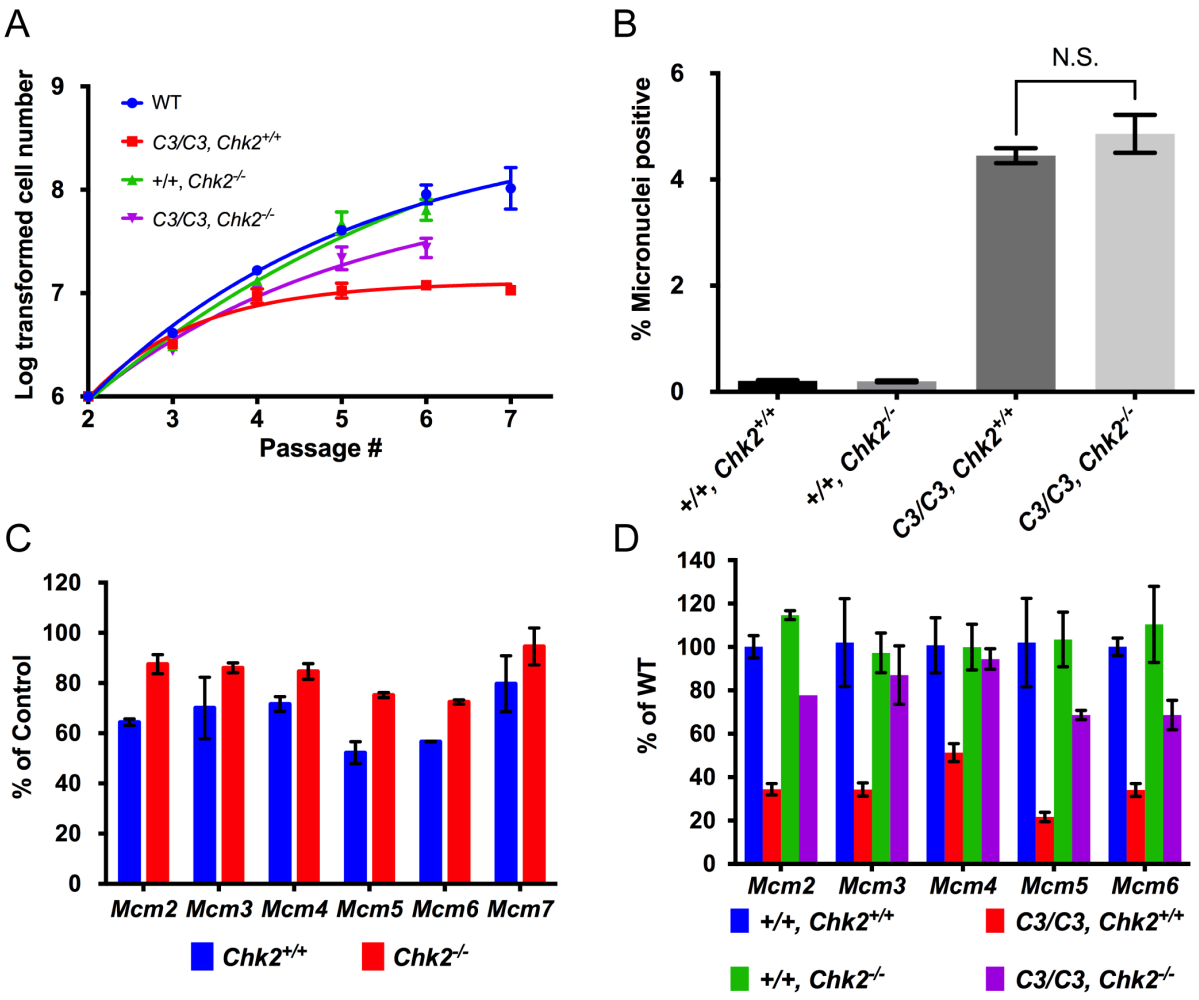


Figure 2.5 *Chk2* deletion modifies phenotypes in the *Chaos3* mutant. (A) Replicative lifespans of *Chaos3* and WT primary MEFs with additional *Chk2* deletion. Cells were maintained under regular culture conditions. Error bar = SEM. **(B)** *Chk2* deletion in the *Chaos3* mutant did not rescue micronuclei formation (two-sided t-test). Numbers of individuals analyzed for each genotype (from left to right) were 12, 6, 31 and 6 respectively. Error bar = SEM. **(C)** Results of qRT-PCR measuring *Mcm2-7* mRNA expression in *Chk2* mutants after 300μM HU treatment for 48h. *Chk2* deletion (*Chk2*$^{-/-}$) partially rescued HU induced *Mcm2-7* pan-reduction in the *Chk2*$^{+/-}$ mutant. Error bar = SEM. **(D)** Results of qRT-PCR measuring *Mcm2-7* mRNA expression in *Chaos3 (C3)* / *Chk2* compound mutant MEFs at passage 5. Error bar = SEM. 2 littermate pairs were measured for each genotype.

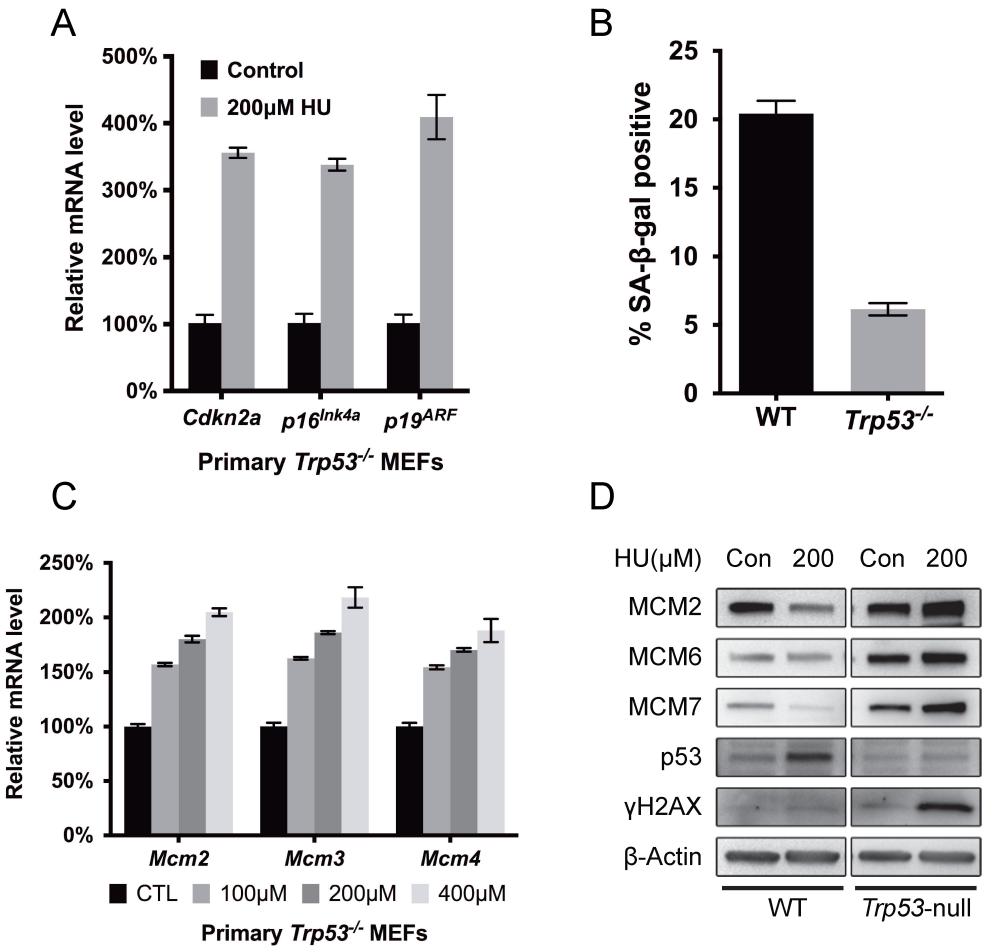


Figure 2.6 RS-induced MCM2-7 downregulation and cellular senescence is dependent upon TRP53. (A) RS treatment in primary *Trp53*$^{-/-}$ MEFs upregulates senescence-associated transcripts similarly to WT MEFs. *Trp53*$^{-/-}$ primary MEFS were treated (or not) with 200μM HU for 72h. Primers specific to the *p16*^{Ink4a} and *p19*^{ARF}, products of the *Cdkn2a* locus, as well as the primary transcript of *Cdkn2a* were quantified by qRT-PCR and normalized to β-actin levels. **(B)** TRP53 deficiency prevents cellular senescence induction in response to chronic RS. Percentage of cells positive for SA-β-gal after 72 hours of exposure to 200μM HU is plotted. Error bar = SEM. **(C)** HU-induced RS causes *Mcm2-4* up-regulation in P2 *Trp53* null primary MEFs. Cells were cultured in the indicated concentrations of HU for 72h. Plotted are mRNA levels as assessed by qRT-PCR. Control cells received no HU, and these mRNA levels are considered to be 100%. *Mcm5-7* responded similarly (not shown). Error bar = SEM. **(D)** RS induces DNA damage and increases MCM protein in *Trp53*-null primary MEFs. Cells were treated with 200μM HU for 72h, and then total protein was extracted for immunoblotting with antibodies against the indicated proteins. The phosphorylated form of H2AX (γH2AX) is an indicator of double strand breaks and replication fork stalling.

2.5. Conclusions and Discussion

In this chapter, I discussed the MCM2-7 regulation in primary cells as an authentic cellular response to DNA replication stress (RS). I started by confirming the previous observations that chronic RS can inhibit normal cell proliferation and resulted in premature senescence in cultured primary mouse cells. In the meantime, the expression of MCM2-7 replication genes gradually reduced as the primary cells lost proliferation potential and becoming senescence. I further confirmed that active MCM2-7 repression is part of the normal cellular response to chronic RS. This regulation is mediated by DNA damage response (DDR), which also depends on normal *Trp53* function. Since moderately reduced MCM expression *per se* does not inhibit normal cell proliferation and cell cycle progression, the RS-induced MCM suppression is likely to serve as a molecular memory in the cells of how much RS has been experienced by the proliferating cells accumulatively. Reduced MCM expression level thus further sensitizes the proliferating cells to additional RS by activating cellular senescence and arrest cell cycle progression indefinitely. Thus, DDR mediated, *Trp53* controlled MCM2-7 reduction may serve as a malignancy prevention mechanism in the proliferating cells in the event of persistent RS.

The connection between MCM2-7 suppression and senescence was first implied when studying the $Mcm4^{Chaos3}$ (*Chaos3*) mutant. Primary MEFs isolated from the homozygous mutant embryos showed signs of growth defects such as reduced growth rate at the population level. Besides, *Chaos3* primary MEFs also displayed abnormal cell morphologies in culture, such as enlarged cell size and multinuclear formation. These are all known phenotypes associated with cellular senescence. Clearly *Chaos3* mutant cells suffer from endogenous RS and the associated DNA damage, since markers for RS and DNA damage are elevated in them (251). Cellular

senescence has long been recognized as a DDR (187), thus it was not hard to draw a connection between elevated endogenous RS and senescence response in the *Chaos3* mutant. Since MCM2-7 pan-reduction, though to a moderate degree, was the most striking feature of the *Chaos3* mutant, endogenous RS, and the related cellular growth defects, were thus attributed to reduced MCM expression.

Indeed, as essential DNA replication genes, MCM2-7 are only expressed in proliferating cells, but not in non-proliferating and quiescent cells. Reduced MCM expression could compromise normal cellular response to RS and led to cell cycle arrest and apoptosis in cultured cancer cells (125-127). These observations thus led to the impression that any perturbation to the MCM expression level, even to a moderate degree, would negatively affect cell proliferation. MCM2-7 proteins assemble into MCM2-7 complexes to participate in DNA replication: it licenses replication origins and also serves as the catalytic core of the replication helicase. Only a pair of the MCM2-7 complexes is needed at each origin to assemble into two replication forks that migrate bidirectionally to replicate the parental genome, while each replication fork is able to cover large genomic regions greater than the distance between two licensed origins. MCM2-7 proteins are so highly expressed in proliferating cells that it ensures completeness of DNA replication at two major levels: 1) MCM2-7 complexes license a large number of replication origins and 2) each origin is loaded with more than a single pair of MCM2-7 complexes. Besides, each individual MCM protein is so highly expressed to ensure enough copies of the MCM2-7 complexes can form and being transported into the nucleus to perform its functions (119). Thus, theoretically speaking, a moderate reduction to the MCM expression level would have minimal impact, if any, on regular cell growth, which is confirmed here by the study on primary cells isolated from MCM2 gene-trap mouse mutant. Severer MCM reduction (by ~70%) also seems to

support normal cell proliferation (130). Thus, for the case of *Chaos3* mutant cells, reduced MCM expression is most definitely not the cause of reduced cell proliferation and premature senescence. MCM2-7 pan-reduction is related to RS, possibly as a consequence of elevated level of RS in the *Chaos3* mutant, rather than the cause of it, which was confirmed by the experiments performed on WT primary cells exposed to generic DNA replication inhibitor hydroxyurea (HU). Furthermore, this RS-induced MCM2-7 regulation is mediated by the conserved DDR mechanism that is responsible for the prevention of malignant cellular transformations, suggesting MCM2-7 downregulation also participates in this process. Indeed, I found that reduced MCM expression can sensitize the cells to RS by accelerating senescence induction. In proliferating cells, DNA damage associated with the replication process should be detected and signaled through the DDR. After resolving these damage, DDR activity should be reset to normal to allow for continued cell cycle progression. However, unlike many other genotoxic stresses that are usually transient and induce DNA damage in a rather acute manner, RS is usually persistent and happens at relatively low level. Replisomes constantly and inevitably face challenges posed by RS during DNA replication. Due to the mild nature of RS, not all of the RS associated DNA damage can be detected and repaired by DDR, thus they may accumulate in the proliferating cell lineages. These alterations to the genomic information are great threat to the maintenance of genomic integrity, not to mention that some of them may have oncogenic potential. Since the molecular memory of RS and its associated DNA damage carried by DDR activation is erased once cells divide and reenters cell cycle, replicating cells must depends on other sensory mechanisms to record their struggling history dealing with RS. The expression level of MCM proteins appears to be a perfect candidate as the carrier of such memory: 1) MCM are licensing proteins, which are needed before actual replication process. Cells can evaluate

MCM expression level directly or the replication licensing to determine whether to proceed into S phase and commit to DNA replication. Licensing inhibition in primary cells through MCM knockdown can deplete the cyclins and cyclin dependent kinases needed for S phase entry, thus leading to G1/S phase arrest. The presence of "licensing checkpoint" was thus proposed which evaluates the licensing process. It appears normal licensing checkpoint also depends on normal *Trp53* function, and is only present in primary cells, but not in transformed cells such as cancer cells. Licensing checkpoint is thus the preliminary DNA replication related checkpoint to evaluate the readiness of committing to DNA replication process (202). 2) MCM proteins are so highly expressed. As previously mentioned, cells can tolerate substantial MCM loss without compromising normal DNA replication and cell proliferation. Thus a rich reserve of MCM proteins ensures proliferating cells can tolerate large amount of chronic RS and having an extended replicative lifespan. This character of high MCM expression might have great importance in stem cells, as they are devoted to self-renew. Stem cells usually have a replicative lifespan as long as the lifespan of the whole organism. They should still be able to support self-renew at later stage of their replicative lifespan after experiencing the inevitable RS during DNA replication at earlier replicative ages. Selective downregulation of MCM genes was found in hematopoietic stem cells (HSCs) isolated from old mice, which also displayed signs of elevated level of RS (255). Besides, the active MCM suppression mechanism due to elevated level of RS as I described here in proliferating somatic cells should also be present in stem cells (259). Genetic mutations that moderately inhibit replication licensing often lead to stem cell defects and developmental syndromes in human (129, 136). Unlike somatic proliferating cells that persistent RS will eventually lead to termination of cell cycle progression by inducing senescence or apoptosis, stem cells can also choose to exit the cell cycle by committing to differentiation. It is

thus interesting to know how reduced replication licensing, either by inhibiting licensing reaction *per se* or as a result of prolonged exposure to RS, would affect the maintenance of "stemness" in stem cells.

The definitive evidence that proves MCM2-7 are regulated in response to RS is that it is mediated by DDR. Disruption of DDR function led to suppressed MCM2-7 reduction in response to RS. Furthermore, MCM2-7 regulation appears to be completely depending on the normal *Trp53* function, which is activated due to DDR activation in response to RS. I further interrogated the potential upstream DDR components that led to this *Trp53* activation. Surprisingly, the ATM-CHK2 pathway appears to be mainly responsible for chronic RS induced cell cycle arrest and MCM2-7 suppression. This observation is consistent with the mouse genetics study, in which ablation of ATM-CHK2 axis of the DDR, rather than ATR-CHK1 axis led to aggravated cancer predisposition in the *Chaos3* mutant mice, suggesting ATM-CHK2 side of the DDR was required to prevent cancer development (260). In cultured *Chaos3* cells, molecular markers that usually associate with the stalled replication forks were only marginally activated comparing to the WT cells, while CHK1 activation was also absent (251). Similar observation was made in MCM mutant that has severe MCM reduction (252). According to the current view on DDR mechanism involved in RS response, RS induced replication fork stalling often results in excessive ssDNA exposure, which activates ATR-CHK1 axis of the DDR. However, in the event of persistent RS associated with reduced MCM expression, ssDNA exposure at stalled replication forks may not be the major form of DNA lesions. Similarly, chronic RS caused by dNTP depletion (HU treatment) also seems to depend on ATM-CHK2 pathway to repress MCM expression. Thus, chronic RS may not abruptly stall replisome movement and activate ATR-CHK1 pathway of the DDR, but by generating DNA damage that

are usually recognized by the ATM-CHK2 side of the DDR, presumably due to DNA double strand breaks resulted from uncompleted DNA replication caused by chronic RS. Since CHK2 deletion failed to fully rescue RS-induced MCM2-7 repression, other DDR mechanisms must be involved in TRP53 activation to repress MCM expression.

In conclusion, MCM2-7 repression is an authentic cellular response to chronic RS, which participates in RS induced terminal cell cycle arrest such as senescence to prevent malignant cellular transformation. It is mainly mediated by the CHK2 related DDR mechanism, and completely depends on normal TRP53 function. Cells can tolerate moderate MCM reduction and sustain normal DNA replication and cell proliferation, however they are sensitized to additional RS with reduced MCM dosage. Thus, highly expressed MCM2-7 genes are not only important for dormant origin licensing to react to acute RS that stalls replication forks, their expression levels also serve as a molecular memory of how much chronic low level RS has been experienced, or to be tolerated during the replicative lifespan of proliferating cells.

CHAPTER 3

CHAOS3: A DNA REPLICATION HELICASE MUTANT DISRUPTS HELICASE FUNCTION[2]

Gongshi Bai [1, 2], Marcus B. Smolka [1, 3, 4], and John C. Schimenti [1, 2, 4]

[1] Cornell University, Dept. of Molecular Biology & Genetics, Ithaca, NY 14853, USA

[2] Cornell University, Dept. of Biomedical Sciences, Ithaca, NY 14853, USA

[3] Cornell University, Weill Institute for Cellular and Molecular Biology, Ithaca, NY 14853, USA

[4] Cornell University Center for Vertebrate Genomics, Ithaca, NY 14853, USA

3.1. Abstract

Minichromosome maintenance 2-7 (MCM2-7) proteins form heterohexameric complex and play essential roles during every aspect of eukaryotic DNA replication. MCM2-7 are loaded onto chromatin to "license" the genome for DNA replication. The entire complex also possesses ATPase activity and serves as the core of DNA replication helicase that unwinds double stranded DNA to generate single stranded DNA, which is used by DNA polymerase as template to synthesize new DNA strands based on Watson-Crick pairing. MCM2-7 complex stability is essential for normal replication fork progression and completeness of DNA replication. Disruption of MCM2-7 complex integrity leads to immediate DNA replication arrest and cell death. Moreover, replisome can only reassemble to restart replication when MCM2-7 complex is maintained at stalled replication forks. *Mcm4*^{Chaos3} ("*Chaos3*") mutant displays constant low level of DNA replication stress (RS), which promotes genomic instability and cancer predisposition in mice. The *Chaos3* point mutation disrupts MCM2-7 complex stability *in vitro*, which is also

---

[2] Work presented in this chapter has been submitted for publication. Authors' contributions: G.B. & J.C.S conceived & designed the research; G.B. & M.B.S performed experiments and analyzed data; G.B. & J.C.S wrote the paper.

accompanied by a moderate MCM2-7 pan-reduction. Here we report that *Chaos3* mutation disrupts helicase complex stability and association with replication forks *in vivo*, which compromises its function under normal and stressful DNA replication conditions that drives genomic instability and oncogenesis.

3.2. Introduction

Minichromosome maintenance 2-7 (MCM2-7) proteins form a stable heterohexameric complex *in vitro* and *in vivo*, which plays pivotal roles throughout eukaryotic DNA replication. "Licensing" must occur to prepare the genome for DNA replication, during which MCM2-7 complexes are loaded onto chromatin (7, 94). Origin Recognition Complexes (ORCs) bind to genomic loci known as DNA replication origins, which further recruit CDC6 and CDT1 to load MCM2-7 (5). Two MCM2-7 hexamers are loaded during each round of licensing reaction, which can both assemble into functional replication helicase and initiate bidirectional replication from a single origin (47). "Re-licensing" is prohibited during S phase to prevent over replication of any parts of the genome. As a result, cells must rely on the MCM2-7 complexes already loaded onto chromatin before S phase to accomplish DNA replication. During DNA replication initiation, Cyclin-dependent kinases (CDKs) and Dbf4-dependent kinase (DDK) phosphorylate MCM2-7 subunits and recruit replication helicase cofactors CDC45 and GINS (Sld5, Psf1, Psf2, and Psf3) complex. The entire CMG complex (Cdc45, MCM2-7 & GINS) has *in vitro* helicase activity and functions as replication helicase *in vivo* (55), which melts origin DNA to further accommodate replisome assembly. During replication elongation, replication helicase unwinds the double stranded parental DNA and provides template for replication polymerases to duplicate genome information semi-conservatively.

MCM2-7 complex stability is essential for eukaryotic DNA replication. MCM2-7 proteins often co-precipitate with the others *in vitro*, as they form a stable heterohexameric complex (77, 84). Mutations in MCM2-7 proteins can disrupt MCM2-7 complex formation (78, 79) and cause cell viability problems (80, 81). Even a single point mutation in one of MCM2-7 subunits can change the overall conformational of the entire complex and disrupts its normal function (261). MCM degradation during S phase disrupted MCM2-7 integrity and led to immediate DNA replication fork stalling, cell cycle arrest and eventual cell death (98). MCM mutation that disrupted helicase function was found in mice, which promoted high incidence of tumorigenesis (134). Evaluation of human cancer genomes revealed high frequency of MCM mutations that might have deleterious effects on protein structure and functions (130).

Chaos3 mutants have DNA replication defects, which promote genomic instability and malignant transformation in mice (135, 249-251, 256). This allele encodes a single amino acid change (Phe345Ile) in a highly conserved residue in the MCM4 protein (135). It also resides close to the conserved zinc-finger domain in the MCM proteins, which is important for MCM2-7 hexameric complex formation (78, 79). A further review on the *Chaos3* mutation based on sequence conservation and the known archaea MCM crystal structure revealed that it might also reside on the interactive surface between neighboring MCM subunits during MCM2-7 complex formation (250). Consistent with these predictions, *in vitro* analysis suggested that *Chaos3* mutation disrupts MCM2-7 and helicase complex formation (250, 251). Despite that it does not negatively affect helicase activity *in vitro*, little is known about the behavior of *Chaos3* containing helicase *in vivo* (251). To study this, methods that can specifically isolate and further analyze proteins directly associate with the nascent DNA at replication forks must be used.

Multiple procedures have been established to study proteins associated with the DNA replication forks in eukaryotic cells (177, 262, 263). All of these methods rely on the labeling of nascent DNA at or closely behind replication forks using nucleotide analogues in cultured proliferating cells. Replication proteins are then cross-linked to nascent DNA and to each other, or native condition (no cross-linking) can be applied as well (264). After cell lysis and chromatin sonication, nucleotide labeled nascent DNA can be affinity purified, when the replication proteins co-precipitate with it. Finally, the proteins can be recovered and analyzed individually using immunochemistry, or by proteomic approaches such as mass-spectrometry (262, 265, 266). Multiple reports have used these methods to discover new replication proteins, study protein dynamics at normal or stalled replication forks, and investigate chromatin maturation and the transmission/maintenance of epigenetic modifications during DNA replication (262, 264-266).

Two of these methods: iPOND (isolation of Proteins On Nascent DNA) and Dm-ChP (DNA-mediated Chromatin Pull-down) were established independently and were later published simultaneously, which were designed around the very similar ideologies (177, 263). They both pulse label the nascent DNA with thymidine analog 5-ethynyl-2′-deoxyuridine (EdU) in cultured mammalian cells. “Click” chemistry was performed to covalently link a biotin tag to the EdU molecule, which facilitates nascent DNA isolation through biotin: streptavidin affinity purification. The only major difference between the two methods is that iPOND incubates whole cell lysate with streptavidin-coated beads during nascent DNA purification, while Dm-ChP uses chromatin fraction enriched lysate.

Here we used Dm-ChP combined with western blotting and mass-spectrometry to study MCM2-7 protein association with DNA replication forks under normal and stressed conditions. Our results suggest that *Chaos3* mutation not only disrupts MCM2-7 complex integrity *in vitro*,

it also compromises helicase complex association with replication fork and thus its function *in vivo*. These results were further backed up by observations that *Chaos3* mutation inhibits helicase unwinding activity *in vivo*, which may contribute to mutations that have oncogenic potentials.

3.3. Materials and methods

SV40 immortalization

An SV40 large T antigen-encoding construct (pBABE-puro SV40 LT) was packaged into lentivirus particles and used to infect primary MEFs at early passages. Cells were then selected and maintained in media containing 1.25μg/mL puromycin. Four pairs of littermate MEFs were transformed in total with this method for downstream assays.

Cell culture and stable isotope labeling of amino acids in culture (SILAC)

General cell maintenance in culture is as described in the previous chapter. Immortalized MEFs used for Dm-ChP and *in vivo* helicase assay was further maintained in the media containing 1.25μg/mL puromycin. For SILAC experiments, regular FBS was substituted by dialyzed FBS (Thermo Scientific), and SILAC DMEM was used as the base media (Thermo Scientific). For amino acid labeling, synthetic L-Lysine & L-Arginine was added to the media to a final concentration of 0.798mM and 0.398mM, respectively. For "heavy" isotope labeled cells, L-Lysine ($^{13}C_6$, $^{15}N_2$) and L-Arginine ($^{13}C_6$, $^{15}N_4$) were used (Thermo Scientific), while for "light" isotope labeled culture, regular L-Lysine & L-Arginine were added. To perform the SILAC labeling, two pairs of immortalized WT and *Chaos3* littermate MEFs were cultured in "heavy" or "light" isotope supplemented media in a 6 well plate for at least 4 passages. Cells are trypsinized and replated at the same density during each passaging, and 3~4 days of culturing was achieved between each passaging. Before the Dm-ChP experiment, SILAC labeled cells

were expanded into 150mm dishes to achieve higher cell number yields. “Heavy” and “light” isotope was reversed between the two WT and *Chaos3* MEFs pairs during labeling before Dm-ChP isolation and mass-spectrometry analysis.

Protein extraction and cell fractionation

Whole cell protein extraction was performed in RIPA buffer. Protein extraction from tissues was performed using T-PER Tissue Protein Extraction Reagent (Thermo Scientific). For fractionation, cultured cells were trypsinized and counted. After two PBS washes, cells were resuspended in Buffer A (10mM HEPES [pH7.9], 10mM KCl, 1.5mM $MgCl_2$, 340mM sucrose, 10% glycerol, 1mM DTT) with 0.1% Triton X-100 and incubated on ice for 5min. Low speed centrifugation (1,300g x 4min at 4°C) was performed to separate the supernatant (S1) and the pellet (P1). The pellet was washed once with Buffer A, then further extracted in Buffer B (3mM EDTA, 0.2mM EGTA, 1mM DTT) on ice for 30min. After centrifugation (1,700g x 4min at 4°C), the supernatant contained the nuclear fraction (S3), and the pellet (P3) containing the chromatin-bound proteins was washed once with Buffer B and then finally extracted in RIPA buffer. During the fractionations, cells were resuspended at 2.5 x 10^4 /μL at each step.

EdU incorporation assay

EdU incorporation and staining was performed as described in the previous chapter. Different cell seeding densities were tested to optimize EdU labeling condition for Dm-ChP and iPOND experiments, to maximize the absolute number of EdU positive cells during the designated labeling period of 30min.

***In vivo* helicase activity assay**

Cells were grown on coverslips and cultured in the presence of 10μM BrdU for up to 72h under standard culture condition. The BrdU-containing media was replaced with media

containing 3mM HU for 30 min. After PBS washes, BrdU staining was performed, except the HCl denaturing step was omitted. Nuclei were counterstained with Hoechst 33342. Mounted coverslips were then examined by fluorescence microscopy. Distinct BrdU staining patterns were used to determine early vs. mid/late S-phase DNA replication. Cells with detectable BrdU staininbg over the background were recognized to have normal helicase unwinding activity.

Isolation of protein on nascent DNA (iPOND)

iPOND was performed essentially as described (264). To elute nascent DNA associated protein purified using streptavidin beads, equal volume of 2x sample buffer was added to the washed and dried beads bed and boiled in water bath for 10min. 1/3 to 1/2 the volume of the final elution was analyzed using western blotting for known nascent DNA associated proteins.

DNA-mediated chromatin pull-down (Dm-ChP)

Dm-ChP was performed essentially as described (263) with the following modifications. EdU pulse labeling was performed for 30 min. For all experiments. One mg of nuclear-enriched protein lysate was incubated with 100μL pre-blocked wet streptavidin agarose beads (Novagen). Pull-down was performed at 4°C for 16~20h with constant rotation, then sequentially washed with RIPA, Wash Buffer 1 (10mM Tris [pH 8.0], 200mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1mM EDTA, 360mM NaCl), Wash Buffer 2 (Wash Buffer 1 without NaCl), and finally TE (10mM Tris pH=7.6, 1mM EDTA). All washing was performed at 4°C for 10 min. with constant rotation. Each washing step was performed twice with the volume of the washing buffer at 10 times the volume of the dried beads. After the final wash, equal amount of 2X Laemmli sample buffer was added to the dried beads and boiled for 10min to elute the EdU-bound fraction for western blot analysis. For mass spectrometry, beads were eluted in 100mM Tris [pH 8.0], 1% SDS, 10mM DTT by boiling at 95° for 10 min.

SDS-PAGE gel staining

Protein fraction isolated using iPOND or Dm-ChP was separated on 4~20% gradient SDS-PAGE gel using electrophoresis (constant 20mA for 75min) to maintain proteins with the size range of 10~250kDa. SDS-PAGE gel is either stained with Coomassie-blue using GelCode Blue Safe Protein Stain (Thermo Scientific) for low-sensitivity purpose, or Pierce Silver Stain Kit (Thermo Scientific) for higher sensitivity detection.

Replication timing (RT) analysis

RT profile of cultured mouse mammary epithelial cells was downloaded from http://www.replicationdomain.com/ and plotted. Previously identified deletion "hotspots" in *Chaos3* mammary tumors, and the nearby genomic regions (± 5Mbps) are studied to interrogate the location of these hotspots in relation to the RT program of the nearby genomic region.

3.4. Results

SV40 immortalization rescues MCM2-7 expression and growth defects of *Chaos3* MEFs

Primary *Chaos3* MEFs senescence prematurely in culture due to elevated level of endogenous RS. Immortalized *Chaos3* and WT littermate MEFs cell lines were established using SV40 large T antigen transformation to study the biochemical characters of *Chaos3* containing mutant helicase. The established immortalized *Chaos3* MEFs displayed no signs of growth defects comparing to their WT littermate lines (Figure 3.1A). EdU incorporation assay also suggested that similar proportions of the WT and *Chaos3* populations were engaged in DNA replication (Figure 3.1B). Thus, SV40-induced cell transformation completely rescued growth defects found in the primary *Chaos3* MEFs.

Accelerated MCM2-7 pan-reduction was observed in the primary *Chaos3* MEFs comparing to their WT littermates as a function of increased passaging in culture. This reduction

(40~60%) happens at both the mRNA and protein levels. As we established that MCM reduction is closely related to senescence and cell growth in a *Trp53*-dependent manner, we wanted to determine whether SV40 induced transformation also rescues MCM2-7 reduction in the immortalized *Chaos3* MEFs. Consistent with the observation that SV40 transforms cells in part by inhibiting *Trp53* activity (267), we found that both the *Mcm2-7* mRNA and protein expression was rescued in the immortalized *Chaos3* MEFs (Figure 3.1C & D). However, many of the MCMs showed slightly reduced chromatin association, probably due to MCM2-7 complex instability (250).

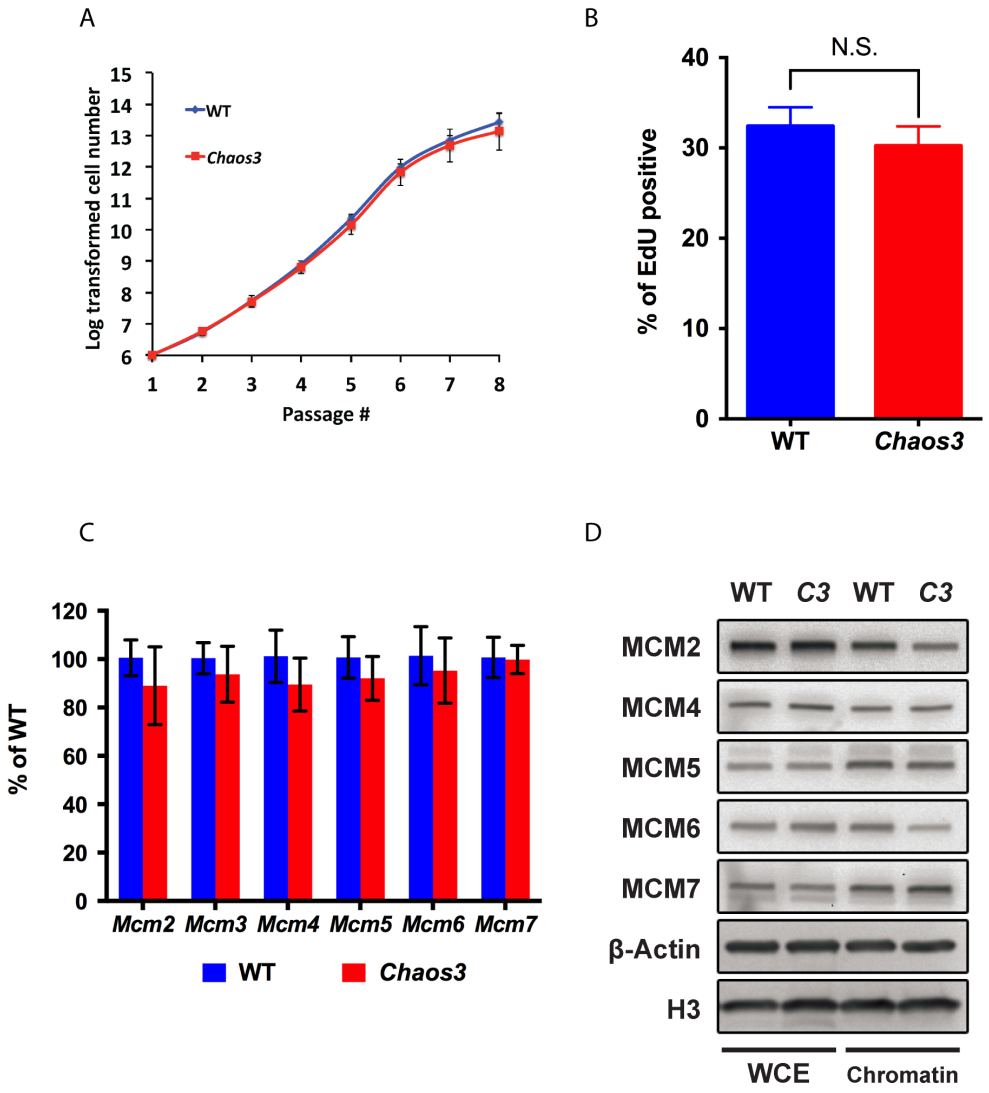


Figure 3.1 SV40 transformation rescued growth defects and MCM2-7 pan-reduction in *Chaos3* MEFs. (A) Replicative lifespans of SV40 immortalized *Chaos3* and WT MEFs. Error bar = SEM. Growth of 3 WT and *Chaos3* littermate pairs were shown. **(B)** Percentage of EdU positive cells after a 30min pulse labeling for both the WT and *Chaos3* immortalized MEFs. Littermate pairs studied in (A) was evaluated here. N.S. = not significant (two-sided t-test, p = 0.477). **(C)** *Mcm2-7* mRNA levels were rescued in *Chaos3* MEFs after SV40 immortalization as shown by qRT-PCR. Error bar = SEM. **(D)** SV40 large T antigen immortalization of primary MEFs rescues MCM2-7 expression in *Chaos3* (C3) cells. Shown are western blots of whole cell proteins extract (WCE) and chromatin-bound (chromatin) fractions isolated from the WT and littermate *Chaos3* immortalized MEFs. The slightly lower level of chromatin-bound MCM2 & 6 in C3 cells likely reflects helicase destabilization or an excess amount of these subunits relative to the others. β-Actin and histone H3 serve as loading controls for total and chromatin-bound proteins, respectively.

Optimized EdU incorporation into immortalized MEFs

Both the iPOND and Dm-ChP procedures rely on EdU pulse labeling of replicating cells (177, 263). A large number of EdU incorporated cells are needed for either methodology thus a sufficient amount of nascent DNA associated proteins could be isolated for downstream analysis. To balance this with the technical challenge of performing EdU pulse labeling and subsequent processing of all the samples in parallel, it is practical to maximize the cell density within a few tissue culturing vessels without severely compromising EdU incorporation efficiency. Thus, I performed experiment to optimize EdU incorporation when different initial cell plating densities were achieved.

Increased cell plating density indeed reduced the percentage of cells positive for EdU incorporation as expected, while low cell density also negatively affected EdU labeling (Figure 3.2A). However, higher initial plating density also resulted in larger final cell number yields. Finally, $4x10^6$ / 100mm culturing dish ($7.27x10^4$ / cm^2) initial plating density was selected for all the downstream iPOND and Dm-ChP experiments to maximize the absolute number of EdU labeled cells.

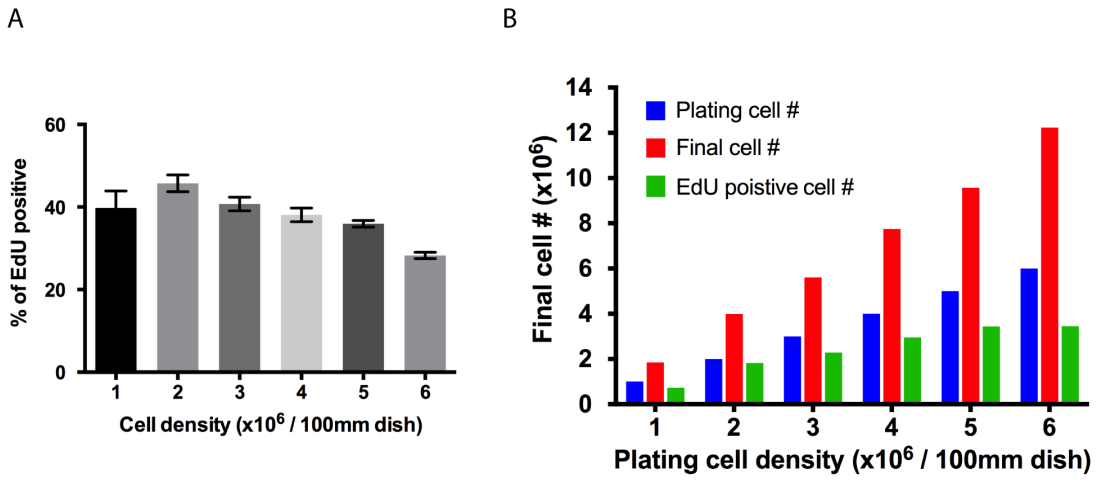


Figure 3.2 Optimize EdU incorporation at different initial cell plating density. (A) Immortalized MEFs were plated at the indicated density 20h before experiment. EdU pulse labeling was performed for 30min. Percentage of EdU positive cells are calculated after standard EdU detection. Error bar = SEM. **(B)** Cell number count for initial number of cells plated (Plating cell #), and cell number 20h after initial plating (Final cell #). Number of EdU positive cells was estimated based on final cell count and the percentage of EdU positive cells shown in (A).

Isolation of MCM protein associated with replication fork using iPOND and Dm-ChP

Both iPOND and Dm-ChP were performed to isolate nascent DNA associated proteins. Despite successful isolation of known replication proteins such as PCNA using iPOND with various EdU pulse labeling periods, we failed to detect MCM protein in the same iPOND purified fractions (Figure 3.3A). This observation is consistent with the original iPOND publication reporting the absence of MCM protein after iPOND enrichment (177). It was further supported by another report published by the same group using mass-spectrometry to analyze iPOND isolated replication proteins (264). Interestingly, another group that also employed iPOND and mass-spectrometry was able to identify many members of the MCM2-7 complex, suggesting minor alterations to this methodology might change the isolated protein spectrum (266).

Since Dm-ChP was capable of isolating MCM proteins associated with the replication forks (263), this method was further used. Similar to iPOND results, Dm-ChP was able to isolate many known replication and nascent DNA associated proteins (i.e. PCNA and histone H3, Figure 3.3B). Replication fork stalling could lead to immediate accumulation of γH2AX in its close proximity as observed using iPOND (177), which was recapitulated by Dm-ChP (Figure 3.4B). Moreover, we were able to detect MCM proteins such as MCM5 & 7 using western blotting after Dm-ChP isolation (Figure 3.3B, 3.4B). Immunoblot detection of other MCMs on Dm-ChP isolated samples was not successful, as the immunochemistry detected bands often appeared in a very strong "smeary" background (Figure 3.3B). We ruled out the possibility that it was due to protein degradation during Dm-ChP sample processing, since cell lysate prepared and immediately analyzed after the same formaldehyde treatments (without Dm-ChP procedure) showed similar western blot results for MCMs. Strong denaturing sample preparation condition

(8M urea) was not able to reverse the background noise issue linked to formaldehyde crosslinking (data not shown). This strong background was neither linked to immunochemistry detection, as formaldehyde treated cell lysate resolved on a SDS-PAGE gel also showed "smeary" band migration patterns after direct gel staining (Figure 3.3D). Proteins with a size over 100kDa was mostly affected, as indicated by both the SDS-PAGE gel staining and western blotting on several proteins with different sizes.

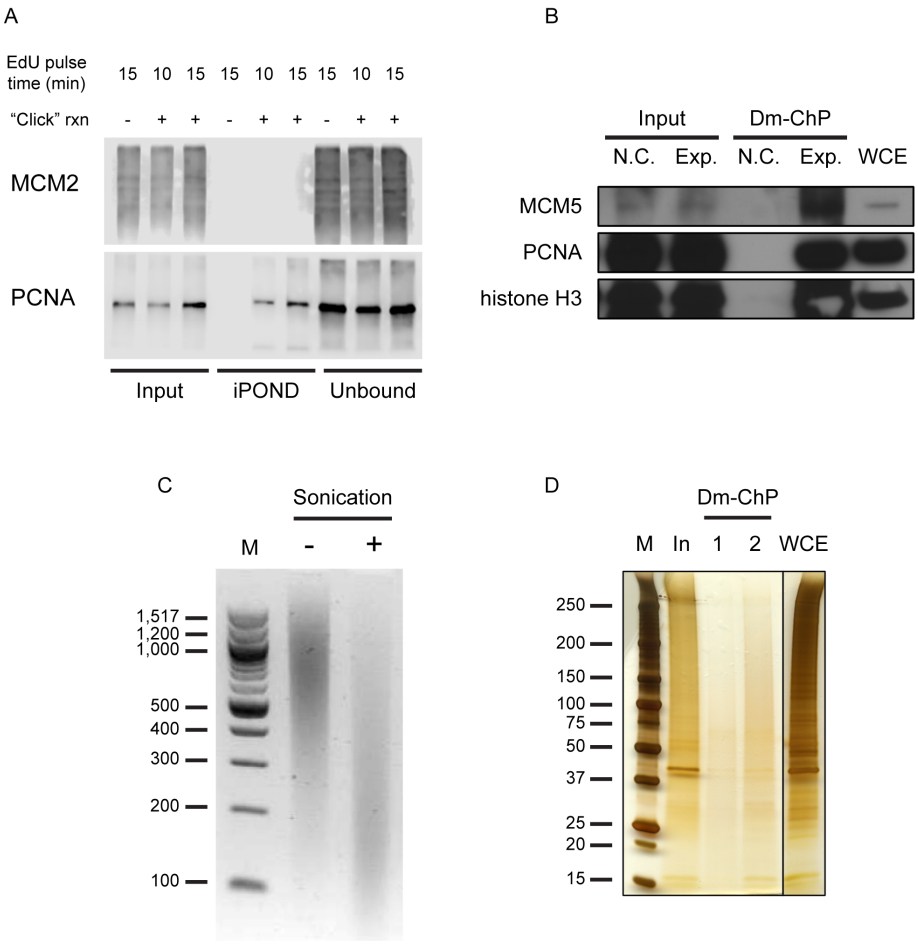


Figure 3.3 Isolation of nascent DNA associated proteins using iPOND and Dm-ChP. (A) iPOND was able to isolate known replication proteins such as PCNA with various EdU pulse labeling times, but not MCM2. Processed lysate was directly analyzed as in "Input" lanes, or incubated with streptavidin beads for affinity purification. Purified fraction was analyzed in "iPOND" lanes, while the unbound lysate was analyzed in "Unbound" lanes. **(B)** Analysis of Dm-ChP isolated proteins using western blot. Both experiment (Exp.) and negative control (N.C.) samples showed immuno detection of nascent DNA associated proteins such as MCM5, PCNA and H3 in the input lanes. In the negative control sample, no "Click" reaction was performed, thus no signal detected in the Dm-ChP purified samples. Regular whole cell extract (WCE) was also loaded as a control to demonstrate the expected position of antibody detection. Note that formaldehyde crosslinking produced strong background signal comparing to WCE. **(C)** Chromatin DNA isolated during Dm-ChP sample preparation before (-) or after (+) chromatin fractionation using sonication. Sonicated samples showed chromatin fragment size around 200~300bp. **(D)** Silver staining of SDS-PAGE with the Dm-ChP prepared samples resolved through electrophoresis. Processed lysate was directly analyzed before (In) or after (Dm-ChP lanes, 1 & 2) streptavidin beads incubation. Regular whole cell extract (WCE) was also loaded as a control. Distinct protein bands can be seen in the WCE lane, but not in any of the other formaldehyde treated samples, especially for the protein over 100kDa.

***Chaos3* mutation disrupts MCM association with normal and stalled replication forks**

As described in the previous chapter, MCM2-7 pan-reduction is not the primary cause, and rather a consequence of RS in the *Chaos3* mutant cells. The primary cause, however, is directly linked to the *Chaos3* mutation *per se,* which causes MCM2-7 complex instability *in vitro*. Since MCM2-7 complex integrity at replication forks is essential for replisome progression (98, 241), we hypothesized that the replication defect in the *Chaos3* cells is due to MCM2-7 helicase instability that compromises its association with replication forks. *In vivo* evidence supported this hypothesis, as increased incidence of spontaneous replication fork stalling was observed in *Chaos3* cells (251). To further test this idea, we used Dm-ChP to isolate and compare replication proteins associated with nascent DNA in both the WT and *Chaos3* cells. We circumvented the issue with immunochemistry detection of many MCMs due to formaldehyde crosslinking by taking a proteomic approach. SV40 immortalized WT and *Chaos3* MEFs were first subjected to stable isotope labeling of amino acid in culture (SILAC) to enable quantitative mass spectrometry analysis after Dm-ChP isolation. We found that MCM2-7 association with nascent DNA was consistently reduced in the *Chaos3* samples (~50% reduction from two independent experiments; Figure 3.4A). Many of the other known replication proteins were also identified in this analysis, and their levels were similar between the WT and *Chaos3* samples. These observations indicate that the *Chaos3* mutation compromises MCM2-7 heterohexamer association with active replication forks *in vivo*.

Since licensing is prohibited during S phase, MCM2-7 maintenance at stalled replication fork is important for proper fork recovery / restart (98, 241). A proportion of the spontaneously stalled replication forks in *Chaos3* cells was not processed and persisted into M phase (251). This observation suggested collapsing of stalled replication forks in the *Chaos3* cells, probably

due to failure of MCM helicase complex maintenance. To test this, we assessed MCM2-7 retention at stalled replication forks by first pulse labeling immortalized MEFs with EdU, followed by a high concentration of hydroxyurea (HU) to induce replication fork stalling, before replication protein isolation using Dm-ChP. We chose a 30min HU treatment to stall replication forks, as it is sufficient to block replication fork movement and induce fork stalling and collapse as demonstrated by a similar method (177). In both WT and *Chaos3* cells, HU caused an increase in γH2AX and decrease of PCNA association, consistent with replication fork stalling (177). A small fraction of MCM7 dissociated from stalled replication forks in WT cells (Figure 3.4B, WT Dm-ChP lanes, "-" *vs*. "+" HU, 25%), which is also reported using a similar method under similar experimental conditions (268). However, *Chaos3* cells suffered an even bigger loss of MCM7 retention after replication fork stalling (Figure 3.4B, *Chaos3* Dm-ChP lanes, "-" *vs*. "+" HU, 47%), on the base where MCM7 association with unchallenged replication fork is already severely reduced (Figure 3.4B, WT *vs*. *Chaos3* Dm-ChP lanes, "-" HU). Altogether, these results suggest that *Chaos3* mutation disrupts MCM2-7 complex integrity, thus its association with normal replication forks, and retention at stalled replication forks.

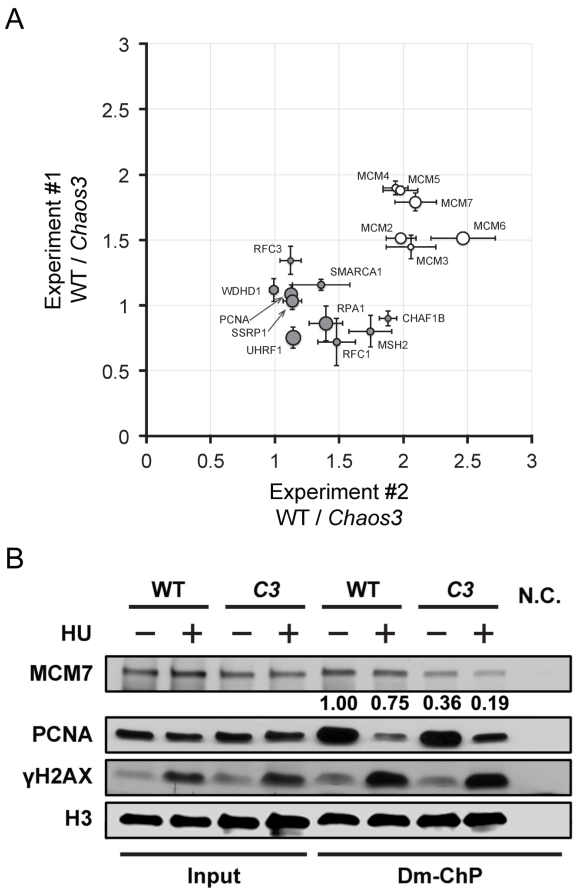


Figure 3.4 The *Chaos3* mutation disrupts MCM2-7 association with replication forks. (A) Decreased MCM2-7 association at unchallenged replication forks in *Chaos3* mutant MEFs. SILAC labeling was performed in immortalized WT and *Chaos3* littermate MEFs followed by Dm-ChP and mass spectrometry. The data represents two experiments, each involving two different cell lines for each genotype. The isotope labeling was reversed in the two SILAC experiments. Relative protein quantity ratios (WT/*Chaos3* cells) were plotted for both experiments. The size of each circle represents the relative peptide enrichment overall in the MS analysis. Open circles are the six MCM2-7 proteins; filled circles are other indicated replication proteins. Error bar = SEM for each individual experiment. **(B)** The *Chaos3* mutation causes MCM2-7 disengagement from replication forks under RS. Shown are western blots containing proteins isolated before (Input) or after Dm-ChP. SV40 T antigen-immortalized cells were pulsed with EdU for 30min to label the replication forks, and then treated with 3mM HU for another 30min to arrest replication forks. The numbers are generated after normalizing MCM7 to H3 loading control. WT cells lose minor fraction of MCM7 at stalled replication forks (WT Dm-ChP lanes, with or without HU treatment, from 1.00 to 0.75, 25% reduction). However, *Chaos3* cells showed bigger dissociation of MCM7 (*Chaos3* Dm-ChP lanes, with or without HU treatment, from 0.36 to 0.19, 47% reduction) on the base of already reduced MCM7 association with normal replication forks (WT and *Chaos3* Dm-ChP lanes, without HU treatment, from 1.00 to 0.36, 64% reduction). C3 = *Chaos3*. N.C. = no 'Click' reaction performed during Dm-ChP.

Chaos3 mutation disrupts helicase unwinding activity *in vivo*

In normal cells, disruption of coordinated DNA replication helicase and polymerases progression through polymerase stalling leads to excessive helicase unwinding and single stranded DNA (ssDNA) exposure, which becomes vulnerable to genotoxic attack (171). As expected, disengagement of MCM from stalled replication forks in the *Chaos3* cells would completely disrupt helicase function, as visualized by immunofluorescence analysis of cells that were pulse-chased with BrdU. Cultured proliferating cells constantly incorporated BrdU into double stranded genomic DNA during DNA replication. Normal helicase unwinds genomic DNA continuously after polymerase stalling and generates ssDNA track between helicase and stalled polymerases, which can be detected by anti-BrdU immunofluorescent staining. However, BrdU incorporated into the dsDNA will not be stained under non-denaturing conditions. Thus, the proportion of proliferating cells that displays detectable BrdU staining foci after replication fork stalling under native BrdU staining condition reveals the degree of helicase activity (269). Unlike WT cells, a significant fraction of the replicating *Chaos3* cells failed to display detectable ssDNA accumulation after HU-induced polymerase stalling (Figure 3.5B), consistent with the observation that MCM protein further dissociate from stalled replication forks (Figure 3.4B).

DNA replication timing program regulates the periodical replication of different genomic domains throughout S-phase. Different replication domains can be visualized using BrdU incorporation and fluorescent microscopy (27). Based on the distinguished BrdU staining patterns (Figure 3.5C), we can determine when the *Chaos3* helicase function was most affected during S phase using the same BrdU pulse-chase *in vivo* helicase function assay. Helicase unwinding through early replicated domains was moderately disrupted in both WT and *Chaos3* cells to a similar degree after replication fork stalling. However, helicase function during

mid/late S-phase replication was greatly compromised in *Chaos3* cells, but not in WT cells (Figure 3.5D). These results suggested that *Chaos3* mutation disrupts helicase unwinding through or near mid/late replicated genomic regions. Since lack of helicase unwinding activity is due to loss of MCM2-7 association with replication forks in *Chaos3* cells, *Chaos3* mutant helicase may stall near mid/late replicated replication domains, which further directly contributes to DNA damage through replication fork collapsing or leads to incomplete DNA replication.

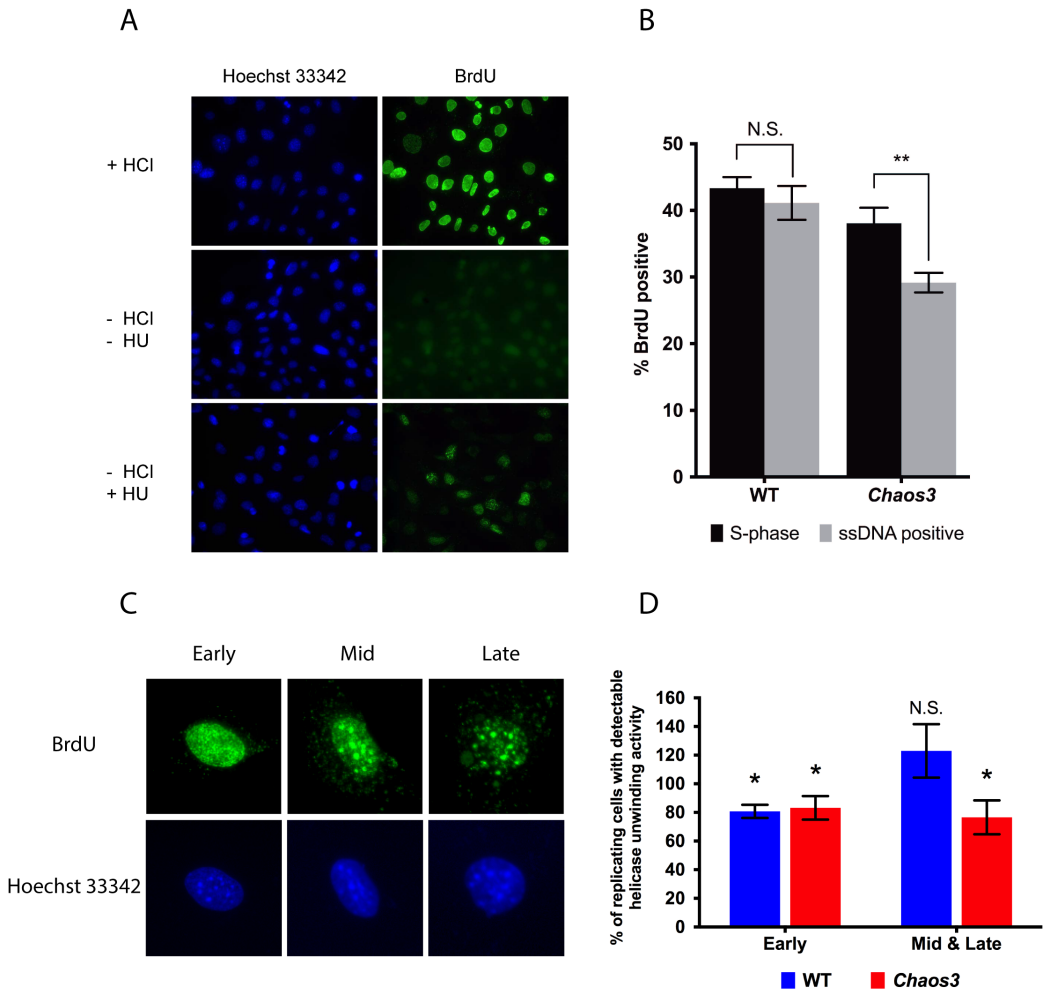


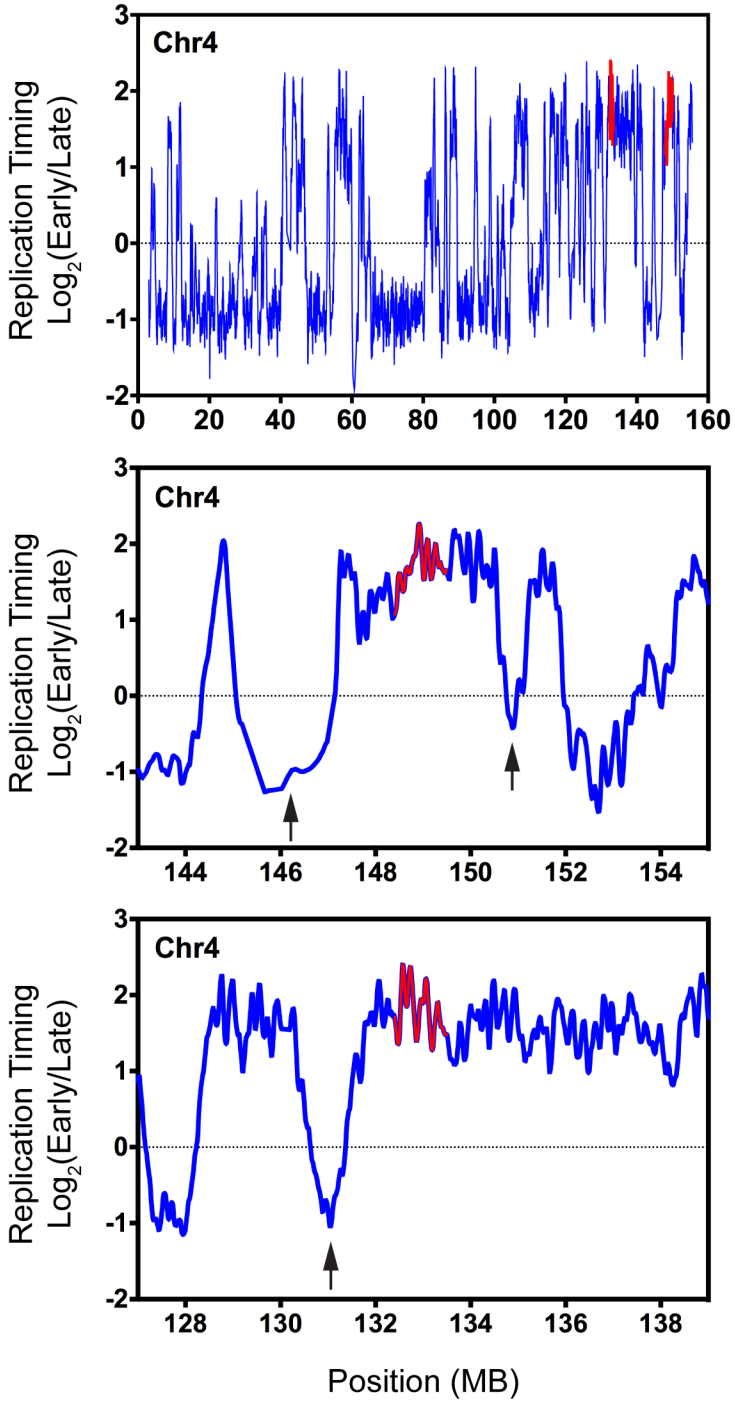
Figure 3.5 *Chaos3* mutation disrupts helicase unwinding activity. (A) Representative image of helicase exposed ssDNA in cells detected with anti-BrdU following HU stalling replication forks. **(B)** *Chaos3* disrupts DNA replicative helicase function *in vivo* following polymerase arrest. This assay is designed to assess the fraction of cells bearing replication forks that have helicases capable of continued unwinding upon stalling of the replicative polymerase. Immortalized cells of the indicated genotypes were labeled with BrdU for up to 72 hours, followed by HU treatment before immunostaining for ssDNA that contains BrdU. The fraction of untreated, BrdU-positive cells reflects cells that are actively replicating (S-phase), and this did not differ between WT and *Chaos3*. HU-treated cells that have detectable BrdU foci contain a sufficient amount of functional helicase that decoupled from the polymerases and continued unwinding (ssDNA positive). WT cells did not show a significant loss in helicase function, but *Chaos3* cells did (two-sided t-test, $p \approx 0.003$). **(C)** Sample images of different BrdU staining pattern due to replication timing (Early, Mid & Late). **(D)** Percentage of replicating cells that displays detectable helicase unwinding activity after HU induced polymerase stalling. Error bar = SEM. * = $p<0.05$, two-sided t-test. N.S. = not significant.

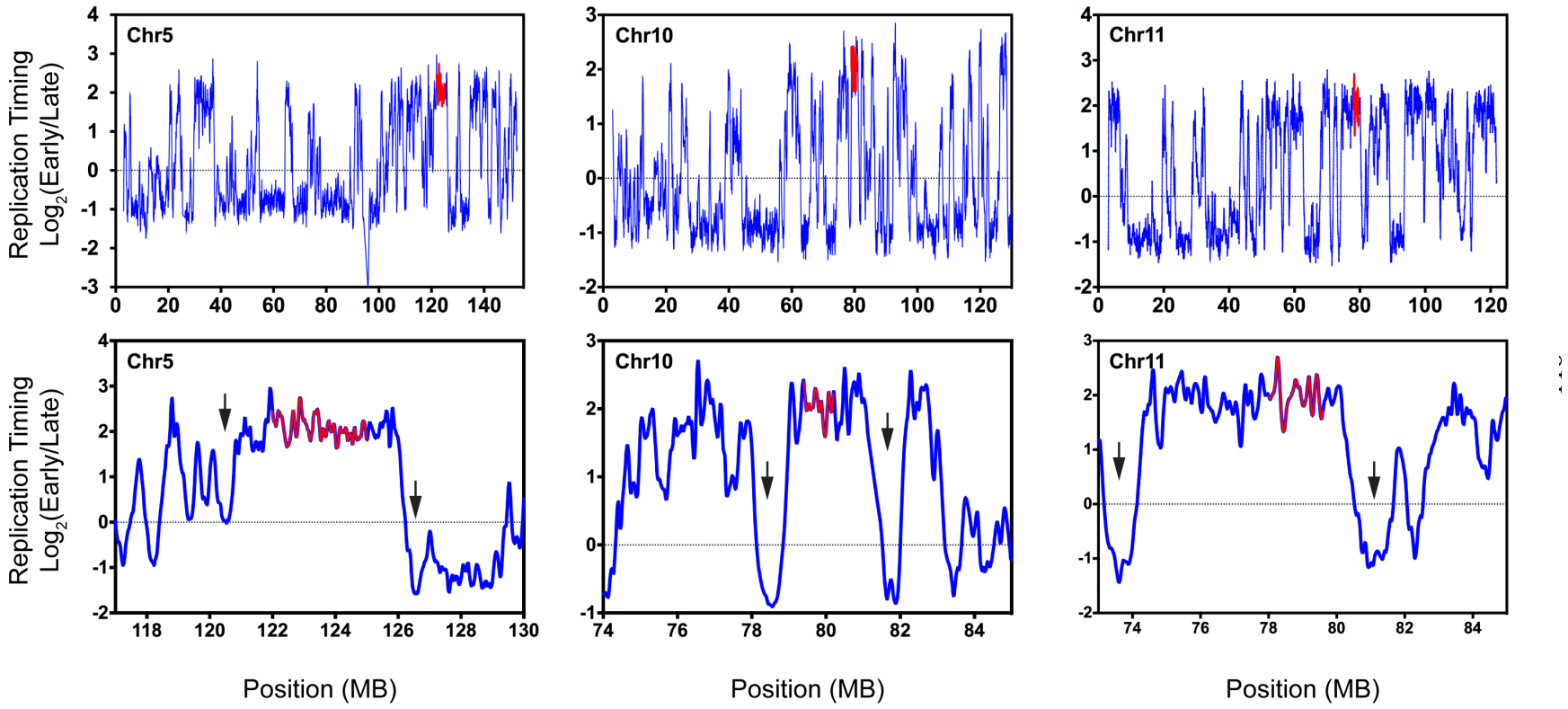
Disrupted *Chaos3* helicase function contributes to chromosome instability and oncogenesis

Recurrent mutations at the chromosomal level, such as deletions were frequent in *Chaos3* mammary tumors, which also overlap with regions of copy number alterations (CNAs) found in human breast cancer. Many genes that can suppress cancerous transformation and/or the maintenance of tumor growth reside in these deleted regions, suggesting partial or complete loss of these genes may contribute to malignant transformation and promote tumor growth in the *Chaos3* mutant (260). Since *Chaos3* mutation disrupts normal helicase stability and function, especially through or near mid/late replicated genomic regions, I suspected that replication fork stalling within certain genomic regions due to *Chaos3* helicase defects may cause loss of genomic information and contribute to chromosome / genomic instability in this mutant. To understand the replication dynamics of frequently deleted genomic regions found in the *Chaos3* mammary tumors, I studied the DNA replication timing profile of these regions. Replication timing program differs in different cell types. Based on the DNA replication timing profile of cultured mouse mammary epithelial cells where the mammary tumor originates (270), commonly deleted genomic regions in primary *Chaos3* mammary tumors are replicated early during S phase, whereas closely flanked (within 5Mb) by late replicated regions (Figure 3.6). A recent study had demonstrated that timing transition regions (TTR) links the early and late replicated domains, which often shares topological chromosome structures similar to late replicated domains (38). Furthermore, late replicated domains often display repressive chromosome signatures such as histone modifications that are associated with the silenced transcription. Since *Chaos3* helicase stability is disrupted, it might have difficulty to replicate through the mid/late replicated domains that have “closed” chromosome structure. The *Chaos3* mutant helicase may also have higher tendency to stall and collapse in these regions, thus leading to DNA damage and incomplete

DNA replication which may propagate into genomic instability and promotes malignant transformation.

A


Chr4
Replication Timing
Log2(Early/Late)
3
2
1
0
-1
-2
0
20
40
60
80
100
120
140
160
Chr4
Replication Timing
Log2(Early/Late)
144
146
148
150
152
154
Chr4
Replication Timing
Log2(Early/Late)
128
130
132
134
136
138
Position (MB)


Chr5
Chr10
Chr11
Replication Timing
Log2(Early/Late)
Position (MB)

B

Figure 3.6 DNA replication timing around the recurrently deleted genomic regions in primary *Chaos3* mammary tumors. (A) Replication timing profile of the entire chromosome 4 (top panel) of cultured mouse mammary epithelial cells (C127). Genomic regions replicated early during S phase are above 0 line while late replicated regions are below 0. The recurrently deleted regions found in *Chaos3* mammary tumors are highlighted in red. The bottom panels display the zoomed in view of the replication timing profile for genomic regions closely flanking the recurrent deleted regions (up and downstream 5Mb). Arrows mark late replicated regions. **(B)** Similar display of replication timing profile of recurrently deleted regions found on chromosome 5, 10 & 11 in the primary *Chaos3* mammary tumors. Replication timing dataset was downloaded from http://www.replicationdomain.com/. Recurrent deleted *Chaos3* mammary tumor regions are found in (260).

3.5. Conclusions and Discussion

In this chapter, I discussed how the *Chaos3* mutation causes DNA replication defects by destabilizing DNA replication helicase association with active and stalled replication forks. Using DNA mediated chromatin pull-down (Dm-ChP), it was the first time to study *Chaos3* containing MCM2-7 complex stability during DNA replication *in vivo*. In addition, I also used BrdU pulse-chase based experiment to demonstrate that *Chaos3* helicase unwinding activity is compromised due to MCM2-7 complex instability, which further supports the Dm-ChP observations.

The F345 amino acid in mice is highly conserved throughout eukaryotes, which can also be found in the archaeal MCM protein (135). The equivalent F391 residue in budding yeast is 15 amino acids downstream of the conserved zinc finger domain, which is essential for MCM interactions, MCM2-7 complex formation and cell viability (78, 81). Based on the crystal structure of *Methanothermobacter thermautotrophicus* MCM (271), the homologous F171 amino acid sits on the MCM interactive surface. The latest Cryo-EM imaging provided high resolution mapping of the budding yeast MCM2-7 complex structure to near atomic level (92). Similar to archaeal MCM, the *Chaos3* mutated residue in yeast also reside between two interacting MCMs (MCM4 & 6, Figure 3.7), suggesting that this mutation may alter the physical interaction between the MCM subunits, thus disrupt MCM2-7 complex stability. This hypothesis was confirmed by two independent *in vitro* analyses interrogating MCM2-7 complex stability (249, 251). These observations suggest that the *Chaos3* mutation may disrupt MCM2-7 complex stability and thus helicase function. Contradictory to this hypothesis, replication fork progression was not affected in the *Chaos3* cells *in vivo*, and the recombinant CMG helicase complex bearing the *Chaos3* mutation had even better performance in terms of helicase unwinding activity *in vitro*

comparing to the WT equivalent (251). Based on these results, it was concluded that *Chaos3* mutation did not affect helicase function and it was reduced MCM2-7 association with chromatin due to lowered expression of these proteins in the *Chaos3* mutant led to compromised dormant origin licensing, thus causing cellular growth defects, genomic instability and cancer predisposition in this mutant (249-251, 256).

Many lines of evidences, however, argue against this conclusion. For example, the growth defect of *Chaos3* primary cells was not due to MCM2-7 repression. Reduced MCM2 expression by at least 50% or to an even larger extend failed to compromise normal cellular proliferation (130). Genomic instability is a hallmark of the *Chaos3* mutant, as indicated by an over 25-fold increase in micronuclei (MN) formation in the homozygous mutant animals comparing to the WT littermates (135). However, combined genetic removal of MCM dosage in the animals with MCM gene-trap mutations only developed 2~5 fold increase of MN, which is similar to *Chaos3* heterozygous mice that have no MCM2-7 reduction (135, 249). Finally, none of the MCM gene trap mutants developed cancer predisposition phenotypes as the *Chaos3* animals, despite they have comparable levels of MCM reduction (249). These arguments proposed that such level of reduced MCM2-7 expression *per se* in the *Chaos3* mutant is not sufficient to cause all the defects at cellular and animal levels. Other characteristics of the *Chaos3* containing MCM2-7 complex must be the major cause of DNA replication defects and highly elevated genomic instability in this mutant. Besides, as discussed in the previous chapter that active MCM2-7 repression is tightly linked to cellular response to RS through *Trp53* centered DDR mechanisms. The progressive loss of MCM2-7 expression due to chronic exposure to RS is accelerated in the *Chaos3* mutant. Thus, reduced MCM2-7 expression is a consequence, rather than the cause of RS in the *Chaos3* mutant. However, it may still contribute

to genomic instability, as reduced MCM expression will indeed compromise optimal replication licensing.

CMG complex containing the *Chaos3* mutation was expressed and assembled *in vitro*, which demonstrated similar to higher helicase activity comparing to its WT parallel, led to the conclusion that *Chaos3* helicase function was not affected in this mutant (251). However, this assay was performed on circular naked DNA *in vitro*, which cannot fully depict the *in vivo* situation when replication helicase has to unwind DNA under different chromosome contexts. Thus, we decided to directly isolate and analyze MCM associated with the replicating DNA *in vivo* using the newly developed iPOND and Dm-ChP procedures. These two methods, together with nascent chromatin capture (NCC) are all developed around the similar ideology, whereas nascent DNA synthesized at the replication forks is labeled with nucleotide analogue. Both iPOND and Dm-ChP use EdU, while NCC relies on biotin-dUTP. All the three methods utilize the strong affinity between streptavidin and biotin to perform the final enrichment of nascent DNA. Thus, both iPOND and Dm-ChP require an extra step to conjugate biotin molecule to the EdU labeled nascent DNA through the highly efficient "Click" reaction, which can be performed under mild chemical conditions to minimize protein degradation and disruption to replisome integrity during sample processing. Mass spectrometry (MS) can be coupled with all of these three methods to comprehensively analyze isolated nascent DNA associated proteins. We took advantage of the SILAC labeling to differentiate the WT and *Chaos3* samples, which facilitates quantification and comparison of Dm-ChP isolated replication proteins in a single MS analysis. The SILAC-MS procedure also circumvented the problem with immunochemistry detection after formaldehyde crosslinking. Using the Dm-ChP method, we found loss of MCM2-7 association with nascent DNA under normal replication conditions in *Chaos3* cells. MCM2-7 destabilization

leads to immediate DNA replication arrest (98), thus this observation explains the increased frequency of spontaneous replication fork stalling observed in the unchallenged *Chaos3* cells (251). Additional RS that stalls replication forks further disrupts MCM association. Since MCM maintenance at stalled replication structure is essential for replication fork restart, this result further confirms the *in vivo* observation that stalled replication forks persist into M phase in *Chaos3* cells (251). These results were also reinforced by the BrdU pulse-labeling assay, suggesting loss of *Chaos3* helicase unwinding activity.

Chaos3 mutation disrupts genome integrity through DNA replication and causes mutations that lead to cellular transformation at susceptible genomic loci. *Chaos3* mutant modeled in budding yeast recapitulated the mouse model, whereas the mutant strain acquired aneuploidy and improved growth through DNA replication. A few mutations due to chromosomal aberrations acquired by the mutant yeast strain were responsible for cancer-like improved growth, while the chromosome breakpoints often associate with certain chromosome structures (272). Many genomic regions are recurrently deleted in the *Chaos3* primary mammary tumors in mice, which contain genes essential for the prevention of malignant transformation and maintenance of tumor growth (260). Though these regions are replicated early during S phase, they are in close proximity to late replicated regions in the mammary epithelial cells, where the mammary malignancies are originated. Recent study revealed that genomic positions of DNA replication domains were determined by 3-dimenstional topological arrangements of chromosomes within the nucleus (38). The boundaries between early and late replicated regions shares similar chromosome states as the late replicated domains, which also diffuse into the early replicated domains. Furthermore, repetitive DNA sequences, which were often recognized as a source of RS, were also enriched within these boundaries. Finally, late replicated domains are

often arranged into repressive nucleus compartments. After all, many signatures that are associated with RS were found in or near late replicated domains. By studying the BrdU staining patterns in the BrdU pulse-labeling experiment, I was able to determine that the *Chaos3* helicase unwinding through or near mid / late replicated genomic regions were mostly affected. Since replicative helicase is the pioneering component of the replisome during replication fork progression, disrupted helicase complex stability may hinder replication process through "hard to replicate" genomic regions due to loss of helicase association. Thus, incomplete DNA replication may occur, and replication structure can collapse due to loss of MCM2-7 association after replisome stalling, which may all contribute to DNA damage and loss of genomic material. Eventually, *Chaos3* associated helicase defects may happen at many genomic loci, which propagates into genomic instability that promotes tumorigenesis.

In summary, *Chaos3* mutation disrupts MCM2-7 association with replication fork and thus the replicative helicase function *in vivo*. Maintenance of MCM2-7 complex at stalled replication forks is also compromised due to helicase instability. Disrupted helicase stability and function may acerbate near genomic regions that are intrinsically non-permissive for replication fork progression, leading to mutations and chromosome aberration near these regions, thus contributes to genomic instability and cancer predisposition in *Chaos3* mutant.

3.6. Acknowledgement

The authors thank Dr. F. Oliveira for guidance with the mass spectrometry experiments.

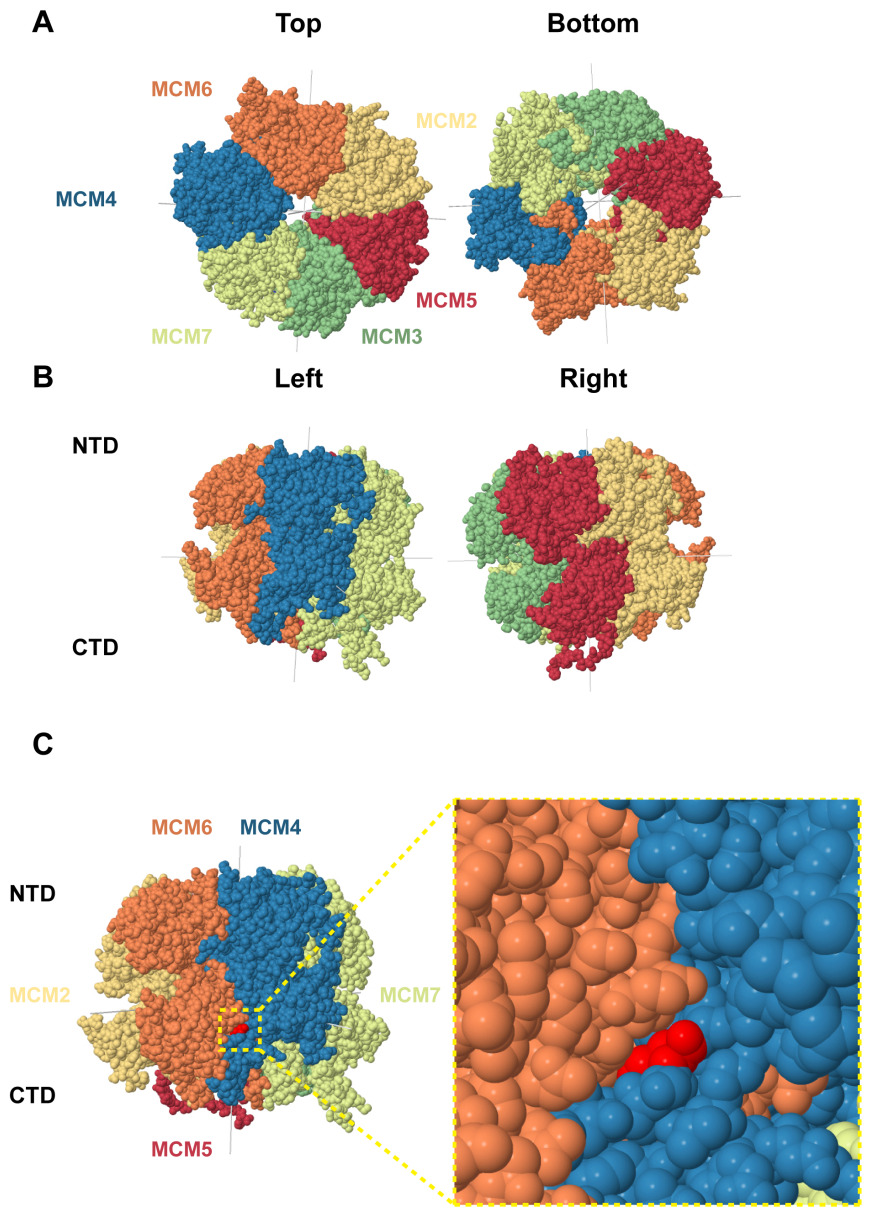


Figure 3.7 cryo-EM structure of eukaryotic MCM2-7 complex. (A) Top & bottom and **(B)** side views of *Saccharomyces cerevisiae* MCM2-7 complex (space fill model, PDB ID: 3JA8) as determined by cyro-EM with a resolution of 3.8Å (92). Each different MCM2-7 subunits are displayed in different colors. NTD = N-terminal domain; CTD = C-terminal domain. **(C)** *Chaos3* mutation modeling using MCM2-7 complex from *Saccharomyces cerevisiae*. The homologous F345 residue in mice is highlighted on the buddying yeast MCM4 protein (F391) in red. This residue resides on the MCM4:6 interactive surface. The F→I mutation does not change hydrophobic side chain character of the amino acid, however, the resulting isoleucine (I) residue lacks the bulky side chain found in phenylalanine (F). Besides, the benzene ring in phenylalanine appears to protrude into the neighboring MCM6 protein.

CHAPTER 4

ROLE OF MICRORNA DURING CHRONIC DNA REPLICATION STRESS[3]

Gongshi Bai[1, 2] and John C. Schimenti[1, 2, 3]

[1] Cornell University, Dept. of Molecular Biology & Genetics, Ithaca, NY 14853, USA

[2] Cornell University, Dept. of Biomedical Sciences, Ithaca, NY 14853, USA

[3] Cornell University Center for Vertebrate Genomics, Ithaca, NY 14853, USA

4.1. Abstract

Minichromosome maintenance 2-7 (Mcm2-7) are essential eukaryotic DNA replication genes, which are highly expressed in proliferating cells. The six subunits form heterohexameric complexes that are loaded onto chromatin and license the genome to become competent for DNA replication. MCM2-7 complex also serves as the core of replication helicase, which unwinds double stranded DNA to provide single stranded DNA template for semi-conserved DNA replication. Mutations that disrupt normal helicase function or MCM expression can lead to DNA replication defects, genomic instability and cancer. Pre-malignant tissues and cancer cells also express high level of MCM2-7 proteins. MCM2-7 expression can be actively repressed in normal cells experiencing replication stress (RS) that constantly arise during DNA replication. This regulation is connected to normal DNA damage response (DDR) through intact *Trp53* function, which can sensitize cells to additional RS by inducing terminal cell cycle arrest such as senescence. However, how MCM2-7 repression is achieved is currently unknown. Here we report the discovery of a group of *Trp53*-dependent miRNAs that represses MCM2-7 expression in response to RS. Modulation of the expression of these miRNAs significantly reduced RS

---

[3] Work presented in this chapter has been submitted for publication. Authors' contributions: G.B. & J.C.S conceived & designed the research; G.B. performed experiments and analyzed data; G.B. & J.C.S wrote the paper.

associated genomic instability in mice, yet to modify tumorigensis phenotypes. It reveals important yet redundant role of miRNAs in regulating MCM expression as a cellular response to chronic RS.

4.2. Introduction

Proliferating cells constantly suffer from replication stress (RS) originated from endogenous and exogenous sources that inhibits normal replication fork progression and the duplication of the genomic information (163). Genomic lesions arise during DNA replication, are detected by DNA damage response (DDR) in normal cells, which promotes DNA damage repair and transient delay cell cycle progression. When the lesion is severe or cannot be resolved in a timely manner, DDR can also induce permanent cell cycle arrest (senescence) or programed cell death (apoptosis). Thus, in higher metazoan, DDR is essential for eliminating cells that are subject to malignant transformation and protecting the whole organism from diseases such as cancer.

Minichromosome maintenance 2-7 (MCM2-7) are essential DNA replication genes, which are involved in every aspects of eukaryotic DNA replication. MCM2-7 proteins are highly expressed in proliferating cells, which assemble into a heterohexameric complex to perform their biological functions as replication licensing factors and replicative helicase. Mutations that disrupt normal MCM function, including helicase activity and protein expression can lead to genomic instability, stem cell defects and cancer predisposition in mice and human (128, 134-136, 249, 251, 252, 259). Elevated MCM expression was found in pre-malignant tissue, which serves as a biomarker that is more sensitive than the conventional cell proliferation markers to identify the early lesions and prognosis of the disease such as cancer (132). Selective MCM2-7 reduction was also observed in hematopoietic stem cells (HSCs) isolated from old mice. MCM2-

7 repression in the HSCs limited their expansion once transplanted into animals (255). These previous observations suggested that MCM2-7 expression level is tightly connected to proliferation status and replicative lifespan of the cell. Indeed, as discussed in the previous chapters, MCM2-7 expression is actively reduced in response to chronic low level of RS, which is regulated by *Trp53* centered DDR. A moderate MCM2-7 repression does not affect normal cell proliferation and the general health of animals, while it sensitized proliferating cells to persistent RS by inducing cellular senescence. However, the mechanism that suppresses MCM2-7 expression in response to RS activated DDR is unknown.

Chaos3 mutant carries a single point mutation in the *Mcm4* gene that disrupts normal MCM2-7 complex formation *in vitro* and helicase function *in vivo,* which is the primary source of endogenous RS that promotes genomic instability and tumorigensis in this mutant mouse (135, 249, 251). Another major characteristic of this mutant is moderately reduced MCM2-7 expression (135, 249, 251), which is part of the normal cellular response to helicase instability induced RS in this mutant. Genetic ablation of *Trp53* function in the *Chaos3* mutant rescued MCM2-7 suppression (256). We also found that RS induced MCM2-7 reduction in normal primary cells depends on *Trp53* status. MCM2-7 repression in the *Chaos3* mutant is an active process, which is originated from the specific removal of *Mcm2-7* mRNAs at the post-transcriptional level. Furthermore, disruption of small non-coding RNA biogenesis, including microRNAs (miRNAs) also resulted in increased MCM2-7 expression in the *Chaos3* mutant cells (250). These observations suggested that endogenously expressed small non-coding RNAs, presumably miRNAs whose expression can be regulated by *Trp53* is responsible for moderate MCM2-7 suppression in response to RS.

miRNAs belongs to a class of the small non-coding RNAs that represses gene expression and regulate cellular functions (214). miRNAs are endogenously expressed and processed from the primary RNA transcripts by RNase III type endonucleases *Drosha* and *Dicer* (206). The "seed" sequence (2~7 nucleotides on the 5' end) of the mature small RNA product (~22nt long) recognizes target mRNAs based on Watson-Crick pairing, which recruits the miRNA induced silencing complex (miRISC) to repress gene expression post-transcriptionally (219). miRNA targeting sites, also known as miRNA responsive elements (MRE) often resides in the 3'UTR of the repressed mRNA transcripts (218). A combination of the same and different miRNAs recognition of multiple MREs within the same target mRNA transcript can greatly repress its expression. It also enables differential gene expression regulation under diverse cellular context with altered miRNA expression profiles (218, 225). miRNA mediated gene expression regulation is often subtle and some times reversible (221). Cells can exploit this character of miRNA to moderately and transiently regulate gene expressions and their corresponding cellular functions without committing to terminal cell fate decisions.

Ectopic miRNA expression was found in various types of human diseases, including cancer. Altered miRNA expression can either promote or suppress cancer development (273, 274). Moreover, normal miRNA regulation is intrinsically connected to the cellular responses to various types of stresses experienced by the cells, including genotoxic stresses that generates DNA lesions and promotes oncogenesis. DDR recognizes these damage and transmits the response signals that converge on the central tumor suppressor *Trp53,* which is a transcriptional regulator of many downstream events. Expressions of many miRNAs are directly targeted by *Trp53*, which can also regulate miRNA biogenesis to regulate their functions (275-277).

In this chapter, I will discuss the discovery of a group of *Trp53* responsive miRNAs that represses MCM2-7 expression in response to chronic low level of RS. Modulation of the miRNA expressions can significantly impact on cellular defects associated with RS such as genomic instability. However, such impact failed to transform at the animal level to affect RS induced tumorigenesis, supporting the idea that miRNAs often play redundant roles during DDR and tumor suppression (273).

4.3. Materials and methods

Small RNA sequencing library preparation

Total RNA, including small RNA was extracted from primary MEFs using miRNAeasy kit (Qiagen). 1μg of total RNA from each sample is used to prepare small RNA sequencing library using the TruSeq small RNA sample preparation kit (Indexes set 1~12, Illumina) per manufacture's instruction. A total number of 24 samples were processed simultaneously. TruSeq indexes 1~12 were used for each lane. A list of the samples and their indexes added during sample preparation can be found in Table 4.1. After library construction, sample #1~12, and #13~24 are pooled into two separate lanes and sequenced in two lanes on the Illumina HiSeq platform using single-end High Output mode. Each pooled lanes was first separated on a TBE-PAGE gel to isolate the prepared library product with the proper size (between 145-160bp), and later analyzed on Experion 1kb DNA chip for quality control.

Small RNA sequencing data analysis

Illumina platform generated reads were filtered based on sequencing quality and then aligned to miRBase database (v19) to generate read counts. Counts for each miRNA was normalized to the total number of reads for each sample. miRNA with a read counts less than

100 per $1x10^6$ total reads (<0.01%) were not further analyzed. Normalized miRNA reads for each sample were used as input for data analysis using DESeq package (278).

miRNA quantitative RT-PCR (qRT-PCR)

Total RNA was extracted from cultured MEFs using miRNAeasy kit (Qiagen). cDNA of small RNA was synthesized from 1ug of total RNA using qScript microRNA cDNA Synthesis Kit (Quanta). PCR amplification and real-time detection was performed with a Bio-Rad CFX96 Real-Time system and data analysis was performed with the Bio-Rad CFX Manager software (Bio-Rad). Relative gene expression was calculated using the ddCq method with RNU6 as endogenous control. A technical replicate was performed for each sample during qPCR reaction. The primers for microRNA amplification were purchased from PerfeCTa microRNA assays (Quanta).

***In silico* prediction of miRNA targets**

To determine the potential miRNAs that target at *Mcm2-7* mRNAs, we used a combination of miRmap (http://mirmap.ezlab.org/) and miRanda (http://www.microrna.org/) software.

Transfection of miRNA mimics

Primary WT MEFs were transfected with 50nM of miRNA mimic (Dharmacon) using DharmaFECT 1 transfection reagent per manufacturer's instructions. Control cells were transfected in parallel with negative control miRNA mimics (based on cel-miR-67), which has minimal sequence similarity with miRNAs in mice. 48h after transfection, cells were harvested for RNA and protein analysis.

Luciferase assay

HeLa cells were cotransfected using Lipofectamine 2000 in a 96 well format with 50nM of miRNA mimic and 100ng of pmirGLO luciferase construct with or without a 3'UTR insert. The 3'UTRs of *Mcm2-7* were cloned into *Nhe*I + *Sbf*I digested pmirGLO vector. Cloning primers are listed in Table 4.2. Empty vector was used as control. 24h after transfection, cells were changed to fresh media. 48h after transfection luciferase activities were measured using the Dual Luciferase Assay System (Promega) and Synergy 2 Multi-Mode Reader (BioTek) per manufacturer's instructions. *Renilla* luciferase activity was normalized to *Firefly* luciferase activity in each well. The ratio of normalized *Renilla* luciferase activity between construct that bears the 3'UTR and control construct (empty vector) reflects the level of miRNA mediated suppression through its interaction with the 3'UTR.

Generation of *Chaos3* / *miR-34abc*-TKO compound mutants

Animals with *miR-34abc* knockout alleles were acquired from Dr. A. Nikitin (273). They were crossed to the *Chaos3* mutant (C3Heb/FeJ background) for at least 5 generations. Male breeders from each generation were selected based on congenic status as evaluated by the DartMouse speed congenic service. A list of genotyping primers is presented in Table 4.2.

Animal Research

The use of animals in this study was performed under a protocol (2004-0038) approved by Cornell's IACUC. Mice were euthanized via CO_2 asphyxiation according to IACUC-approved conditions.

Table 4.1 Sample information for miRNA sequencing

Lane #	Sample #	Sample ID	Index	Description
1	1	5G1-P2	ATCACG	WT MEFs, P2
	2	5G4-P2	CGATGT	*Chaos3* MEFs, P2
	3	5H5-P2	TTAGGC	*Chaos3* MEFs, P2
	4	5H6-P2	TGACCA	WT MEFs, P2
	5	WT-C1	ACAGTG	WT MEFs, treated with 0μM HU for 48h
	6	WT-C2	GCCAAT	
	7	WT-T1	CAGATC	WT MEFs, treated with 200μM HU for 48h
	8	WT-T2	ACTTGA	
	9	Chk2-C1	GATCAG	*Chk2*$^{-/-}$ MEFs, treated with 0μM HU for 48h
	10	Chk2-C2	TAGCTT	
	11	Chk2-T1	GGCTAC	*Chk2*$^{-/-}$ MEFs, treated with 200μM HU for 48h
	12	Chk2-T2	CTTGTA	
2	13	5G1-P3	ATCACG	WT MEFs, P3
	14	5G4-P3	CGATGT	*Chaos3* MEFs, P3
	15	5H5-P3	TTAGGC	*Chaos3* MEFs, P3
	16	5H6-P3	TGACCA	WT MEFs, P3
	17	5G1-P4	ACAGTG	WT MEFs, P4
	18	5G4-P4	GCCAAT	*Chaos3* MEFs, P4
	19	5H5-P4	CAGATC	*Chaos3* MEFs, P4
	20	5H6-P4	ACTTGA	WT MEFs, P4
	21	5G1-P5	GATCAG	WT MEFs, P5
	22	5G4-P5	TAGCTT	*Chaos3* MEFs, P5
	23	5H5-P5	GGCTAC	*Chaos3* MEFs, P5
	24	5H6-P5	CTTGTA	WT MEFs, P5

Table 4.2 Primers used for genotyping *miR-34abc* mutants and *Mcm2-7* 3'UTR cloning

Name	Sequence (5' to 3')	Purpose
miR-34a_A	GAATGTGTATACGTGTTTTGCCTGA	*miR-34a* genotyping
miR-34a_F	TGGCCCCTTTAATTTACAAGCCCA	
miR-34a_D	AGCTGACATGCCAGGAATGCTGA	
miR-34bc_A	CTGCGCTTCTTTCTTCGATGTAGC	*miR-34bc* genotyping
miR-34bc_B	TGGCTTTAGGATCTCCATTTCAGC	
miR-34bc_D	ACCTGGTTAAGTGGGCTGAGTTCC	
pmirGLO.Seq_F	GACGAGGTGCCTAAAGGAC	Sequencing primers for pmirGLO
pmirGLO.Seq_R	CCAACTCAGCTTCCTTTCG	
Mcm2 3'UTR F	AAATTTGCTAGCGACCAGCACAGGGGCCTC	Primers for cloning full-length *Mcm2* 3'UTR into pmirGLO
Mcm2 3'UTR R	AAATTTCCTGCAGGTAGCGCTATTTAGGTTTTATTC	
Mcm3 3'UTR F	AAATTTGCTAGCAGTTGTTGCTACCAAGTAC	Primers for cloning full-length *Mcm3* 3'UTR into pmirGLO
Mcm3 3'UTR R	AAATTTCCTGCAGGTACAAAATAGTTATATTTACTCAGTATG	
Mcm4 3'UTR F	AAATTTGCTAGCGCTGCATGGCCCTCGGAC	Primers for cloning full-length *Mcm4* 3'UTR into pmirGLO
Mcm4 3'UTR R	AAATTTCCTGCAGGAATTCAAATGGTTTAGAGATTTATTG	
Mcm5 3'UTR F	AAATTTGCTAGCGCCCATTGCCCATCAACC	Primers for cloning full-length *Mcm5* 3'UTR into pmirGLO
Mcm5 3'UTR R	AAATTTCCTGCAGGGGTGGCCCACACTTTTAATCC	
Mcm6 3'UTR F	AAATTTGCTAGCGGTATTGAAGGTAACTTGATGG	Primers for cloning full-length *Mcm6* 3'UTR into pmirGLO
Mcm6 3'UTR R	AAATTTCCTGCAGGCACAGAACAAGTTGTTTTATTTTCATG	
Mcm7 3'UTR F	AAATTTGCTAGCTAGCCAGTTTTTACACCCTCC	Primers for cloning full-length *Mcm7* 3'UTR into pmirGLO
Mcm7 3'UTR R	AAATTTCCTGCAGGAACAAGCAAGAGGCAATCAAAAC	

4.4. Results

Overview of small RNA sequencing results

Small RNA sequencing libraries were generated from total RNA samples isolated from primary WT and mutant *Chaos3* MEFs exposed to endogenous and exogenous replication stress (list of samples can be found in Table 4.1). Size selected product after library preparation was sequenced on next generation sequencing platform (Figure 4.1A). Each library generates around on average of 15 million raw reads. After quality control to filter out reads that have poor sequencing read quality or missing linker sequences after library construction, 94.1% of the reads were used as input to align to the mature mouse miRNA / hairpin database and generated number of reads for each miRNA (miRBase, mouse miRNA v19). Data analysis was performed using DESeq package designed to analyze small RNA sequencing datasets (278). Briefly, number of reads for each mature miRNA was first normalized to the total number of reads of the library, and analyzed using DESeq package to identify miRNAs whose expressions are significantly changed during RS response. Since we were trying to identify miRNAs that can regulate MCM2-7 expression in response to RS, which are highly expressed in proliferating cells, later we only considered miRNAs that are relatively abundant in each of the samples (with more than 100 reads / 1 million total reads, or more than 0.01% of the overall miRNAs expression).

To determine the miRNAs that are upregulated in response to endogenous RS, we analyzed WT and littermate *Chaos3* MEFs miRNA expression profile from total RNA isolated at each passage. We hypothesized that with increased passage number, miRNAs that repress MCM2-7 expression should be over-represented in the small RNAs samples isolated from *Chaos3* cells, which have reduced MCM2-7 expression comparing to their WT littermate

primary MEFs. However, miRNA expression profile of WT and *Chaos3* MEFs are very similar even at later passages, thus the statistical analysis revealed little information on miRNAs that are significantly misregulated. We then focused on the miRNA expression profile of WT primary MEFs after exogenous RS treatment. Total RNA from WT primary MEFs receiving 200μM HU treatment for 48h was analyzed. This time point was picked over shorter or longer HU exposure is because 1) with 48h of treatment, MCM2-7 levels started to reduce. However, cell cycle genes related to S phase entry and progression were yet to be affected (Figure 4.1C) and 2) *E2f* family transcription factors regulate MCM2-7 expression at the transcription level (279), which maintains normal expression after HU exposure comparing to the control cells (Figure 4.1D). A quick examination suggested that miRNAs belonging to the *miR-34 (miR-34b, c), miR-10 (miR-10b), miR-27 (miR-27b)* & *miR-181 (miR-181a, b & c)* families are upregulated in response to HU treatment, which are also induced in *Chaos3* cells, albeit not reaching statistical significant level (Figure 4.1E). Members of these miRNA families that have high basal expression level (*miR-34b, miR-34c, miR-181a, miR-10b & miR-27b*) were further studied to examine whether they target at MCM2-7 to repress their expression. *miR-34a* is a well characterized member of the *miR-34* miRNA family, which was demonstrated to target at multiple MCM transcripts in the context of miRISC in human cells (280). Thus, *miR-34a* was included in the following studies as a positive control.

Among the miRNAs downregulated in response to RS, *miR-17, 20a & 92a* are outstanding. Interestingly, these miRNAs are expressed from a polycistron cluster in both mouse and human genome, which have been demonstrated to have oncogenesis potentials (281). This miRNA cluster, also known as oncomiR-1, is frequently amplified or overexpressed in human tumor samples, was identified as a potent cancer driver in mice (274). Our findings that these

miRNAs are repressed in response to genotoxic treatment in normal primary cells support the idea that they may promote malignant transformation.

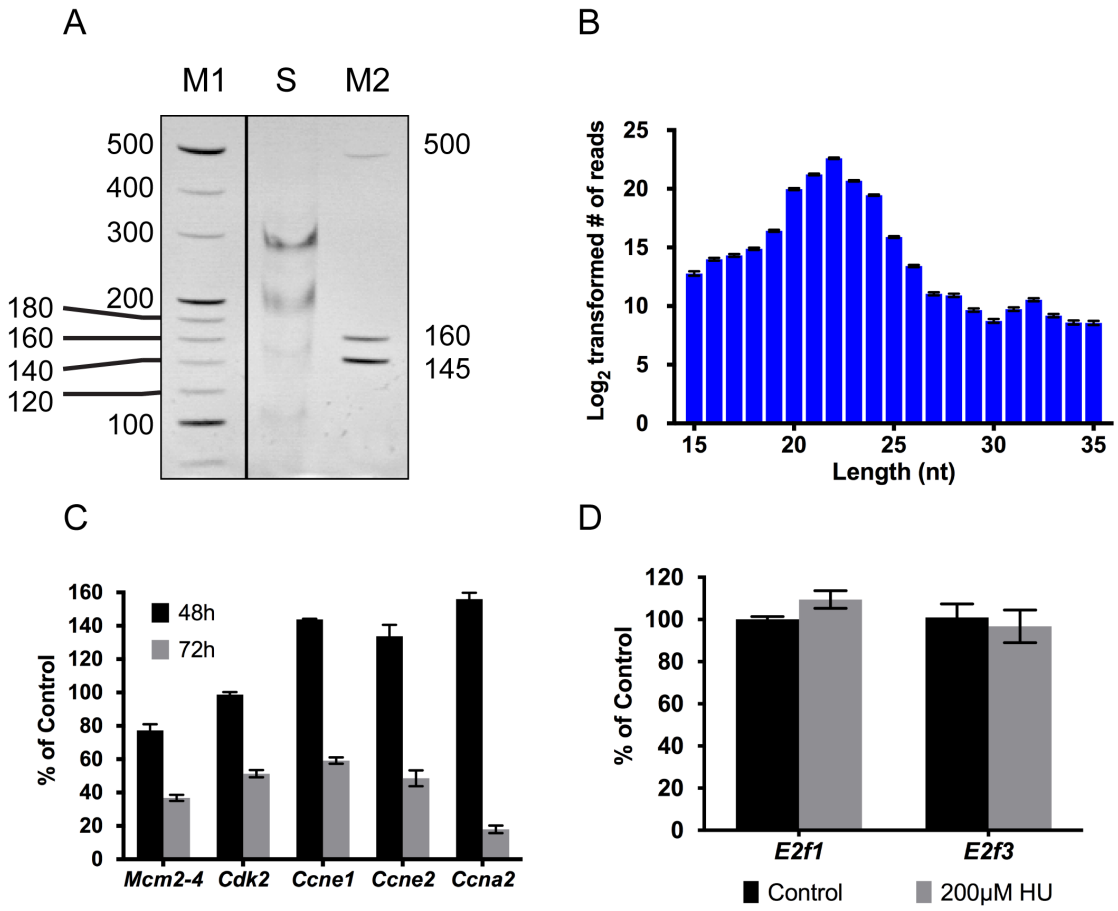
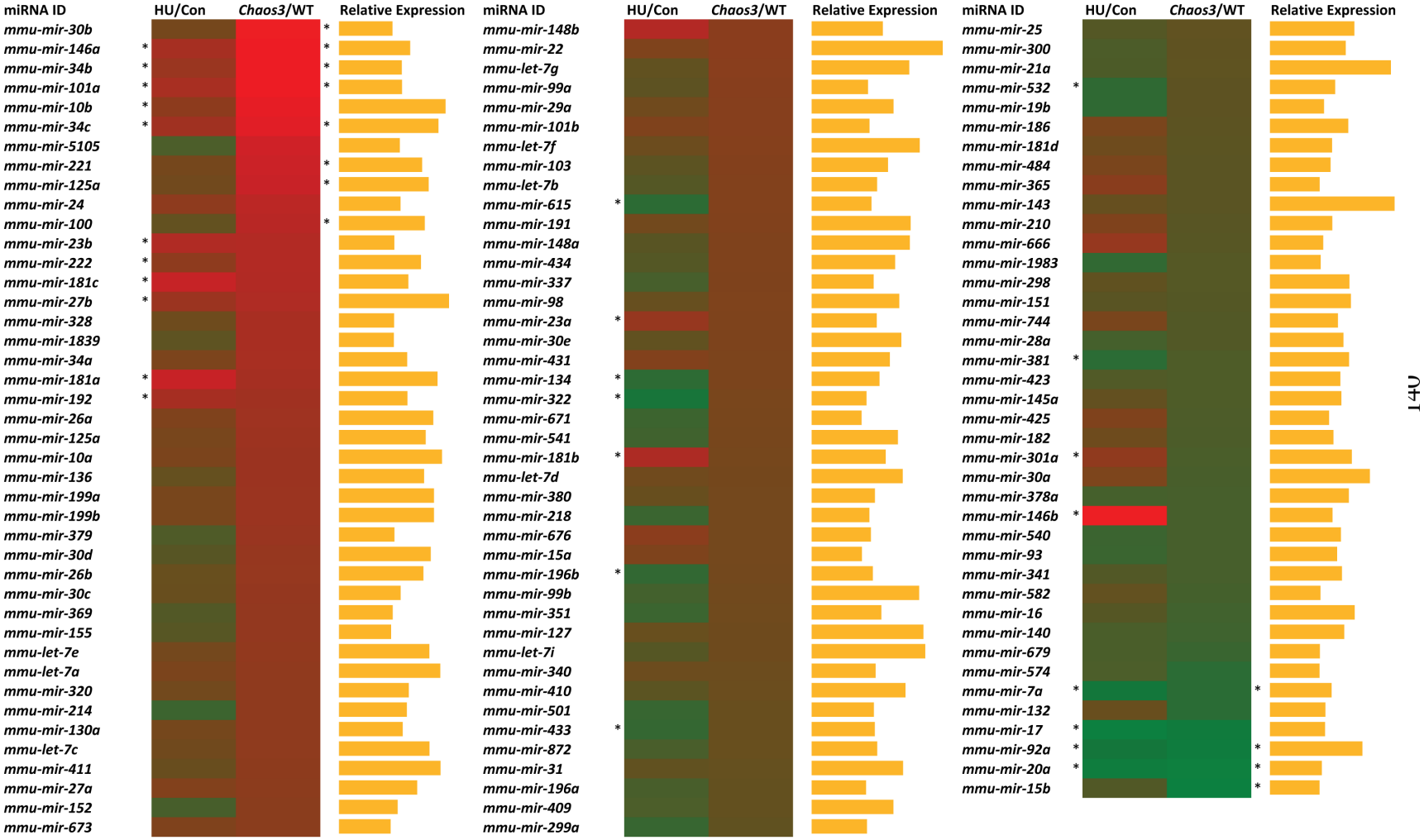

A
M1
S
M2
500
400
300
200
180
160
140
120
100
500
160
145
B
Log2 transformed # of reads
25
20
15
10
5
0
15
20
25
30
35
Length (nt)
C
% of Control
160
140
120
100
80
60
40
20
0
48h
72h
Mcm2-4
Cdk2
Ccne1
Ccne2
Ccna2
D
% of Control
120
100
80
60
40
20
0
E2f1
E2f3
Control
200μM HU

Figure 4.1 Overview of small RNA sequencing. (A) TBE-PAGE electroporation of constructed small RNA sequencing library. Custom markers (M1 & M2) were used to identify the library product with the proper size. Mature miRNAs with sequencing primers and barcodes added during library construction have a size between 145~160bp. DNA within this size range was purified from the PAGE gel and sequenced. **(B)** Size distribution of the small RNA sequenced for all the 24 libraries. The majority of the small RNAs have a size around 22nt. **(C)** qRT-PCR measurement of mRNA expression of cell cycle related genes after prolonged HU exposure. Results for WT primary MEFs exposed to 200μM HU for 48 or 72h were shown after normalized to their respective controls. Cyclins responsible for G1/S phase transition (cyclin E1 & E2) and S phase progression (cyclin A2, somatic cells) are accumulated after 48h HU treatment. 72h HU exposure repressed these cyclins' expression, indicating a transient cell cycle arrest program was implemented. *Cdk2* is the kinase that is regulated by these cyclins to decide S phase entry / progression. Similar to cyclins, *Cdk2* level was not affected at all with 48h HU treatment, but reduced after longer exposure. Error bar = SEM. **(D)** Expression of the *E2f* transcription factor family members are not affected after 200μM HU treatment for 48h in WT primary MEFs. Relative mRNA expression levels as measured by qRT-PCR were shown. Error bar = SEM. **(E)** Heat map of the miRNA expression profile in response to endogenous and exogenous RS. The *Chaos3*/WT column shows differential miRNA expression between *Chaos3* and WT MEFs at passage 5. The HU/Con column demonstrated altered miRNA expression with and without 200μM HU treatment for 48h in WT primary MEFs. Asterisks mark the miRNAs whose expressions are changed at a statistically significant level. The orange bar on the far right represents the relative expression level (Log_2 transformed) of each miRNA in WT MEFs without any treatment. Only the miRNAs that have more than 100 reads / 1 million total reads within each library were shown.

E


miRNA ID
HU/Con
Chaos3/WT
Relative Expression
mmu-mir-30b
mmu-mir-146a
mmu-mir-34b
mmu-mir-101a
mmu-mir-10b
mmu-mir-34c
mmu-mir-5105
mmu-mir-221
mmu-mir-125a
mmu-mir-24
mmu-mir-100
mmu-mir-23b
mmu-mir-222
mmu-mir-181c
mmu-mir-27b
mmu-mir-328
mmu-mir-1839
mmu-mir-34a
mmu-mir-181a
mmu-mir-192
mmu-mir-26a
mmu-mir-125a
mmu-mir-10a
mmu-mir-136
mmu-mir-199a
mmu-mir-199b
mmu-mir-379
mmu-mir-30d
mmu-mir-26b
mmu-mir-30c
mmu-mir-369
mmu-mir-155
mmu-let-7e
mmu-let-7a
mmu-mir-320
mmu-mir-214
mmu-mir-130a
mmu-let-7c
mmu-mir-411
mmu-mir-27a
mmu-mir-152
mmu-mir-673
miRNA ID
HU/Con
Chaos3/WT
Relative Expression
mmu-mir-148b
mmu-mir-22
mmu-let-7g
mmu-mir-99a
mmu-mir-29a
mmu-mir-101b
mmu-let-7f
mmu-mir-103
mmu-let-7b
mmu-mir-615
mmu-mir-191
mmu-mir-148a
mmu-mir-434
mmu-mir-337
mmu-mir-98
mmu-mir-23a
mmu-mir-30e
mmu-mir-431
mmu-mir-134
mmu-mir-322
mmu-mir-671
mmu-mir-541
mmu-mir-181b
mmu-let-7d
mmu-mir-380
mmu-mir-218
mmu-mir-676
mmu-mir-15a
mmu-mir-196b
mmu-mir-99b
mmu-mir-351
mmu-mir-127
mmu-let-7i
mmu-mir-340
mmu-mir-410
mmu-mir-501
mmu-mir-433
mmu-mir-872
mmu-mir-31
mmu-mir-196a
mmu-mir-409
mmu-mir-299a
miRNA ID
HU/Con
Chaos3/WT
Relative Expression
mmu-mir-25
mmu-mir-300
mmu-mir-21a
mmu-mir-532
mmu-mir-19b
mmu-mir-186
mmu-mir-181d
mmu-mir-484
mmu-mir-365
mmu-mir-143
mmu-mir-210
mmu-mir-666
mmu-mir-1983
mmu-mir-298
mmu-mir-151
mmu-mir-744
mmu-mir-28a
mmu-mir-381
mmu-mir-423
mmu-mir-145a
mmu-mir-425
mmu-mir-182
mmu-mir-301a
mmu-mir-30a
mmu-mir-378a
mmu-mir-146b
mmu-mir-540
mmu-mir-93
mmu-mir-341
mmu-mir-582
mmu-mir-16
mmu-mir-140
mmu-mir-679
mmu-mir-574
mmu-mir-7a
mmu-mir-132
mmu-mir-17
mmu-mir-92a
mmu-mir-20a
mmu-mir-15b

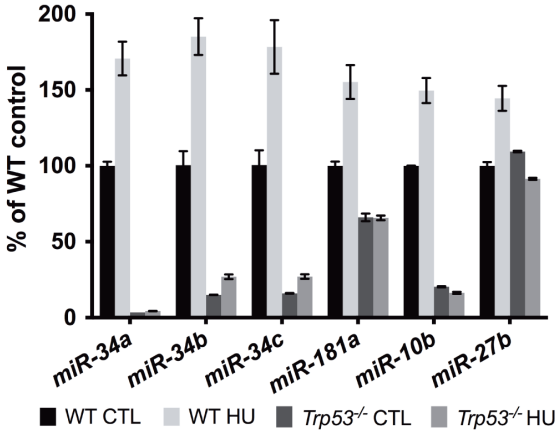
***Trp53*-dependent microRNAs represses MCM2-7 expression in response to RS**

We further studied the miRNAs as mentioned before to determine if they are responsible for repressing MCM2-7 expression in response to RS. We confirmed the sequencing results using miRNA qRT-PCR. HU treatment in WT primary MEFs induces the expression of *miR-34a, b & c, miR-10b, miR-27b* and *miR-181a*. Interestingly, *Trp53* deletion abolished the HU induced miRNA upregulation in the primary MEFs (Figure 4.2A), confirming these candidate miRNAs are *Trp53*-dependent (277).

miRNAs usually interact with miRNA recognition elements (MREs) in the 3'UTR of the target mRNAs to repress their expression. Full-length 3'UTRs of each *Mcm2-7* gene were cloned into a dual-luciferase reporter plasmid. Overexpression of all the aforementioned miRNAs except *mir-10b* repressed luciferase activity regulated by the corresponding MCM 3'UTR (Figure 4.2C & D), confirming *in silico* predictions (Figure 4.2B). We also found potential targeting of the *Mcm7* 3'UTR by *miR-34s,* despite the lack of *in silico*-predicted binding sites (Figure 4.2C). These results indicate that the candidate miRNAs target MCMs transcripts to reduce their expression, consistent with a subset of observations in human cells (280). We also evaluated *Mcm2-7* mRNA and protein levels after miRNA over-expression. *Mcm2-7* mRNA levels were not reduced after *miR-10b, 181a* or *27b* over-expression (Figure 4.2E), while *miR-27b* and *181a* repressed MCM4 protein expression but not other MCMs studied (Figure 4.2F). Although *miR-34* mainly repressed MCM5 & 7 as indicated by the luciferase assay (Figure 4.2C), supporting the finding that it is a direct target of *miR-34a* in the context of the RISC (280), overexpression of the *miR-34s* individually diminished MCM2-7 mRNA and protein (Figure 4.2G & H). Interestingly, siRNA knockdown of MCM5 also caused MCM2-7 pan-

reduction (Figure 4.2I). In sum, we identified a group of *Trp53*-dependent miRNAs that can regulate MCM2-7 expression directly or indirectly in response to RS.

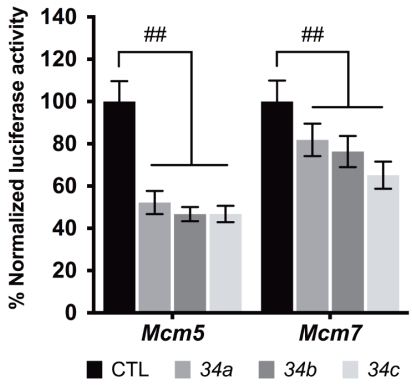
A



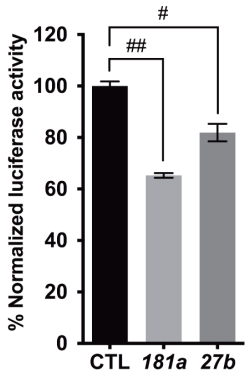
B

```
82: 5'-ACCUCUUCUAUUCCUGGAAUGUA-3' Mcm4 (wt)
       || : ::     | | ||||||
    3'-UGAGUGGCUGUCGCAACUUACAA-5' mmu-miR-181a

385: 5'-CCGUACAUCCUUUACUGUGAU-3' Mcm4 (wt)
        |: |  |   ::|||||||
     3'-CGUCUUGAAUCGGUGACACU-5' mmu-miR-27b

100:5'-AATAAAGTGTTTT--GATTGCCTC-'3 Mcm7 (wt)
          |  :  |      | |||||
    3'-CGUUAGUCGAUUGAUGUGACGGA-5' mmu-miR-34c

275:5'-GCUCCCGGGCCACUCACUGCCA-3' Mcm5 (wt)
       :|  ||:|   |  ||||||||
    3'-UGUUGGUCGAUUCUGUGACGGU-5' mmu-miR-34a

276:5'-CUCCCGGGCCACU--CACUGCCA-3' Mcm5 (wt)
       :    :|| | |  |||||||
    3'-UGUUAGUCGAUUAAUGUGACGGA-5' mmu-miR-34b

275:5'-GCUCCCGGGCCACU-CACUGCCA-3' Mcm5
       ||   |:|   ||| |||||||
    3'-CGUUAGUCGAUUGAUGUGACGGA-5' mmu-miR-34c
```

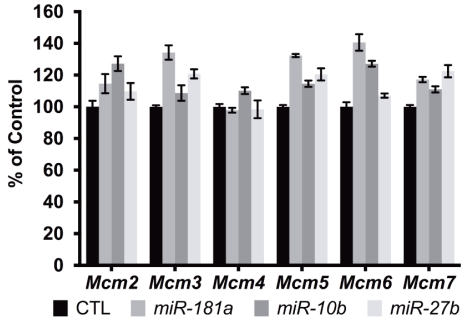
C



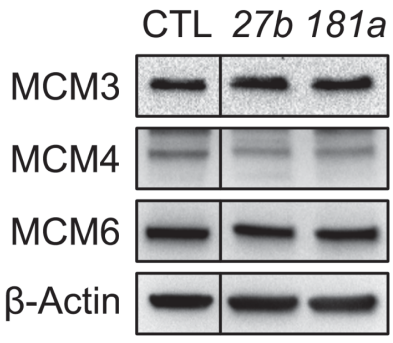
D



E



F



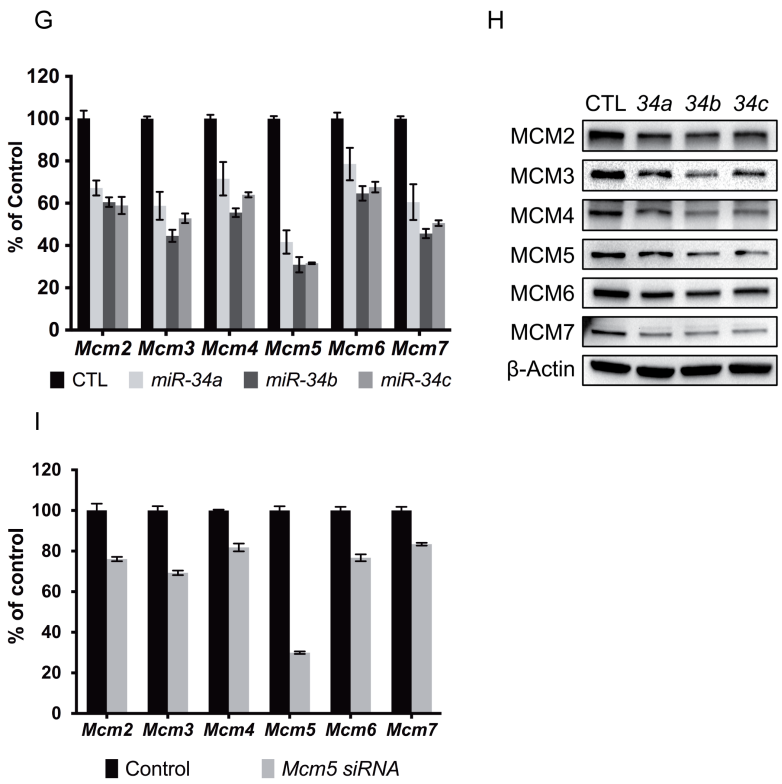


Figure 4.2 RS responsive miRNAs repress MCM2-7 expression. (A) RS induced miRNA upregulation depends on *Trp53*. miRNA levels were quantified by qRT-PCR and normalized to U6 small nuclear RNA (RNU6). The HU treatment was 200μM for 72h. Error bar = SEM. CTL = control. **(B)** miRNA targeting sites within the 3'UTR of MCM genes based on *in silico* prediction. Numbers on the left are starting positions of shown sequence within the 3'UTR. Seed sequences of miRNA are in red. **(C)** *miR-34abc* acts upon the 3′UTR of Mcm5 & 7 in HeLa cells co-transfected with a dual luciferase reporter bearing the full-length 3′UTR of the genes. *miR-34a, 34b,* and *34c* were transfected individually and assays were performed after 48h. **(D)** Same as **(C)**, but the impact of co-transfecting *miR-181a* or *miR-27b* with a reporter containing the full-length 3'UTR of *Mcm4* was assessed. **(E)** *miR-10b, 27b & 181a* overexpression through miRNA mimic transfection individually cannot reduce *Mcm2-7* mRNA expression. mRNA level of each gene was measured by qRT-PCR and normalized to β-actin levels. *Mcm2-7* mRNA levels were considered 100% in the control cells which were transfected with negative control miRNA mimics (based on cel-miR-67). Error bar= SEM. **(F)** Immunoblot showing MCM4 protein levels declined 48 hrs after *miR-27b, 181a* transfection. **(G)** *miR-34* family miRNA overexpression through miRNA mimic transfection individually reduced *Mcm2-7* mRNA expression. **(H)** Same as **(F)**, but *miR-34a, b* or *c* miRNAs were transfected. Note that the control lane in **(H)** is the same as in **(F). (I)** *Mcm5*-siRNA knockdown caused reduced expression of other MCM mRNAs. 50nM of siRNA specifically against *Mcm5* was transfected into primary WT MEFs for 48h. mRNA levels of *Mcm2-7* were measured by qRT-PCR and normalized to β-actin levels. *Mcm2-7* mRNA levels were considered 100% in the control cells which were mock transfected.

miR-34abc knockout in *Chaos3* mice rescues MCM2-7 expression and reduces genomic instability

Since MCM dosage impacts RS and *Trp53*-dependent RS-responsive miRNAs can regulate MCM expression, we tested whether modulating miRNA expression affects cellular responses to RS. Among the RS responsive miRNAs we studied, only the *miR-34* family miRNAs caused MCM2-7 pan-reduction upon ectopic expression (Figure 4.2G & H), a scenario similar to MCM2-7 repression after RS induction. Furthermore, overexpression of the *miR-34* miRNAs, but not other miRNAs, significantly inhibited DNA replication (Figure 4.3A). Numerous reports demonstrated that *miR-34* miRNAs impact cell cycle progression partly by targeting at DNA replication genes, including MCMs (277, 280, 282). To determine if these miRNAs impact cellular response to RS *in vivo*, we generated *miR-34abc* triple knockout (*34abc*-TKO) mice, in conjunction with the *Chaos3* mutation, and examined genomic instability and MCM levels in these mice and cells derived from them.

miR-34 deletion partially rescued MCM2-7 pan-reduction in HU-treated primary WT MEFs (Figure 4.3B), complementing the previous experiments in which overexpression of *miR-34s* decreased MCM levels. To determine if *miR-34s* deficiency could also rescue RS phenotypes *in vivo*, we measured MCM protein levels in various tissues. *Chaos3* mice had dramatically lower MCMs in multiple tissues compared to WT, similar to primary MEFs at later passages (Figure 4.3C; liver is shown). This observation indicates that the MCM2-7 pan-reduction in MEFs is not a culture artifact. MCM expression in *Chaos3* mice was partially rescued by *miR-34abc* deletion (Figure 4.3C). These results indicate that *miR-34* expression contributes to both endogenous and exogenous RS-induced MCM2-7 repression *in vivo* and *in vitro*.

A hallmark of the *Chaos3* mutation is highly elevated micronuclei (MN), an indicator of genomic instability (GIN), in reticulocytes and erythrocytes (135). *miR-34* deletion reduced MN levels in *Chaos3* mice by ~20% (Figure 4.3D), and this reduction in MN was sensitive to *miR-34* genetic dosage (Figure 4.3E). However, *Chaos3, miR-34abc* knockout quadruple female mutants still succumb to tumor formation with latency similar to *Chaos3* single mutants (Figure 4.3F). These data, in conjunction with the results showing that *miR-34* ablation increased MCM levels in *Chaos3* mouse tissues, indicates that at least part of the genomic instability in *Chaos3* mice is related to decreased replication origin licensing orchestrated by TRP53 induction of MCM-targeting miRNAs. However, MCM reduction in the *Chaos3* mutant has very minor impact on genomic instability and tumoregensis, supporting our hypothesis that other character of the *Chaos3* mutation, presumably helicase instability, is the major cause of all the defects found in this mutant.

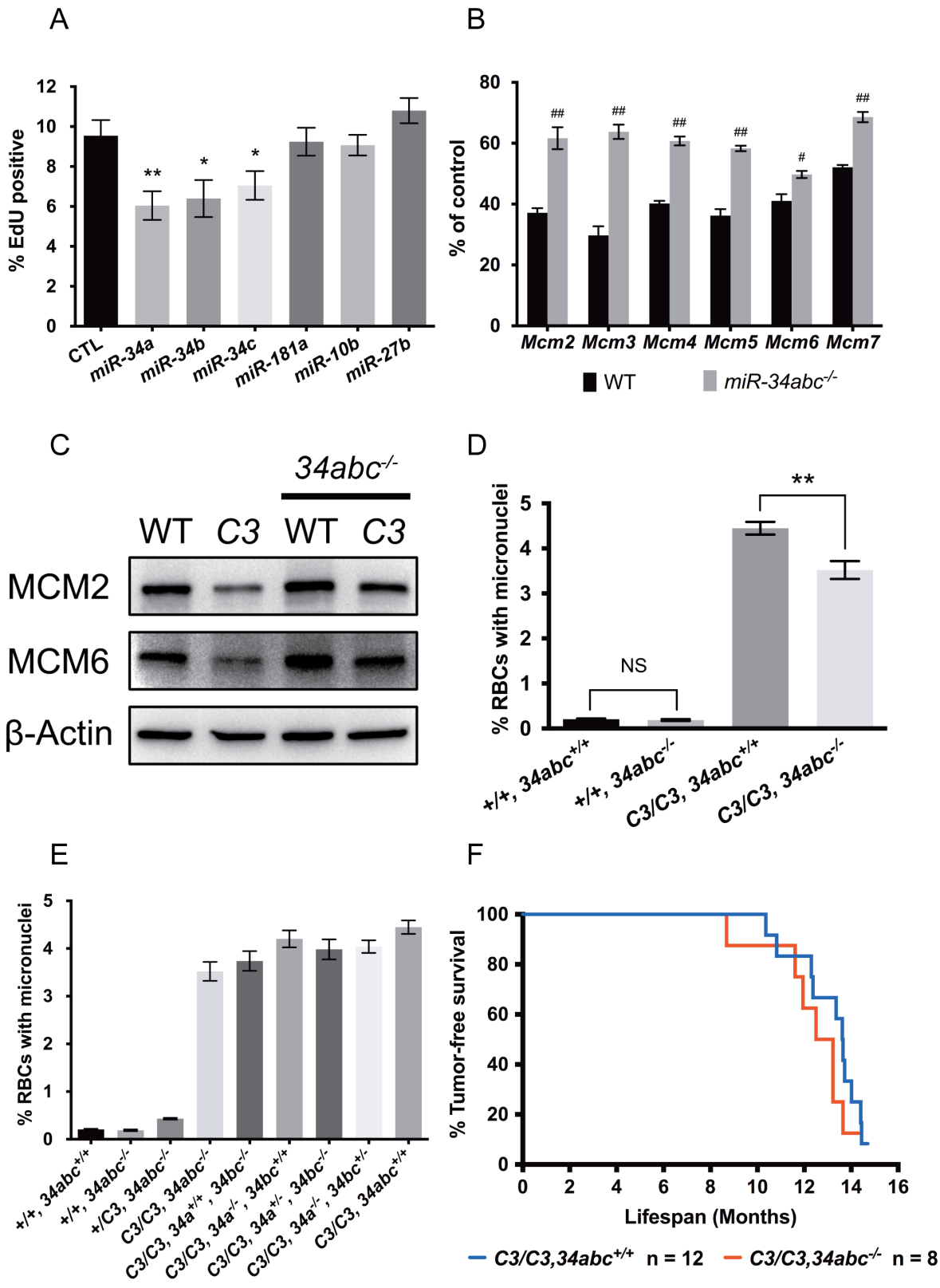

A
% EdU positive
CTL
miR-34a
miR-34b
miR-34c
miR-181a
miR-10b
miR-27b
B
% of control
Mcm2
Mcm3
Mcm4
Mcm5
Mcm6
Mcm7
WT
miR-34abc-/-
C
34abc-/-
WT
C3
WT
C3
MCM2
MCM6
β-Actin
D
% RBCs with micronuclei
NS
+/+, 34abc+/+
+/+, 34abc-/-
C3/C3, 34abc+/+
C3/C3, 34abc-/-
E
% RBCs with micronuclei
+/+, 34abc+/+
+/+, 34abc-/-
+/C3, 34abc-/-
C3/C3, 34abc-/-
C3/C3, 34a+/+, 34bc-/-
C3/C3, 34a-/-, 34bc+/+
C3/C3, 34a+/-, 34bc-/-
C3/C3, 34a-/-, 34bc+/-
C3/C3, 34abc+/+
F
% Tumor-free survival
Lifespan (Months)
C3/C3,34abc+/+ n = 12
C3/C3,34abc-/- n = 8

Figure 4.3 *miR-34abc* deficiency partially rescues RS induced MCM2-7 repression and genomic instability in the *Chaos3* mutant. (A) Ectopically expressed miRNAs through miRNA mimic transfection affects DNA replication. Primary WT MEFs transfected with *miR-34* miRNA mimics for 48h significantly inhibit DNA replication (two sided t-test, *, $p<0.05$, **, $p<0.005$), but not *miR-10b, 27b* or *181a* mimics. Control cells were transfected with the negative control miRNA mimics. Error bar = SEM. **(B)** *miR-34abc* deletion partially rescues *Mcm2-7* mRNA levels after RS induction. WT and *miR-34abc*$^{-/-}$ primary MEFs were treated with 200μM HU for 72h and *Mcm2-7* mRNA levels were measured by qRT-PCR. β-actin levels were also measured to normalize the *Mcm2-7* levels (statistically significant, two-sided t-test. #, $p<0.01$, ##, $p<0.001$). Error bar = SEM. **(C)** *miR-34abc* deficiency partially rescues MCM protein expression in liver from *Chaos3* (*C3/C3*), *miR-34abc*$^{-/-}$ compound mutants. Total protein extracted from age-matched livers of WT and mutants was immunoblotted to measure MCM expression. Protein expression of β-actin was used as loading control. Other tissues including lung, spleen, kidney and small intestine were also studied which showed similar results (data not shown). **(D)** *miR-34abc* deletion in the *Chaos3* mutant significantly rescued micronuclei (MN) formation (two-sided t-test, **, $p<0.005$). Numbers of individuals analyzed for each genotype (from left to right) were 12, 5, 31 and 17 respectively. **(E)** MN levels of different *Chaos3* / *miR-34abc* compound mutants. With increased deletion of *miR-34abc* genetic dosage, MN reduced. **(F)** Kaplan-Meier tumor-free survival plot of the female animals with the indicated genotypes. Additional *miR-34abc* deletion did not affect tumor-free survival of the *Chaos3* mutant (Log-rank / Mantel-Cox test, $p = 0.2600$; Gehan-Breslow-Wilcoxon test $p = 0.2356$). Median tumor-free survival latency *C3/C3, 34abc*$^{+/+}$ = 13.63, *C3/C3, 34abc*$^{-/-}$ = 12.86 months.

4.5. Conclusions and Discussion

In this chapter, I discussed the discovery of a group of miRNAs that represses MCM2-7 expression in response to RS. These miRNAs are regulated by central tumor suppressor gene *Trp53* and thus are connected to the intrinsic cellular response to RS through DDR activation. Each miRNA is modestly induced in primary cells after prolonged exposure to RS. However, they may function cooperatively to significantly reduce MCM2-7 expression. We also confirmed the function of *miR-34* family miRNAs in reducing MCM expression in the context of *Chaos3* mutation induced endogenous RS both *ex vivo* and *in vivo*. Partial rescue of MCM expression in the *Chaos3* mutant animals after *miR-34* miRNA deletion significantly reduced genomic instability at the whole animal level, yet failed to rescue the cancer susceptibility phenotype, confirming that reduced MCM2-7 expression only marginally contributed to genomic defects associated with this mutation.

Post-transcriptional regulation of MCM2-7 expression was first implied when studying the *Chaos3* mutant. Although MCM2-7 repression was considered as the major cause of RS and genomic instability in this mutant based on previous studies (135, 249, 251, 256), it was also suggested that these genes are targeted specifically for downregulation (250). The suppression originated from reduced mRNA expression, which is also reflected at the protein level. mRNA reduction can happen at both transcriptional and post-transcriptional levels. In a detailed report on this subject, multiple independent assays were designed to evaluate *Mcm2-7* expression at the transcriptional level in the *Chaos3* mutant cells (250), which revealed that it was not affected comparing to the WT counterpart. Besides, ablation of microRNA biogenesis through *Drosha* or *Dicer* knockdown resulted in increased *Mcm2-7* mRNA expression in the *Chaos3* mutant. Together, these results pinpointed *Mcm2-7* repression at the post-transcriptional level, at least for

the case of *Chaos3* mutant. Thus we performed small RNA sequencing to identify the culprit small RNAs targeting at MCM2-7 expression, with a focus on the miRNAs as none of the other small RNA classes (siRNA & piRNA) are expressed in the experimental system we used.

Small RNA sequencing results revealed many miRNAs whose expressions are altered significantly at the statistical level during RS response. This result confirmed our previous report that *miR-34* family miRNAs are upregulated in *Chaos3* cells suffering from endogenous RS (250). Moreover, we also identified a few other miRNA families, with their members being upregulated in response to RS. As *Trp53* is connecting the MCM2-7 repression to RS activated DDR, we hypothesized that *Trp53* should also control the expressions of these miRNAs. Indeed, disruption of TRP53 function either abolishes the overall expression (*miR-34s, miR-10b, miR-181a*), or the RS induction of the miRNAs (*miR-27b*). All of these miRNAs were shown to be *Trp53* responsive during acute DDR or ectopic expression of oncogenes (277), suggesting these miRNAs may response to generic genotoxic stress upon *Trp53* activation, and is not limited to RS.

miRNAs mainly target 3'UTR of mRNAs to repress expression, either by inducing mRNA degradation or translational suppression, or a combination of both mechanisms. The 3'UTR analysis using the luciferase assay demonstrated the relative level of suppression on protein expression (luciferase reporter), which can be the consequence of removed mRNA transcript of the reporter gene or repressed production of luciferase protein. In combination with *Mcm2-7* mRNA level analysis upon ectopic miRNA expression, we found that *miR-34* family miRNAs have a strong suppressive power over all of the *Mcm2-7* mRNA and protein expression, despite that only the 3'UTR of *Mcm5 & 7* are targeted by these miRNAs. These results confirmed the previous analysis in human cells, where a physical interaction was confirmed

between *miR-34a* and *Mcm5* 3'UTR (280). Other RS induced miRNAs identified here did not reduce MCM2-7 expression at the mRNA level, though the suppression was present at the protein level, suggesting translational inhibition.

miR-34 family miRNAs mainly target at the *Mcm5* of the *Mcm2-7* genes as suggested by the luciferase assay. *Mcm5* mRNA also showed the highest level of reduction comparing to other *Mcm2-7* genes upon ectopic expression of individual miRNA from the *miR-34* family. This result suggested that severe inhibition of an individual *Mcm2-7* expression, such as *Mcm5* could lead to overall suppression of all the *Mcm2-7* genes, as demonstrated by *Mcm5* siRNA knockdown. This effect is probably not limited to *Mcm5* within the MCM gene family, as severe reduction of MCM2 expression due to genetic mutation also caused lowered expression of other MCMs (128). The MCM2-7 pan-reduction effect is thus probably triggered based on the severity of reduction of an individual Mcm gene. Genetic removal of ~50% of Mcm gene expression failed to cause MCM2-7 pan-reduction in various MCM (*Mcm3, 4, 6* & 7) gene-trap mutants (249). However, *Mcm2* gene-trap mutant with the same genetic disruption of one copy of *Mcm2* locus showed reduced expression in other MCM genes (250), given that eukaryotic cells are likely to express lowest copies of MCM2 comparing to other MCMs (119). Cells can also evaluate MCM expression level directly and make an adjustment to maintain the relative stoichiometry between all the MCMs. For instance, ectopic expression of a single MCM protein reduces the protein expressed from the endogenous locus of that specific MCM gene (283). However, the mechanism that causes MCM2-7 pan-reduction in response to RS is most likely independent of this direct MCM level evaluation in the cell. For instance, all of the *Mcm2-7* mRNA expressions are reduced by ~40-50% comparing to the control in response to RS in

primary WT and *Mcm2* gene-trap cells (Figure 4.4B), including *Mcm2* that is already reduced by ~60% from the normal level due to the genetic disruption in the *Mcm2* mutant (Figure 4.4A).

Regulated miRNA expression can have great impact on cellular functions, since each individual miRNA have many potential targets. However, most of the miRNA only moderately repress the expression of its target genes, since the effect of a single miRNA species is often diluted due to miRNA association with other potential targets. miRNA regulated gene repression can also be reversed when the target mRNAs are released from post-transcriptional repression without being permanently removed from the cells through mRNA degradation. These characters of miRNAs make them good candidates for regulating MCM2-7 expression in the presence of chronic low level of RS. miRNAs can be rapidly induced upon DDR activation through *Trp53* function. All of these RS-inducible miRNAs can function individually or cooperatively to repress MCM2-7 expression. Since the repression is moderate, MCM2-7 expression will not be ablated dramatically; cells can still proliferate normally with this level of MCM2-7 reduction. It thus provides the cells a window of opportunity to evaluate the conditions without committing to terminal cellular decisions such as permanent cell cycle arrest.

Among the identified RS-inducible, *Trp53*-depenedent miRNAs, we further focused on the *miR-34* family miRNAs for the following reasons: 1) these miRNAs were shown to be upregulated in the primary cells having endogenous or exogenously induced RS. 2) *miR-34* family miRNAs are well characterized. They are direct transcriptional target of *Trp53* upon its activation (277), and MCM genes were identified as *bona fide* targets of these miRNAs (280). 3) Ectopic expression of *miR-34* miRNAs can negatively regulate cell cycle progression and induce permanent cell cycle arrest (277), partially due to the fact that *miR-34s* repress cell cycle related genes, including replication genes' expressions. We also found that *miR-34* family miRNAs are

the only ones that can significantly inhibit DNA replication upon ectopic expression (Figure 4.3A). 4) *miR-34* family miRNAs are the only ones we studied that are capable of inducing MCM2-7 pan-reduction at both the mRNA and protein level, the same phenotype we observed in the primary cells upon RS induction, and finally 5) *miR-34* knockout mouse models were generated and readily available for our research (273). *miR-34s* deletion partially rescued MCM2-7 pan-reduction induced by RS originated from endogenous (*Chaos3* mutant background) or exogenous (WT background, HU treatment) sources. However, the level of rescue is limited, suggesting other mechanisms must be involved to reduce MCM2-7 expression: besides *miR-34s*, we also identified other *Trp53*-dependent miRNAs that can target at MCM2-7 genes. Furthermore, we cannot fully rule out the possibility that MCM2-7 genes are downregulated at the transcriptional level in response to RS, or the presence of other gene expression regulatory mechanisms. This mediocre yet significant rescue of MCM2-7 expression in the *Chaos3* mutant background also resulted in significant rescue of genomic instability. This confirms the previous proposals that reduced MCM2-7 expression contribute to the genomic instability defects in this mutant (135, 249, 251). However, since the rescue is very moderate and the level of genomic instability is still vastly elevated, it is clear that MCM2-7 reduction only marginally contribute to genomic defects reported in the *Chaos3* mutant.

In conclusion, *Trp53* regulated miRNAs assisted MCM2-7 repression during cellular response to chronic RS. *miR-34* family miRNAs appear to be the major culprits, yet the involvement of other miRNAs could not be ruled out. RS induced miRNAs moderately repress MCM2-7 expression at the post-transcriptional level, allowing cells to counteract to chronic RS without committing to terminal cellular decision resulted from genotoxic stresses.

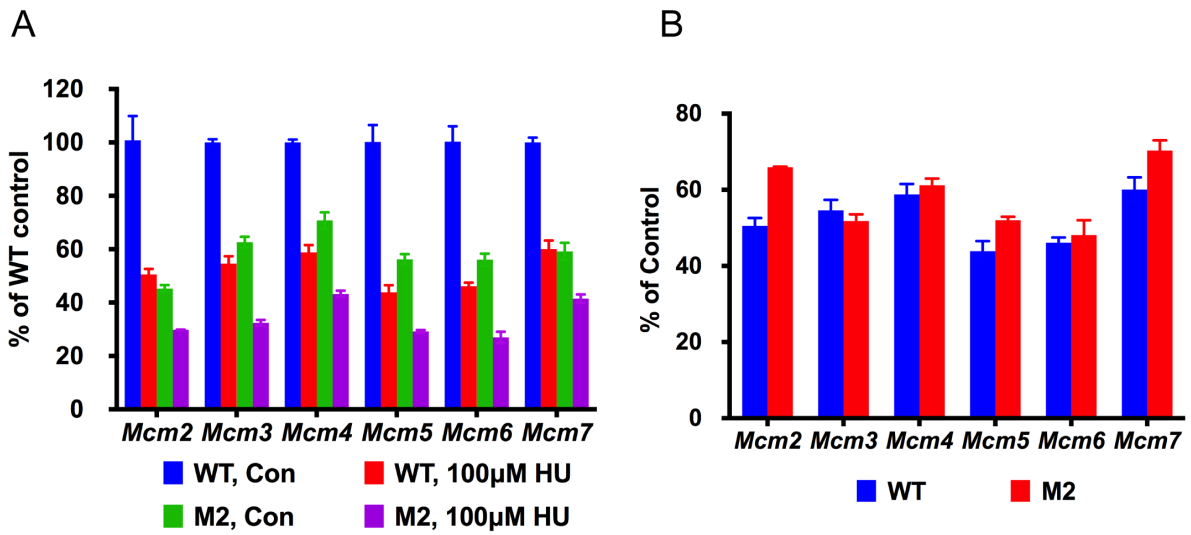


Figure 4.4 Additional MCM2-7 pan-reduction in response to RS is regardless of MCM expression level. (A) Relative mRNA expression level of *Mcm2-7* measured by qRT-PCR in primary WT and *Mcm2* gene-trap (M2) MEFs in response to RS. Cells are treated with 100μM HU for 72h. Control (Con) cell received no HU treatment. Relative mRNA levels are normalized to WT control samples. Notice that M2 cells already have reduced *Mcm2* mRNA expression due to gene-trap insertion, which also have lower expression of all the other *Mcm3-7* mRNAs (WT, Con Vs. M2, Con). Error bar = SEM. **(B)** The same data from (A), but the relative mRNA expression level was presented after comparing to their respective control, i.e. the height of blue bar in (B) represents the ratio between red and blue bar shown in (A). Both WT and M2 cells have reduced expression in each of the *Mcm2-7* mRNAs at the similar levels in response to RS. Error bar = SEM.

4.6. Acknowledgement

The authors thank Dr. A. Grimson for guidance with the small RNA sequencing library construction and sequencing reads alignment.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

5.1. Chronic DNA replication stress reduces replicative lifespan of cells by TRP53-dependent, microRNA assisted MCM2-7 downregulation

DNA replication stress (RS) are conditions that can inhibit DNA replication process by stalling replisome progression during DNA replication. Stalled replisomes, when not recovered in time, can collapse and result in DNA damage and incomplete DNA replication. DNA damage response (DDR) is responsible for the detection and recovery of stalled replisomes, and prevents DNA lesions associated with DNA replication processes. Many types of the RS originate from endogenous sources, thus proliferating cells inevitably experience RS during DNA replication. Endogenous RS often happens at low levels, which results in low incidence of replication associated DNA lesions as well. DDR is only moderately activated to resolve endogenous RS induced DNA damage, which is not sufficient to promote terminal cell cycle arrest of the proliferating cells. Furthermore, some of these damage escape the detection from DNA damage response (DDR) mechanisms and are not properly resolved during each cell cycle, and would accumulate in the proliferating cell lineages throughout their replicative lifespan. A few of these mutations may function independently or collaboratively to initiated malignant transformation, causing chaotic genome state such as genomic instability, which eventually promotes oncogenesis. As DDR activity is reset to normal before cells start another cell cycle, the daughter cells would have no recollection of the RS associated DNA damage from the previous cell cycle.

MCM2-7 are essential eukaryotic DNA replication genes, which participate in DNA replication licensing and function as the core of replicative helicase. MCM2-7 are highly expressed in proliferating cells. One of the purposes of excessive MCM2-7 production is to

thoroughly license the genome at any possible replication origins, including dormant origins that serve as backups. In the event of acute RS that stalls replisomes, nearby dormant origins can be activated to rescue the incomplete DNA replication and ensure the faithful transmission of genetic materials. Based on the results presented here, I propose another major function of MCM2-7 as the carrier of molecular memory of the accumulative RS experienced by proliferating cells over the replicative lifespan. MCM2-7 expression levels are gradually reduced as replicating cells normally response to chronic low levels of RS. Moderate MCM2-7 loss can sustain normal DNA replication and cell proliferation, however sensitize cells to additional RS by accelerating terminal cell cycle arrest, such as through senescence induction. Primary cells with severe MCM loss may fail licensing checkpoint, thus do not enter S phase with insufficiently licensed genome. Reduced MCM expression level also compromises dormant origin licensing, thus cells may not be able to resolve RS induced replisome stalling in a timely manner and arrest cell cycle progression terminally (Figure 5.1). Desensitizing to MCM expression loss and/or insufficient licensing may be one of the key events during cellular transformation. MCM2-7 repression is regulated by DDR in the event of RS. DDR reacts to RS collectively, either by directly detecting stalled replication structure (ATR related), or DNA lesions resulted from collapsed replisome (ATM related) and converge the information on central tumor suppressor *Trp53* to regulate MCM2-7 expression. *Trp53*-dependent microRNAs assist during this process, which moderately regulate MCM2-7 expression level. Since MCM2-7 are highly expressed in normal cycling cells, it can become a rich reserve so the proliferating cells can have extended replicative lifespan even after experiencing substantial amount of RS. This may have import implication to stem cells, as they need to maintain their self-renew ability for the better part of the lifespan of the whole living organism.

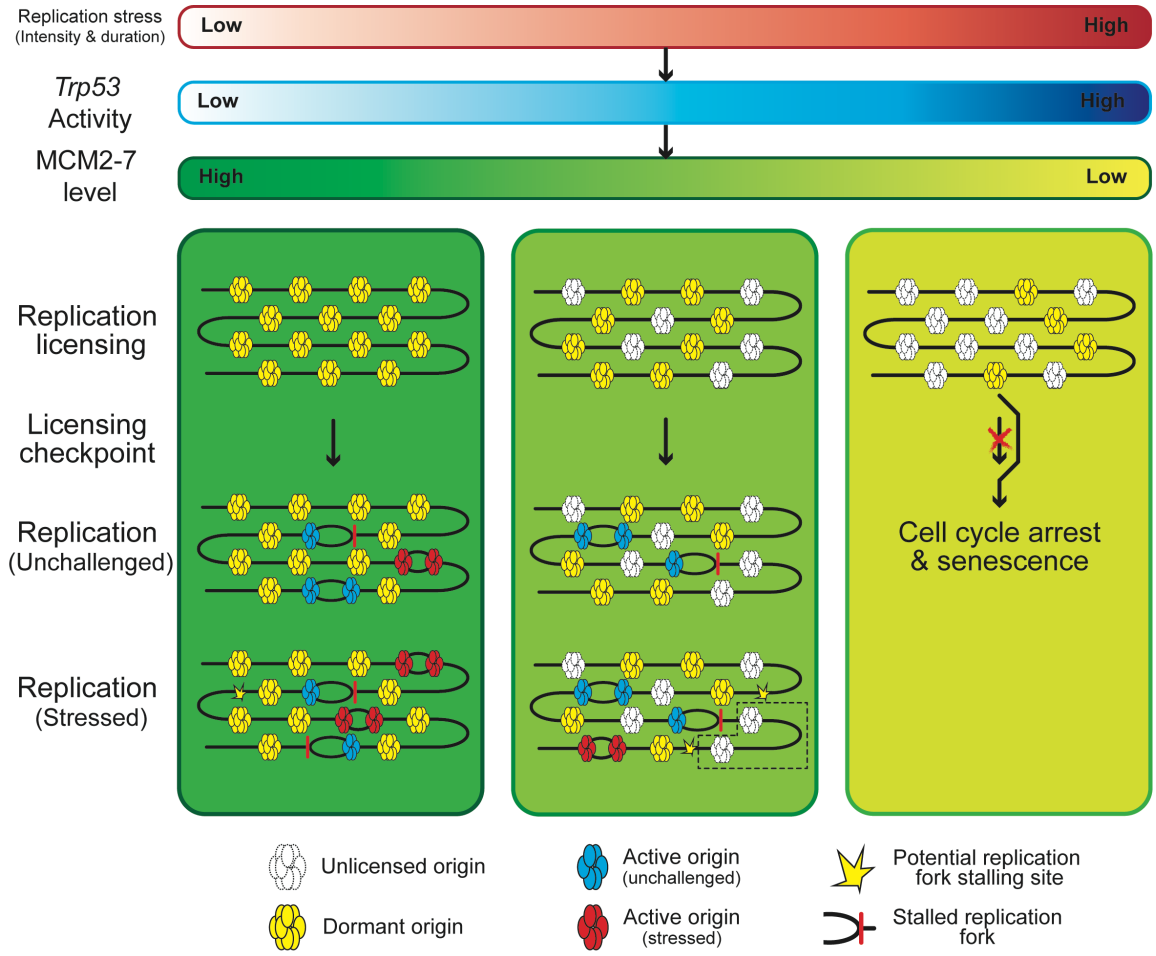


Figure 5.1 Model of replication stress (RS) induced, DDR-mediated MCM2-7 downregulation dictates replicative lifespan of the proliferating cells. Normally, proliferating cells experience low level of spontaneous RS, which maintains high MCM2-7 expression and sufficiently license the genome at all the potential replication origins, including dormant origins (left panel). Once replication is challenged in cells with normal MCM2-7 expression, dormant origins can be activated to rescue nearby stalled replication fork and ensures completeness of genome replication. MCM2-7 expression can be downregulated in proliferating cells experiencing RS (middle panel). The level of reduction reflects the collective effect of RS experienced by such cells. Reduced MCM2-7 expression compromise sufficient genome licensing, especially dormant origin licensing. Once the replication is challenged in these cells, they have limited number of MCM2-7 licensed dormant origins to serve as backup. RS stalled replication fork may not be rescued by the converging replication fork assembled from dormant origins, producing under-replicated genome region and loss of genomic information. Cells with servere MCM2-7 expression loss after experiencing high level of RS will arrest cell cycle progression and cease to proliferate through senescence induction (right panel). DDR, especially tumor suppressor *Trp53* activation response to RS in the cells, which further regulates MCM2-7 expression level in primary cells.

5.2. Modeling Chaos3 mutant using *Saccharomyces cerevisiae*

Saccharomyces cerevisiae, or budding yeast is such an excellent system to study DNA replication processes genetically and biochemically. MCM genes were formerly discovered in genetic screenings performed in this model organism. Since MCM are highly conserved throughout eukaryotes, the *Chaos3* mouse mutant was modeled by generating the genetically equivalent mutation in budding yeast MCM4 protein. Yeast *Chaos3* mutant strain expressed normal levels of MCM2-7 proteins, yet recapitulated many other phenotypes found in the *Chaos3* mouse mutant, such as chromosome aberrations, aneuploidy and genomic instability. It also displayed improved growth phenotype, which is resulted from genetic mutation acquired during genomic instability events (272). Many of the DNA replication reactions have been recapitulated biochemically in vitro using purified budding yeast replication proteins (47, 60). If I were given the chance to study the *Chaos3* mutant using buddying yeast as initially planned, here are some of the projects I would like to work on.

Replication progression can be easily monitored in budding yeast cells since 1) the population can be easily synchronized to enter the S phase simultaneously and 2) the genomic locations of DNA replication origins are determined, thus replication fork progression can be tracked over S phase. Since *Chaos3* mutation disrupts helicase complex stability, it is conceivable to map the sites where the mutant helicase would have a higher tendency to stall. Thus by studying the genes near these susceptible replication stalling points, it will be clearer to determine the primary genomic alteration events that may contribute to aneuploidy and improved growth in the *Chaos3* mutant yeast strain.

The structure of entire MCM2-7 complex was solved (92). Based on the structural information, the ultimate argument in the *Chaos3* mutant: whether reduced MCM expression or

MCM2-7 complex instability is the original source of DNA replication defects can be settled. Consistent with the previous predictions based on the archaeal MCM structure, *Chaos3* mutation in budding yeast also resides on the MCM interactive surface. A further examination of this structure suggested that the benzene side chain missing from the phenylalanine (F) to isoleucine (I) mutation protrudes from the MCM4 protein into the MCM4:6 interactive surface. Genetically engineered mutations in amino acids in the MCM6 protein that may potentially interact with the F391 residue in MCM4 may alter the interaction between MCM proteins in a similar way as the Chaos3 mutant. One of such potential targets for mutation is tyrosine (Y) 450 in MCM6. Its phenol side chain appears to be in close proximity to the benzene side change of F391. Mutants carrying mutation such as Y450I in MCM6 might be able to phenocopy the *Chaos3* mutant.

Another major usage of the budding yeast is to determine the impact of reduced MCM expression on replication licensing and replication dynamics genome wide. Positions of all of the replication origins are well characterized in budding yeast genome. By performing a ChIP-seq experiment using antibodies against MCM proteins on chromosomes isolated from G1 phase budding yeast cells, it will be easy to determine how reduced MCM dosage would affect replication licensing: whether it would fail to license some of the origins, such as dormant origins, or the MCM footprint near licensed origins will be reduced. One can also determine the genome wide usage of origins with reduced licensing.

5.3. Studying the impact of MCM expression stem cell maintenance and cancer

Based on the research presented here on somatic cycling cells, MCM2-7 repression in response to RS is an authentic cellular response. At organismal level, stem cells are the group of cells devoted to self-renew and generate new cells for differentiation. Stemness and the self-renew ability must be maintained in stem cells during the lifespan of the whole living organism,

otherwise leading to stem cell depletion, developmental defects and premature aging. RS induced MCM2-7 downregulation mechanism appears to be present in stem cells as in somatic cells, as stem cells experiencing endogenous RS have lowered MCM2-7 expression (259). MCM2-7 expression level dictates the replicative lifespan of the stem cells. Reduced MCM expression in stem cells could limit their *in vivo* expansion (255). Unlike somatic cells which exit cell cycle by assuming a terminal cell cycle arrest state such as senescence, quiescent or apoptosis, stem cells can choose to lose self-renew ability through differentiation. Thus, reduced replicative lifespan due to MCM2-7 downregulation or insufficient licensing may correlates with the loss of stemness, which could be accelerated by RS. The simplest way to test this hypothesis is by exposing stem cells to chronic low level of RS and then determines whether the loss of MCM expression correlates with reduced expression of stemness maintenance markers. Furthermore, stem cells that express reduced dosage of MCM protein should be more inclined to differentiate during response to RS.

Elevated MCM expressions were often found in precancerous or cancer cells, which had been used as histological markers to diagnose the disease (132). Cancer cells that upregulate MCM expression might be positively selected, or increased MCM expression is a consequence of disrupted tumor suppressors function that usually controls normal MCM expression. Either way, since increased MCM expression can improve cellular response to RS by thoroughly licensing the genome, elevated MCM expression is favorable in cancerous cells as they often experience elevated level of RS due to abnormally activated growth signaling and increased conflict between replisome and transcription machinery. Procedures that can further elevate the RS or reduce cellular response capacity to RS in cancer cells may help to inhibit their proliferation and thus tumor growth. This can be achieved by specifically targeting at MCM

genes or other licensing factors to reduce their expression thus inhibit replication licensing. In combination with a recently discovered MCM2-7 helicase inhibitor (284), MCM function can be completely shut down in highly proliferative cancer cells, which serves as a potential therapeutic procedure of cancer treatment.

REFERENCE

1. Blow JJ, Dutta A. Preventing re-replication of chromosomal DNA. Nature reviews Molecular cell biology. 2005;6(6):476-86.
2. Arias EE, Walter JC. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. Genes & development. 2007;21(5):497-518.
3. Diffley JF, Cocker JH, Dowell SJ, Rowley A. Two steps in the assembly of complexes at yeast replication origins in vivo. Cell. 1994;78(2):303-16.
4. Edenberg HJ, Huberman JA. Eukaryotic chromosome replication. Annual review of genetics. 1975;9:245-84.
5. Bell SP, Dutta A. DNA replication in eukaryotic cells. Annual review of biochemistry. 2002;71:333-74.
6. Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M. Eukaryotic chromosome DNA replication: where, when, and how? Annual review of biochemistry. 2010;79:89-130.
7. Sclafani RA, Holzen TM. Cell cycle regulation of DNA replication. Annual review of genetics. 2007;41:237-80.
8. Mott ML, Berger JM. DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol. 2007;5(5):343-54.
9. Struhl K, Stinchcomb DT, Scherer S, Davis RW. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proceedings of the National Academy of Sciences of the United States of America. 1979;76(3):1035-9.
10. Brewer BJ, Fangman WL. The localization of replication origins on ARS plasmids in S. cerevisiae. Cell. 1987;51(3):463-71.
11. Palzkill TG, Oliver SG, Newlon CS. DNA sequence analysis of ARS elements from chromosome III of Saccharomyces cerevisiae: identification of a new conserved sequence. Nucleic acids research. 1986;14(15):6247-64.
12. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, et al. Replication dynamics of the yeast genome. Science. 2001;294(5540):115-21.
13. Wyrick JJ, Aparicio JG, Chen T, Barnett JD, Jennings EG, Young RA, et al. Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science. 2001;294(5550):2357-60.
14. Dai J, Chuang RY, Kelly TJ. DNA replication origins in the Schizosaccharomyces pombe genome. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(2):337-42.
15. Gilbert DM. In search of the holy replicator. Nature reviews Molecular cell biology. 2004;5(10):848-55.
16. Nieduszynski CA, Knox Y, Donaldson AD. Genome-wide identification of replication origins in yeast by comparative genomics. Genes & development. 2006;20(14):1874-9.
17. Xu W, Aparicio JG, Aparicio OM, Tavare S. Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in S. cerevisiae. BMC Genomics. 2006;7:276.
18. Chuang RY, Kelly TJ. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(6):2656-61.
19. Lee JK, Moon KY, Jiang Y, Hurwitz J. The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by

means of the AT-hook domains of the spOrc4 protein. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(24):13589-94.
20. Gilbert DM. Making sense of eukaryotic DNA replication origins. Science. 2001;294(5540):96-100.
21. Cadoret JC, Meisch F, Hassan-Zadeh V, Luyten I, Guillet C, Duret L, et al. Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(41):15837-42.
22. MacAlpine DM, Rodriguez HK, Bell SP. Coordination of replication and transcription along a Drosophila chromosome. Genes & development. 2004;18(24):3094-105.
23. Pappas DL, Jr., Frisch R, Weinreich M. The NAD(+)-dependent Sir2p histone deacetylase is a negative regulator of chromosomal DNA replication. Genes & development. 2004;18(7):769-81.
24. Pasero P, Bensimon A, Schwob E. Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. Genes & development. 2002;16(19):2479-84.
25. Mantiero D, Mackenzie A, Donaldson A, Zegerman P. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. EMBO J. 2011;30(23):4805-14.
26. Guillou E, Ibarra A, Coulon V, Casado-Vela J, Rico D, Casal I, et al. Cohesin organizes chromatin loops at DNA replication factories. Genes & development. 2010;24(24):2812-22.
27. Mendez J. Temporal regulation of DNA replication in mammalian cells. Crit Rev Biochem Mol Biol. 2009;44(5):343-51.
28. Wu R, Terry AV, Singh PB, Gilbert DM. Differential subnuclear localization and replication timing of histone H3 lysine 9 methylation states. Mol Biol Cell. 2005;16(6):2872-81.
29. Jackson DA, Pombo A. 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. The Journal of cell biology. 1998;140(6):1285-95.
30. Farkash-Amar S, Lipson D, Polten A, Goren A, Helmstetter C, Yakhini Z, et al. Global organization of replication time zones of the mouse genome. Genome Res. 2008;18(10):1562-70.
31. Karnani N, Taylor C, Malhotra A, Dutta A. Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas. Genome Res. 2007;17(6):865-76.
32. Schubeler D, Scalzo D, Kooperberg C, van Steensel B, Delrow J, Groudine M. Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. Nature genetics. 2002;32(3):438-42.
33. Goren A, Tabib A, Hecht M, Cedar H. DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin. Genes & development. 2008;22(10):1319-24.
34. Takagi N, Sugawara O, Sasaki M. Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. Chromosoma. 1982;85(2):275-86.
35. Simon I, Tenzen T, Mostoslavsky R, Fibach E, Lande L, Milot E, et al. Developmental regulation of DNA replication timing at the human beta globin locus. EMBO J. 2001;20(21):6150-7.

36. Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, et al. Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol. 2008;6(10):e245.
37. Dimitrova DS, Gilbert DM. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. Molecular cell. 1999;4(6):983-93.
38. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, et al. Topologically associating domains are stable units of replication-timing regulation. Nature. 2014;515(7527):402-5.
39. Gribnau J, Hochedlinger K, Hata K, Li E, Jaenisch R. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. Genes & development. 2003;17(6):759-73.
40. Grasser F, Neusser M, Fiegler H, Thormeyer T, Cremer M, Carter NP, et al. Replication-timing-correlated spatial chromatin arrangements in cancer and in primate interphase nuclei. J Cell Sci. 2008;121(Pt 11):1876-86.
41. Klemm RD, Austin RJ, Bell SP. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. Cell. 1997;88(4):493-502.
42. Speck C, Chen Z, Li H, Stillman B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. Nat Struct Mol Biol. 2005;12(11):965-71.
43. Speck C, Stillman B. Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. The Journal of biological chemistry. 2007;282(16):11705-14.
44. You Z, Masai H. Cdt1 forms a complex with the minichromosome maintenance protein (MCM) and activates its helicase activity. The Journal of biological chemistry. 2008;283(36):24469-77.
45. Sun J, Evrin C, Samel SA, Fernandez-Cid A, Riera A, Kawakami H, et al. Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2-7 bound to DNA. Nat Struct Mol Biol. 2013;20(8):944-51.
46. Sun J, Fernandez-Cid A, Riera A, Tognetti S, Yuan Z, Stillman B, et al. Structural and mechanistic insights into Mcm2-7 double-hexamer assembly and function. Genes & development. 2014;28(20):2291-303.
47. Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, Diffley JF. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell. 2009;139(4):719-30.
48. Takara TJ, Bell SP. Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2-7 helicases. EMBO J. 2011;30(24):4885-96.
49. Drury LS, Perkins G, Diffley JF. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. Curr Biol. 2000;10(5):231-40.
50. Jallepalli PV, Brown GW, Muzi-Falconi M, Tien D, Kelly TJ. Regulation of the replication initiator protein p65cdc18 by CDK phosphorylation. Genes & development. 1997;11(21):2767-79.
51. Nguyen VQ, Co C, Li JJ. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. Nature. 2001;411(6841):1068-73.
52. Sugimoto N, Tatsumi Y, Tsurumi T, Matsukage A, Kiyono T, Nishitani H, et al. Cdt1 phosphorylation by cyclin A-dependent kinases negatively regulates its function without affecting geminin binding. The Journal of biological chemistry. 2004;279(19):19691-7.
53. Wohlschlegel JA, Dwyer BT, Dhar SK, Cvetic C, Walter JC, Dutta A. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science. 2000;290(5500):2309-12.

54. McGarry TJ, Kirschner MW. Geminin, an inhibitor of DNA replication, is degraded during mitosis. Cell. 1998;93(6):1043-53.
55. Ilves I, Petojevic T, Pesavento JJ, Botchan MR. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Molecular cell. 2010;37(2):247-58.
56. Labib K. How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes & development. 2010;24(12):1208-19.
57. Costa A, Hood IV, Berger JM. Mechanisms for initiating cellular DNA replication. Annual review of biochemistry. 2013;82:25-54.
58. Hardy CF, Dryga O, Seematter S, Pahl PM, Sclafani RA. mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(7):3151-5.
59. Tanaka S, Araki H. Helicase activation and establishment of replication forks at chromosomal origins of replication. Cold Spring Harb Perspect Biol. 2013;5(12):a010371.
60. Yeeles JT, Deegan TD, Janska A, Early A, Diffley JF. Regulated eukaryotic DNA replication origin firing with purified proteins. Nature. 2015;519(7544):431-5.
61. Costa A, Ilves I, Tamberg N, Petojevic T, Nogales E, Botchan MR, et al. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nat Struct Mol Biol. 2011;18(4):471-7.
62. Walter J, Newport J. Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. Molecular cell. 2000;5(4):617-27.
63. Santocanale C, Foiani M, Lucchini G, Plevani P. The isolated 48,000-dalton subunit of yeast DNA primase is sufficient for RNA primer synthesis. The Journal of biological chemistry. 1993;268(2):1343-8.
64. Garg P, Burgers PM. DNA polymerases that propagate the eukaryotic DNA replication fork. Crit Rev Biochem Mol Biol. 2005;40(2):115-28.
65. Chilkova O, Stenlund P, Isoz I, Stith CM, Grabowski P, Lundstrom EB, et al. The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. Nucleic acids research. 2007;35(19):6588-97.
66. Zhang G, Gibbs E, Kelman Z, O'Donnell M, Hurwitz J. Studies on the interactions between human replication factor C and human proliferating cell nuclear antigen. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(5):1869-74.
67. Podust LM, Podust VN, Sogo JM, Hubscher U. Mammalian DNA polymerase auxiliary proteins: analysis of replication factor C-catalyzed proliferating cell nuclear antigen loading onto circular double-stranded DNA. Mol Cell Biol. 1995;15(6):3072-81.
68. Bravo R, Frank R, Blundell PA, Macdonald-Bravo H. Cyclin/PCNA is the auxiliary protein of DNA polymerase-delta. Nature. 1987;326(6112):515-7.
69. Johnson RE, Klassen R, Prakash L, Prakash S. A Major Role of DNA Polymerase delta in Replication of Both the Leading and Lagging DNA Strands. Molecular cell. 2015;59(2):163-75.
70. Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Sugino A. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proceedings of the National Academy of Sciences of the United States of America. 1968;59(2):598-605.

71. Garg P, Stith CM, Sabouri N, Johansson E, Burgers PM. Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. Genes & development. 2004;18(22):2764-73.
72. Goulian M, Richards SH, Heard CJ, Bigsby BM. Discontinuous DNA synthesis by purified mammalian proteins. The Journal of biological chemistry. 1990;265(30):18461-71.
73. Kao HI, Veeraraghavan J, Polaczek P, Campbell JL, Bambara RA. On the roles of Saccharomyces cerevisiae Dna2p and Flap endonuclease 1 in Okazaki fragment processing. The Journal of biological chemistry. 2004;279(15):15014-24.
74. Maine GT, Sinha P, Tye BK. Mutants of S. cerevisiae defective in the maintenance of minichromosomes. Genetics. 1984;106(3):365-85.
75. Sinha P, Chang V, Tye BK. A mutant that affects the function of autonomously replicating sequences in yeast. J Mol Biol. 1986;192(4):805-14.
76. Tye BK. Minichromosome maintenance as a genetic assay for defects in DNA replication. Methods. 1999;18(3):329-34.
77. Forsburg SL. Eukaryotic MCM proteins: beyond replication initiation. Microbiology and molecular biology reviews : MMBR. 2004;68(1):109-31.
78. You Z, Ishimi Y, Masai H, Hanaoka F. Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex. The Journal of biological chemistry. 2002;277(45):42471-9.
79. Sherman DA, Pasion SG, Forsburg SL. Multiple domains of fission yeast Cdc19p (MCM2) are required for its association with the core MCM complex. Mol Biol Cell. 1998;9(7):1833-45.
80. Forsburg SL, Sherman DA, Ottilie S, Yasuda JR, Hodson JA. Mutational analysis of Cdc19p, a Schizosaccharomyces pombe MCM protein. Genetics. 1997;147(3):1025-41.
81. Yan H, Gibson S, Tye BK. Mcm2 and Mcm3, two proteins important for ARS activity, are related in structure and function. Genes & development. 1991;5(6):944-57.
82. Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus. Nature. 1997;390(6658):364-70.
83. Smith DR, Doucette-Stamm LA, Deloughery C, Lee H, Dubois J, Aldredge T, et al. Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics. J Bacteriol. 1997;179(22):7135-55.
84. Tye BK. MCM proteins in DNA replication. Annual review of biochemistry. 1999;68:649-86.
85. Hennessy KM, Lee A, Chen E, Botstein D. A group of interacting yeast DNA replication genes. Genes & development. 1991;5(6):958-69.
86. Thommes P, Kubota Y, Takisawa H, Blow JJ. The RLF-M component of the replication licensing system forms complexes containing all six MCM/P1 polypeptides. EMBO J. 1997;16(11):3312-9.
87. Adachi Y, Usukura J, Yanagida M. A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. Genes Cells. 1997;2(7):467-79.
88. Richter A, Knippers R. High-molecular-mass complexes of human minichromosome-maintenance proteins in mitotic cells. Eur J Biochem. 1997;247(1):136-41.
89. Su TT, Feger G, O'Farrell PH. Drosophila MCM protein complexes. Mol Biol Cell. 1996;7(2):319-29.

90. Ishimi Y. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. The Journal of biological chemistry. 1997;272(39):24508-13.
91. Sherman DA, Forsburg SL. Schizosaccharomyces pombe Mcm3p, an essential nuclear protein, associates tightly with Nda4p (Mcm5p). Nucleic acids research. 1998;26(17):3955-60.
92. Li N, Zhai Y, Zhang Y, Li W, Yang M, Lei J, et al. Structure of the eukaryotic MCM complex at 3.8 A. Nature. 2015.
93. Rao PN, Johnson RT. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. Nature. 1970;225(5228):159-64.
94. Blow JJ, Laskey RA. A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature. 1988;332(6164):546-8.
95. Kubota Y, Mimura S, Nishimoto S, Takisawa H, Nojima H. Identification of the yeast MCM3-related protein as a component of Xenopus DNA replication licensing factor. Cell. 1995;81(4):601-9.
96. Aparicio OM, Weinstein DM, Bell SP. Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell. 1997;91(1):59-69.
97. Wechsler JA, Gross JD. Escherichia coli mutants temperature-sensitive for DNA synthesis. Mol Gen Genet. 1971;113(3):273-84.
98. Labib K, Tercero JA, Diffley JF. Uninterrupted MCM2-7 function required for DNA replication fork progression. Science. 2000;288(5471):1643-7.
99. You Z, Komamura Y, Ishimi Y. Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity. Mol Cell Biol. 1999;19(12):8003-15.
100. Chong JP, Hayashi MK, Simon MN, Xu RM, Stillman B. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(4):1530-5.
101. Lutzmann M, Maiorano D, Mechali M. Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. Gene. 2005;362:51-6.
102. Maiorano D, Cuvier O, Danis E, Mechali M. MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. Cell. 2005;120(3):315-28.
103. Gambus A, Blow JJ. Mcm8 and Mcm9 form a dimeric complex in Xenopus laevis egg extract that is not essential for DNA replication initiation. Cell cycle. 2013;12(8):1225-32.
104. Lutzmann M, Grey C, Traver S, Ganier O, Maya-Mendoza A, Ranisavljevic N, et al. MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. Molecular cell. 2012;47(4):523-34.
105. Nishimura K, Ishiai M, Horikawa K, Fukagawa T, Takata M, Takisawa H, et al. Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. Molecular cell. 2012;47(4):511-22.
106. Hartford SA, Luo Y, Southard TL, Min IM, Lis JT, Schimenti JC. Minichromosome maintenance helicase paralog MCM9 is dispensible for DNA replication but functions in germ-line stem cells and tumor suppression. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(43):17702-7.
107. Blanton HL, Radford SJ, McMahan S, Kearney HM, Ibrahim JG, Sekelsky J. REC, Drosophila MCM8, drives formation of meiotic crossovers. PLoS genetics. 2005;1(3):e40.

108. Kohl KP, Jones CD, Sekelsky J. Evolution of an MCM complex in flies that promotes meiotic crossovers by blocking BLM helicase. Science. 2012;338(6112):1363-5.
109. Park J, Long DT, Lee KY, Abbas T, Shibata E, Negishi M, et al. The MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination. Mol Cell Biol. 2013;33(8):1632-44.
110. Lee KY, Im JS, Shibata E, Park J, Handa N, Kowalczykowski SC, et al. MCM8-9 complex promotes resection of double-strand break ends by MRE11-RAD50-NBS1 complex. Nat Commun. 2015;6:7744.
111. Wood-Trageser MA, Gurbuz F, Yatsenko SA, Jeffries EP, Kotan LD, Surti U, et al. MCM9 mutations are associated with ovarian failure, short stature, and chromosomal instability. Am J Hum Genet. 2014;95(6):754-62.
112. AlAsiri S, Basit S, Wood-Trageser MA, Yatsenko SA, Jeffries EP, Surti U, et al. Exome sequencing reveals MCM8 mutation underlies ovarian failure and chromosomal instability. J Clin Invest. 2015;125(1):258-62.
113. Merchant AM, Kawasaki Y, Chen Y, Lei M, Tye BK. A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. Mol Cell Biol. 1997;17(6):3261-71.
114. Thu YM, Bielinsky AK. MCM10: one tool for all-Integrity, maintenance and damage control. Seminars in cell & developmental biology. 2014;30:121-30.
115. Sawyer SL, Cheng IH, Chai W, Tye BK. Mcm10 and Cdc45 cooperate in origin activation in Saccharomyces cerevisiae. J Mol Biol. 2004;340(2):195-202.
116. Heller RC, Kang S, Lam WM, Chen S, Chan CS, Bell SP. Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. Cell. 2011;146(1):80-91.
117. Zhu W, Ukomadu C, Jha S, Senga T, Dhar SK, Wohlschlegel JA, et al. Mcm10 and And-1/CTF4 recruit DNA polymerase alpha to chromatin for initiation of DNA replication. Genes & development. 2007;21(18):2288-99.
118. Thu YM, Bielinsky AK. Enigmatic roles of Mcm10 in DNA replication. Trends Biochem Sci. 2013;38(4):184-94.
119. Lei M, Kawasaki Y, Tye BK. Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae. Mol Cell Biol. 1996;16(9):5081-90.
120. Mahbubani HM, Chong JP, Chevalier S, Thommes P, Blow JJ. Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor. The Journal of cell biology. 1997;136(1):125-35.
121. Wong PG, Winter SL, Zaika E, Cao TV, Oguz U, Koomen JM, et al. Cdc45 limits replicon usage from a low density of preRCs in mammalian cells. PLoS One. 2011;6(3):e17533.
122. Edwards MC, Tutter AV, Cvetic C, Gilbert CH, Prokhorova TA, Walter JC. MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. The Journal of biological chemistry. 2002;277(36):33049-57.
123. Madine MA, Khoo CY, Mills AD, Musahl C, Laskey RA. The nuclear envelope prevents reinitiation of replication by regulating the binding of MCM3 to chromatin in Xenopus egg extracts. Curr Biol. 1995;5(11):1270-9.
124. Krude T, Musahl C, Laskey RA, Knippers R. Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. J Cell Sci. 1996;109 (Pt 2):309-18.

125. Woodward AM, Gohler T, Luciani MG, Oehlmann M, Ge X, Gartner A, et al. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. The Journal of cell biology. 2006;173(5):673-83.
126. Ibarra A, Schwob E, Mendez J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(26):8956-61.
127. Ge XQ, Jackson DA, Blow JJ. Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. Genes & development. 2007;21(24):3331-41.
128. Pruitt SC, Bailey KJ, Freeland A. Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. Stem Cells. 2007;25(12):3121-32.
129. de Munnik SA, Otten BJ, Schoots J, Bicknell LS, Aftimos S, Al-Aama JY, et al. Meier-Gorlin syndrome: growth and secondary sexual development of a microcephalic primordial dwarfism disorder. Am J Med Genet A. 2012;158A(11):2733-42.
130. Kunnev D, Freeland A, Qin M, Leach RW, Wang J, Shenoy RM, et al. Effect of minichromosome maintenance protein 2 deficiency on the locations of DNA replication origins. Genome Res. 2015;25(4):558-69.
131. Blow JJ, Hodgson B. Replication licensing--defining the proliferative state? Trends in cell biology. 2002;12(2):72-8.
132. Giaginis C, Vgenopoulou S, Vielh P, Theocharis S. MCM proteins as diagnostic and prognostic tumor markers in the clinical setting. Histol Histopathol. 2010;25(3):351-70.
133. Das M, Singh S, Pradhan S, Narayan G. MCM Paradox: Abundance of Eukaryotic Replicative Helicases and Genomic Integrity. Mol Biol Int. 2014;2014:574850.
134. Bagley BN, Keane TM, Maklakova VI, Marshall JG, Lester RA, Cancel MM, et al. A dominantly acting murine allele of Mcm4 causes chromosomal abnormalities and promotes tumorigenesis. PLoS genetics. 2012;8(11):e1003034.
135. Shima N, Alcaraz A, Liachko I, Buske TR, Andrews CA, Munroe RJ, et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nature genetics. 2007;39(1):93-8.
136. Hughes CR, Guasti L, Meimaridou E, Chuang CH, Schimenti JC, King PJ, et al. MCM4 mutation causes adrenal failure, short stature, and natural killer cell deficiency in humans. J Clin Invest. 2012;122(3):814-20.
137. Giaginis C, Georgiadou M, Dimakopoulou K, Tsourouflis G, Gatzidou E, Kouraklis G, et al. Clinical significance of MCM-2 and MCM-5 expression in colon cancer: association with clinicopathological parameters and tumor proliferative capacity. Dig Dis Sci. 2009;54(2):282-91.
138. Saydam O, Senol O, Schaaij-Visser TB, Pham TV, Piersma SR, Stemmer-Rachamimov AO, et al. Comparative protein profiling reveals minichromosome maintenance (MCM) proteins as novel potential tumor markers for meningiomas. J Proteome Res. 2010;9(1):485-94.
139. Cobanoglu U, Mungan S, Gundogdu C, Ersoz S, Ozoran Y, Aydin F. The expression of MCM-2 in invasive breast carcinoma: a stereologic approach. Bratisl Lek Listy. 2010;111(1):45-9.
140. Lau KM, Chan QK, Pang JC, Li KK, Yeung WW, Chung NY, et al. Minichromosome maintenance proteins 2, 3 and 7 in medulloblastoma: overexpression and involvement in regulation of cell migration and invasion. Oncogene. 2010;29(40):5475-89.
141. Giaginis C, Giagini A, Tsourouflis G, Gatzidou E, Agapitos E, Kouraklis G, et al. MCM-2 and MCM-5 expression in gastric adenocarcinoma: clinical significance and comparison with Ki-67 proliferative marker. Dig Dis Sci. 2011;56(3):777-85.

142. Nicol AF, Lapa e Silva JR, Cunha CB, Amaro-Filho SM, Oliveira N, Grinsztejn B, et al. Evaluation of MCM-2 expression in TMA cervical specimens. PLoS One. 2012;7(4):e32936.
143. Wharton SB, Chan KK, Anderson JR, Stoeber K, Williams GH. Replicative Mcm2 protein as a novel proliferation marker in oligodendrogliomas and its relationship to Ki67 labelling index, histological grade and prognosis. Neuropathol Appl Neurobiol. 2001;27(4):305-13.
144. Kato H, Miyazaki T, Fukai Y, Nakajima M, Sohda M, Takita J, et al. A new proliferation marker, minichromosome maintenance protein 2, is associated with tumor aggressiveness in esophageal squamous cell carcinoma. J Surg Oncol. 2003;84(1):24-30.
145. Rodins K, Cheale M, Coleman N, Fox SB. Minichromosome maintenance protein 2 expression in normal kidney and renal cell carcinomas: relationship to tumor dormancy and potential clinical utility. Clin Cancer Res. 2002;8(4):1075-81.
146. Chatrath P, Scott IS, Morris LS, Davies RJ, Rushbrook SM, Bird K, et al. Aberrant expression of minichromosome maintenance protein-2 and Ki67 in laryngeal squamous epithelial lesions. Br J Cancer. 2003;89(6):1048-54.
147. Gonzalez MA, Pinder SE, Callagy G, Vowler SL, Morris LS, Bird K, et al. Minichromosome maintenance protein 2 is a strong independent prognostic marker in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2003;21(23):4306-13.
148. Obermann EC, Went P, Zimpfer A, Tzankov A, Wild PJ, Stoehr R, et al. Expression of minichromosome maintenance protein 2 as a marker for proliferation and prognosis in diffuse large B-cell lymphoma: a tissue microarray and clinico-pathological analysis. BMC Cancer. 2005;5:162.
149. Szelachowska J, Dziegiel P, Jelen-Krzeszewska J, Jelen M, Matkowski R, Pomiecko A, et al. Mcm-2 protein expression predicts prognosis better than Ki-67 antigen in oral cavity squamocellular carcinoma. Anticancer Res. 2006;26(3B):2473-8.
150. Gakiopoulou H, Korkolopoulou P, Levidou G, Thymara I, Saetta A, Piperi C, et al. Minichromosome maintenance proteins 2 and 5 in non-benign epithelial ovarian tumours: relationship with cell cycle regulators and prognostic implications. Br J Cancer. 2007;97(8):1124-34.
151. Liu M, Li JS, Tian DP, Huang B, Rosqvist S, Su M. MCM2 expression levels predict diagnosis and prognosis in gastric cardiac cancer. Histol Histopathol. 2013;28(4):481-92.
152. Lee YS, Ha SA, Kim HJ, Shin SM, Kim HK, Kim S, et al. Minichromosome maintenance protein 3 is a candidate proliferation marker in papillary thyroid carcinoma. Exp Mol Pathol. 2010;88(1):138-42.
153. Kikuchi J, Kinoshita I, Shimizu Y, Kikuchi E, Takeda K, Aburatani H, et al. Minichromosome maintenance (MCM) protein 4 as a marker for proliferation and its clinical and clinicopathological significance in non-small cell lung cancer. Lung Cancer. 2011;72(2):229-37.
154. Dudderidge TJ, Kelly JD, Wollenschlaeger A, Okoturo O, Prevost T, Robson W, et al. Diagnosis of prostate cancer by detection of minichromosome maintenance 5 protein in urine sediments. Br J Cancer. 2010;103(5):701-7.
155. Zheng T, Chen M, Han S, Zhang L, Bai Y, Fang X, et al. Plasma minichromosome maintenance complex component 6 is a novel biomarker for hepatocellular carcinoma patients. Hepatol Res. 2014;44(13):1347-56.
156. Ren B, Yu G, Tseng GC, Cieply K, Gavel T, Nelson J, et al. MCM7 amplification and overexpression are associated with prostate cancer progression. Oncogene. 2006;25(7):1090-8.

157. Zhong X, Chen X, Guan X, Zhang H, Ma Y, Zhang S, et al. Overexpression of G9a and MCM7 in oesophageal squamous cell carcinoma is associated with poor prognosis. Histopathology. 2015;66(2):192-200.
158. Tamura T, Shomori K, Haruki T, Nosaka K, Hamamoto Y, Shiomi T, et al. Minichromosome maintenance-7 and geminin are reliable prognostic markers in patients with oral squamous cell carcinoma: immunohistochemical study. J Oral Pathol Med. 2010;39(4):328-34.
159. Fujioka S, Shomori K, Nishihara K, Yamaga K, Nosaka K, Araki K, et al. Expression of minichromosome maintenance 7 (MCM7) in small lung adenocarcinomas (pT1): Prognostic implication. Lung Cancer. 2009;65(2):223-9.
160. Nishihara K, Shomori K, Fujioka S, Tokuyasu N, Inaba A, Osaki M, et al. Minichromosome maintenance protein 7 in colorectal cancer: implication of prognostic significance. Int J Oncol. 2008;33(2):245-51.
161. Branzei D, Foiani M. Maintaining genome stability at the replication fork. Nature reviews Molecular cell biology. 2010;11(3):208-19.
162. Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Molecular cell. 2010;40(2):179-204.
163. Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nature cell biology. 2014;16(1):2-9.
164. Barlow JH, Faryabi RB, Callen E, Wong N, Malhowski A, Chen HT, et al. Identification of early replicating fragile sites that contribute to genome instability. Cell. 2013;152(3):620-32.
165. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. Molecular cell. 2012;46(2):115-24.
166. Le Tallec B, Koundrioukoff S, Wilhelm T, Letessier A, Brison O, Debatisse M. Updating the mechanisms of common fragile site instability: how to reconcile the different views? Cell Mol Life Sci. 2014;71(23):4489-94.
167. Le Tallec B, Millot GA, Blin ME, Brison O, Dutrillaux B, Debatisse M. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. Cell Rep. 2013;4(3):420-8.
168. Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Molecular cell. 2011;44(6):966-77.
169. Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, et al. Nucleotide deficiency promotes genomic instability in early stages of cancer development. Cell. 2011;145(3):435-46.
170. Beck H, Nahse-Kumpf V, Larsen MS, O'Hanlon KA, Patzke S, Holmberg C, et al. Cyclin-dependent kinase suppression by WEE1 kinase protects the genome through control of replication initiation and nucleotide consumption. Mol Cell Biol. 2012;32(20):4226-36.
171. Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes & development. 2005;19(9):1040-52.
172. Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. Nature reviews Molecular cell biology. 2008;9(8):616-27.
173. Nam EA, Cortez D. ATR signalling: more than meeting at the fork. Biochem J. 2011;436(3):527-36.

174. Labib K, De Piccoli G. Surviving chromosome replication: the many roles of the S-phase checkpoint pathway. Philos Trans R Soc Lond B Biol Sci. 2011;366(1584):3554-61.
175. Mahaney BL, Meek K, Lees-Miller SP. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. Biochem J. 2009;417(3):639-50.
176. Mah LJ, El-Osta A, Karagiannis TC. gammaH2AX: a sensitive molecular marker of DNA damage and repair. Leukemia. 2010;24(4):679-86.
177. Sirbu BM, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D. Analysis of protein dynamics at active, stalled, and collapsed replication forks. Genes & development. 2011;25(12):1320-7.
178. Lopes M, Cotta-Ramusino C, Pellicioli A, Liberi G, Plevani P, Muzi-Falconi M, et al. The DNA replication checkpoint response stabilizes stalled replication forks. Nature. 2001;412(6846):557-61.
179. De Piccoli G, Katou Y, Itoh T, Nakato R, Shirahige K, Labib K. Replisome stability at defective DNA replication forks is independent of S phase checkpoint kinases. Molecular cell. 2012;45(5):696-704.
180. Petermann E, Helleday T. Pathways of mammalian replication fork restart. Nature reviews Molecular cell biology. 2010;11(10):683-7.
181. Santocanale C, Diffley JF. A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. Nature. 1998;395(6702):615-8.
182. Shechter D, Costanzo V, Gautier J. ATR and ATM regulate the timing of DNA replication origin firing. Nature cell biology. 2004;6(7):648-55.
183. Trenz K, Errico A, Costanzo V. Plx1 is required for chromosomal DNA replication under stressful conditions. EMBO J. 2008;27(6):876-85.
184. Ge XQ, Blow JJ. Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. The Journal of cell biology. 2010;191(7):1285-97.
185. Maya-Mendoza A, Petermann E, Gillespie DA, Caldecott KW, Jackson DA. Chk1 regulates the density of active replication origins during the vertebrate S phase. EMBO J. 2007;26(11):2719-31.
186. Boutros R, Dozier C, Ducommun B. The when and wheres of CDC25 phosphatases. Current opinion in cell biology. 2006;18(2):185-91.
187. d'Adda di Fagagna F. Living on a break: cellular senescence as a DNA-damage response. Nat Rev Cancer. 2008;8(7):512-22.
188. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. Nature. 1990;345(6274):458-60.
189. Chin L, Artandi SE, Shen Q, Tam A, Lee SL, Gottlieb GJ, et al. p53 deficiency rescues the adverse effects of telomere loss and cooperates with telomere dysfunction to accelerate carcinogenesis. Cell. 1999;97(4):527-38.
190. Parrinello S, Samper E, Krtolica A, Goldstein J, Melov S, Campisi J. Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. Nature cell biology. 2003;5(8):741-7.
191. Di Micco R, Fumagalli M, d'Adda di Fagagna F. Breaking news: high-speed race ends in arrest--how oncogenes induce senescence. Trends in cell biology. 2007;17(11):529-36.
192. Lazzerini Denchi E, Attwooll C, Pasini D, Helin K. Deregulated E2F activity induces hyperplasia and senescence-like features in the mouse pituitary gland. Mol Cell Biol. 2005;25(7):2660-72.

193. d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T, et al. A DNA damage checkpoint response in telomere-initiated senescence. Nature. 2003;426(6963):194-8.
194. Karlseder J, Hoke K, Mirzoeva OK, Bakkenist C, Kastan MB, Petrini JH, et al. The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. PLoS Biol. 2004;2(8):E240.
195. Denchi EL, de Lange T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature. 2007;448(7157):1068-71.
196. Bartkova J, Rezaei N, Liontos M, Karakaidos P, Kletsas D, Issaeva N, et al. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. Nature. 2006;444(7119):633-7.
197. Campisi J, d'Adda di Fagagna F. Cellular senescence: when bad things happen to good cells. Nature reviews Molecular cell biology. 2007;8(9):729-40.
198. Gonzalez-Suarez E, Samper E, Flores JM, Blasco MA. Telomerase-deficient mice with short telomeres are resistant to skin tumorigenesis. Nature genetics. 2000;26(1):114-7.
199. Rudolph KL, Chang S, Lee HW, Blasco M, Gottlieb GJ, Greider C, et al. Longevity, stress response, and cancer in aging telomerase-deficient mice. Cell. 1999;96(5):701-12.
200. Leri A, Franco S, Zacheo A, Barlucchi L, Chimenti S, Limana F, et al. Ablation of telomerase and telomere loss leads to cardiac dilatation and heart failure associated with p53 upregulation. EMBO J. 2003;22(1):131-9.
201. van Deursen JM. The role of senescent cells in ageing. Nature. 2014;509(7501):439-46.
202. Shreeram S, Sparks A, Lane DP, Blow JJ. Cell type-specific responses of human cells to inhibition of replication licensing. Oncogene. 2002;21(43):6624-32.
203. Machida YJ, Teer JK, Dutta A. Acute reduction of an origin recognition complex (ORC) subunit in human cells reveals a requirement of ORC for Cdk2 activation. The Journal of biological chemistry. 2005;280(30):27624-30.
204. Liu P, Slater DM, Lenburg M, Nevis K, Cook JG, Vaziri C. Replication licensing promotes cyclin D1 expression and G1 progression in untransformed human cells. Cell cycle. 2009;8(1):125-36.
205. Nevis KR, Cordeiro-Stone M, Cook JG. Origin licensing and p53 status regulate Cdk2 activity during G(1). Cell cycle. 2009;8(12):1952-63.
206. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell. 2009;136(4):642-55.
207. Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature. 2004;431(7006):343-9.
208. Tomari Y, Zamore PD. Perspective: machines for RNAi. Genes & development. 2005;19(5):517-29.
209. Okamura K, Lai EC. Endogenous small interfering RNAs in animals. Nature reviews Molecular cell biology. 2008;9(9):673-8.
210. Ishizu H, Siomi H, Siomi MC. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes & development. 2012;26(21):2361-73.
211. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993;75(5):843-54.
212. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology. 2009;10(2):126-39.

213. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA. 2005;11(3):241-7.
214. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116(2):281-97.
215. Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, et al. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell. 2006;125(5):887-901.
216. Westholm JO, Lai EC. Mirtrons: microRNA biogenesis via splicing. Biochimie. 2011;93(11):1897-904.
217. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes & development. 2003;17(24):3011-6.
218. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Molecular cell. 2007;27(1):91-105.
219. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009;136(2):215-33.
220. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19(1):92-105.
221. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005;433(7027):769-73.
222. Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science. 2004;304(5670):594-6.
223. Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH. Substrate requirements for let-7 function in the developing zebrafish embryo. Nucleic acids research. 2004;32(21):6284-91.
224. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120(1):15-20.
225. Saetrom P, Heale BS, Snove O, Jr., Aagaard L, Alluin J, Rossi JJ. Distance constraints between microRNA target sites dictate efficacy and cooperativity. Nucleic acids research. 2007;35(7):2333-42.
226. Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell. 2005;123(4):607-20.
227. Mathonnet G, Fabian MR, Svitkin YV, Parsyan A, Huck L, Murata T, et al. MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. Science. 2007;317(5845):1764-7.
228. Kiriakidou M, Tan GS, Lamprinaki S, De Planell-Saguer M, Nelson PT, Mourelatos Z. An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell. 2007;129(6):1141-51.
229. Chendrimada TP, Finn KJ, Ji X, Baillat D, Gregory RI, Liebhaber SA, et al. MicroRNA silencing through RISC recruitment of eIF6. Nature. 2007;447(7146):823-8.
230. Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes & development. 2006;20(14):1885-98.
231. Wu L, Fan J, Belasco JG. MicroRNAs direct rapid deadenylation of mRNA. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(11):4034-9.

232. Petersen CP, Bordeleau ME, Pelletier J, Sharp PA. Short RNAs repress translation after initiation in mammalian cells. Molecular cell. 2006;21(4):533-42.
233. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet. 2008;9(2):102-14.
234. Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. Molecular cell. 2007;25(5):635-46.
235. Adjibade P, Mazroui R. Control of mRNA turnover: implication of cytoplasmic RNA granules. Seminars in cell & developmental biology. 2014;34:15-23.
236. Pillai RS, Bhattacharyya SN, Artus CG, Zoller T, Cougot N, Basyuk E, et al. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. 2005;309(5740):1573-6.
237. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol Cell Biol. 2007;27(11):3970-81.
238. Bell SP, Stillman B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature. 1992;357(6374):128-34.
239. Moyer SE, Lewis PW, Botchan MR. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(27):10236-41.
240. Bochman ML, Schwacha A. The Mcm2-7 complex has in vitro helicase activity. Molecular cell. 2008;31(2):287-93.
241. Errico A, Costanzo V, Hunt T. Tipin is required for stalled replication forks to resume DNA replication after removal of aphidicolin in Xenopus egg extracts. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(38):14929-34.
242. Bailis JM, Luche DD, Hunter T, Forsburg SL. Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. Molecular and cellular biology. 2008;28(5):1724-38.
243. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. Nature reviews Molecular cell biology. 2010;11(3):220-8.
244. Allen C, Ashley AK, Hromas R, Nickoloff JA. More forks on the road to replication stress recovery. Journal of molecular cell biology. 2011;3(1):4-12.
245. Feng D, Tu Z, Wu W, Liang C. Inhibiting the expression of DNA replication-initiation proteins induces apoptosis in human cancer cells. Cancer research. 2003;63(21):7356-64.
246. Blow JJ, Ge XQ, Jackson DA. How dormant origins promote complete genome replication. Trends Biochem Sci. 2011;36(8):405-14.
247. Blow JJ, Gillespie PJ. Replication licensing and cancer--a fatal entanglement? Nature reviews Cancer. 2008;8(10):799-806.
248. Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA, et al. Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell. 1997;91(5):649-59.
249. Chuang CH, Wallace MD, Abratte C, Southard T, Schimenti JC. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. PLoS genetics. 2010;6(9):e1001110.
250. Chuang CH, Yang D, Bai G, Freeland A, Pruitt SC, Schimenti JC. Post-transcriptional homeostasis and regulation of MCM2-7 in mammalian cells. Nucleic acids research. 2012;40(11):4914-24.

251. Kawabata T, Luebben SW, Yamaguchi S, Ilves I, Matise I, Buske T, et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. Molecular cell. 2011;41(5):543-53.
252. Kunnev D, Rusiniak ME, Kudla A, Freeland A, Cady GK, Pruitt SC. DNA damage response and tumorigenesis in Mcm2-deficient mice. Oncogene. 2010;29(25):3630-8.
253. Debacq-Chainiaux F, Erusalimsky JD, Campisi J, Toussaint O. Protocols to detect senescence-associated beta-galactosidase (SA-betagal) activity, a biomarker of senescent cells in culture and in vivo. Nature protocols. 2009;4(12):1798-806.
254. Monasor A, Murga M, Lopez-Contreras AJ, Navas C, Gomez G, Pisano DG, et al. INK4a/ARF limits the expansion of cells suffering from replication stress. Cell cycle. 2013;12(12):1948-54.
255. Flach J, Bakker ST, Mohrin M, Conroy PC, Pietras EM, Reynaud D, et al. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. Nature. 2014;512(7513):198-202.
256. Kawabata T, Yamaguchi S, Buske T, Luebben SW, Wallace M, Matise I, et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. Mammalian genome : official journal of the International Mammalian Genome Society. 2011;22(9-10):506-17.
257. Itahana K, Campisi J, Dimri GP. Mechanisms of cellular senescence in human and mouse cells. Biogerontology. 2004;5(1):1-10.
258. Wallace MD, Southard TL, Schimenti KJ, Schimenti JC. Role of DNA damage response pathways in preventing carcinogenesis caused by intrinsic replication stress. Oncogene. 2014;33(28):3688-95.
259. Ge XQ, Han J, Cheng EC, Yamaguchi S, Shima N, Thomas JL, et al. Embryonic Stem Cells License a High Level of Dormant Origins to Protect the Genome against Replication Stress. Stem Cell Reports. 2015;5(2):185-94.
260. Wallace MD, Pfefferle AD, Shen L, McNairn AJ, Cerami EG, Fallon BL, et al. Comparative oncogenomics implicates the neurofibromin 1 gene (NF1) as a breast cancer driver. Genetics. 2012;192(2):385-96.
261. Hoang ML, Leon RP, Pessoa-Brandao L, Hunt S, Raghuraman MK, Fangman WL, et al. Structural changes in Mcm5 protein bypass Cdc7-Dbf4 function and reduce replication origin efficiency in Saccharomyces cerevisiae. Mol Cell Biol. 2007;27(21):7594-602.
262. Alabert C, Bukowski-Wills JC, Lee SB, Kustatscher G, Nakamura K, de Lima Alves F, et al. Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. Nature cell biology. 2014;16(3):281-93.
263. Kliszczak AE, Rainey MD, Harhen B, Boisvert FM, Santocanale C. DNA mediated chromatin pull-down for the study of chromatin replication. Scientific reports. 2011;1:95.
264. Sirbu BM, Couch FB, Cortez D. Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. Nat Protoc. 2012;7(3):594-605.
265. Sirbu BM, McDonald WH, Dungrawala H, Badu-Nkansah A, Kavanaugh GM, Chen Y, et al. Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. The Journal of biological chemistry. 2013;288(44):31458-67.

266. Lopez-Contreras AJ, Ruppen I, Nieto-Soler M, Murga M, Rodriguez-Acebes S, Remeseiro S, et al. A proteomic characterization of factors enriched at nascent DNA molecules. Cell Rep. 2013;3(4):1105-16.
267. Ahuja D, Saenz-Robles MT, Pipas JM. SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. Oncogene. 2005;24(52):7729-45.
268. Dungrawala H, Rose KL, Bhat KP, Mohni KN, Glick GG, Couch FB, et al. The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. Molecular cell. 2015.
269. Groth A, Corpet A, Cook AJ, Roche D, Bartek J, Lukas J, et al. Regulation of replication fork progression through histone supply and demand. Science. 2007;318(5858):1928-31.
270. Lu J, Li F, Murphy CS, Davidson MW, Gilbert DM. G2 phase chromatin lacks determinants of replication timing. The Journal of cell biology. 2010;189(6):967-80.
271. Brewster AS, Wang G, Yu X, Greenleaf WB, Carazo JM, Tjajadi M, et al. Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(51):20191-6.
272. Li XC, Schimenti JC, Tye BK. Aneuploidy and improved growth are coincident but not causal in a yeast cancer model. PLoS Biol. 2009;7(7):e1000161.
273. Cheng CY, Hwang CI, Corney DC, Flesken-Nikitin A, Jiang L, Oner GM, et al. miR-34 cooperates with p53 in suppression of prostate cancer by joint regulation of stem cell compartment. Cell Rep. 2014;6(6):1000-7.
274. Jin HY, Oda H, Lai M, Skalsky RL, Bethel K, Shepherd J, et al. MicroRNA-17~92 plays a causative role in lymphomagenesis by coordinating multiple oncogenic pathways. EMBO J. 2013;32(17):2377-91.
275. Feng Z, Zhang C, Wu R, Hu W. Tumor suppressor p53 meets microRNAs. Journal of molecular cell biology. 2011;3(1):44-50.
276. Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K. Modulation of microRNA processing by p53. Nature. 2009;460(7254):529-33.
277. He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, et al. A microRNA component of the p53 tumour suppressor network. Nature. 2007;447(7148):1130-4.
278. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(10):R106.
279. Ohtani K, Iwanaga R, Nakamura M, Ikeda M, Yabuta N, Tsuruga H, et al. Cell growth-regulated expression of mammalian MCM5 and MCM6 genes mediated by the transcription factor E2F. Oncogene. 1999;18(14):2299-309.
280. Lal A, Thomas MP, Altschuler G, Navarro F, O'Day E, Li XL, et al. Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling. PLoS genetics. 2011;7(11):e1002363.
281. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, et al. A microRNA polycistron as a potential human oncogene. Nature. 2005;435(7043):828-33.
282. Kaller M, Liffers ST, Oeljeklaus S, Kuhlmann K, Roh S, Hoffmann R, et al. Genome-wide characterization of miR-34a induced changes in protein and mRNA expression by a combined pulsed SILAC and microarray analysis. Mol Cell Proteomics. 2011;10(8):M111 010462.

283. Kuipers MA, Stasevich TJ, Sasaki T, Wilson KA, Hazelwood KL, McNally JG, et al. Highly stable loading of Mcm proteins onto chromatin in living cells requires replication to unload. The Journal of cell biology. 2011;192(1):29-41.
284. Simon N, Bochman ML, Seguin S, Brodsky JL, Seibel WL, Schwacha A. Ciprofloxacin is an inhibitor of the Mcm2-7 replicative helicase. Biosci Rep. 2013;33(5).