Single molecule MATAC-seq reveals key determinants for DNA replication efficiency

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades „Doctor rerum naturalium“ (Dr. rer. nat.) an der Fakultät für Biologie der Ludwig-Maximilians-Universität München

Anna Chanou, M. Sc.

Course of Study

Biology (Dr. rer. nat.)

Ludwig-Maximilians-Universität München

Munich, 2023

Diese Dissertation wurde angefertigt

unter der Leitung von Prof. Dr. Maria-Elena Torres-Padilla

im Bereich von Institute of Epigenetics and Stem Cells

an der Ludwig-Maximilians-Universität München

Ersgutachter/in:	Prof. Dr. Maria - Elena Torres - Padilla
Zweitgutachter/in:	Prof. Dr. Christoph Osman
Tag der Abgabe:	01.03.2023
Tag der mündlichen Prüfung:	27.10.2023

EIDESSTATTLICHE RKLÄRUNG

Ich versichere hier mit an Eides statt, dass meine Dissertation selbständig und ohne unerlaubte Hilfsmittel angefertigt worden ist.

Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungs kommission vorgelegt.

Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 21.11.2023

Anna Chanou

Summary

The stochastic nature of origin activation results in significant variability in the way genome replication is carried out from cell to cell. The reason for the diversity in efficiency and timing of individual origins has remained an unresolved issue for a long time. Cell-to-cell variability has been demonstrated to play a crucial role in cellular plasticity and cancer in mammalian cells. Although population-based methods have provided valuable insight into biological processes, it is necessary to use single molecule techniques to uncover events that are hidden by the population average. Many biological processes, such as DNA replication, transcription, and gene expression, are closely linked to the local chromatin structure. In yeast, although DNA replication origins have conserved DNA sequences, they display remarkable differences in timing and efficiency. Some origins initiate replication earlier during S-phase or more frequently than others, resulting in a high degree of heterogeneity among the cells in a population, with no two cells having the exact same replication profile.

Our hypothesis is that the local nucleosomal structure may affect the DNA replication profile of individual origins. To explore this relationship, we have developed Methylation Accessibility of Targeted Chromatin domain Sequencing (MATAC-Seq) to determine single-molecule chromatin accessibility maps of specific genomic locations after targeted purification in their native chromatin context. Our analysis of selected early-efficient (EE) and late-inefficient (LI) replication origins in *Saccharomyces cerevisiae* using MATAC-Seq revealed significant cell-to-cell heterogeneity in their chromatin states. Disrupting the INO80 or ISW2 chromatin remodeling complexes led to changes at individual nucleosomal positions that correspond to changes in replication efficiency. Our results show that a chromatin state with a narrow size of accessible origin DNA in combination with well-positioned surrounding nucleosomes and an open +2 linker region was a strong predictor for efficient origin activation.

MATAC-Seq provides a single-molecule assay for chromatin accessibility that reveals the large spectrum of alternative chromatin states that coexist at a given locus, which was previously masked in population-based experiments. This provides a mechanistic basis for origin activation heterogeneity that occurs during DNA replication in eukaryotic cells. As a result, our single-molecule assay for chromatin accessibility will be ideal for defining single-molecule heterogeneity across many biological processes, such as transcription, replication, or DNA repair *in vitro* and *ex vivo*.

Zusammenfassung

Die stochastische Natur der Aktivierung von Replikationsursprüngen führt zu einer signifikanten Variabilität in der Art und Weise, wie die DNA Replikation von Zelle zu Zelle durchgeführt wird. Der Grund für die Diversität in Effizienz und Zeitpunkt der individuellen Aktivierung von Ursprüngen blieb lange ein ungelöstes Problem. Es wurde gezeigt, dass die Zell-zu-Zell-Variabilität eine entscheidende Rolle bei der zellulären Plastizität und Krebs in Säugetierzellen spielt. Obwohl populationsbasierte Methoden wertvolle Einblicke in biologische Prozesse geliefert haben, ist es notwendig, Einzelmolekültechniken zu verwenden, um Ereignisse aufzudecken, die durch das Durchschnittsverhalten aller Moleküle verborgen sind. Viele biologische Prozesse wie DNA-Replikation, Transkription und Genexpression sind eng mit der lokalen Chromatinstruktur verbunden. Obwohl die DNA-Replikationsursprünge in Hefe konservierte DNA-Sequenzen aufweisen, zeigen sie bemerkenswerte Unterschiede im Zeitpunkt und Effizienz der Replikation. Einige Ursprünge initiieren die Replikation früher während der S-Phase oder häufiger als andere, was zu einem hohen Grad an Heterogenität zwischen den Zellen in einer Population führt, wobei keine zwei Zellen das exakt gleiche Replikationsprofil aufweisen.

Unsere Hypothese ist, dass die lokale nukleosomale Struktur das DNA-Replikationsprofil beeinflussen kann. Um diese Beziehung zu untersuchen, haben wir Methylation Accessibility of Targeted Chromatin Domain Sequencing (MATAC-Seq) entwickelt, um Einzelmolekül-Chromatin-Zugänglichkeitskarten spezifischer genomischer Orte nach gezielter Reinigung in ihrem nativen Chromatin-Kontext zu bestimmen. Unsere Analyse ausgewählter früh-effizient (EE) und spät-ineffizient (LI) feuernde Replikationsursprünge in *Saccharomyces cerevisiae* mit MATAC-Seq ergab eine signifikante Zell-zu-Zell-Heterogenität in ihren Chromatinzuständen. Die genetische Deletion der INO80- oder ISW2-Chromatin-Remodeling Komplexe führte zu Veränderungen an einzelnen nukleosomalen Positionen, die mit Veränderungen der Replikationseffizienz korrespondierten. Unsere Ergebnisse zeigten, dass ein Chromatinzustand mit einem engen Fenster an zugänglicher Replikationsursprungs-DNA in Kombination mit gut positionierten umgebenden Nukleosomen und einer offenen +2-Linkerregion ein starker Prädiktor für eine effiziente Ursprungsaktivierung war.

MATAC-Seq bietet einen Einzelmolekül-Assay für die Zugänglichkeit von Chromatin, der das große Spektrum alternativer Chromatinzustände aufzeigt, die an einem bestimmten genomischen Lokus koexistieren, der zuvor in populationsbasierten Experimenten maskiert war. Dies liefert eine mechanistische Grundlage für die Heterogenität der Ursprungsaktivierung, die während der DNA-Replikation in eukaryotischen Zellen auftritt. Infolgedessen ist unser Einzelmolekül-Assay

für Chromatin-Zugänglichkeit ideal für die Definition der Einzelmolekül-Heterogenität über viele biologische Prozesse hinweg, wie z. B. Transkription, Replikation oder DNA-Reparatur *in vitro* und *ex vivo*.

1. Chromatin

The outstanding length and complexity of eukaryotic genomes poses multiple challenges to cells. The 3D space in the nucleus is limited and therefore requires a high degree of DNA compaction, while providing sufficient accessibility for gene expression at the same time. In order to overcome this fundamental challenge, cells wrap the genome into a nucleoprotein complex named chromatin (Igo-Kemenes et al., 1982). Chromatin has a center role for all DNA-dependent transactions, including transcription, replication, DNA repair and recombination, thus relying on tight regulation (Kornberg and Lorch, 2020).

The basic unit of chromatin is the nucleosome, a histone octamer protein complex wrapped with 147 bp of DNA. Each octamer contains two histone H3-H4 dimers that are connected with two histone H2A-H2B dimers (Luger, 1997a; McGuffee et al., 2013). These nucleosome core particles (NCPs) are separated by short 15–50 bp of linker DNA, giving chromatin the typical “beads on a string” appearance along linear DNA molecules (**Figure 1**). In general, DNA and octamers form tight interactions along their surface, thereby inhibiting most nuclear processes that need to gain access to the DNA. Thus, nucleosomes must be dynamically repositioned or evicted so that specific effector proteins can access regulatory DNA sites within nucleosomal DNA. Apart from intrinsically favorable or unfavorable DNA sequences for nucleosome formation (Travers et al., 2010), other cellular components, such as transcription factors and ATP-dependent chromatin remodeling machines contribute to the chromatin landscape *in vivo* (Dou and Gorovsky, 2000; Jenuwein, 2001; Lusser and Kadonaga, 2003; Heintzman et al., 2009; Bell et al., 2011). In addition, histones contain unstructured N-terminal tails that extend out of the nucleosomal core. These polypeptide chains contain many residues that can be targeted for addition or removal of a multitude of post-translational modifications (PTMs). A wide number of histone modifications, such as methylation of arginines and lysines, the acetylation of lysines, and the phosphorylation of threonines and serines have been found to be crucial for chromatin state regulation, gene expression, and DNA replication (Tyler et al., 1999; Strahl and Allis, 2000; Jenuwein, 2001; Lusser and Kadonaga, 2003; Gibney and Nolan, 2010). Combining these diverse PTMs in a combinatorial way exponentially increases the heterogeneity of distinct chromatin states that can co-exist at a given genomic locus.

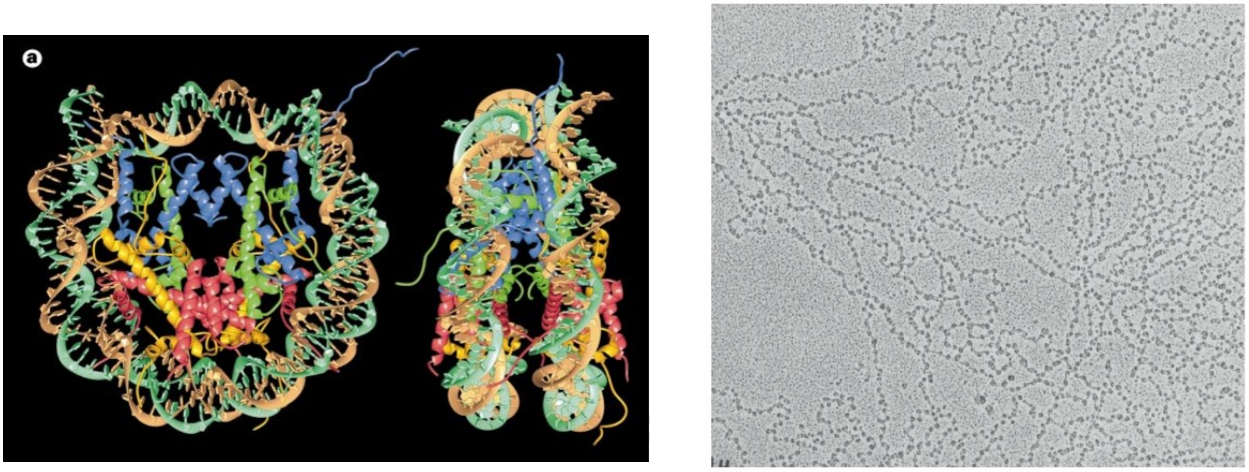


Figure 1. A. Nucleosomal particles: histone proteins are wrapped around 147bp of DNA (brown and green chain). H3 histone is indicated by blue, H4 histone by light green, H2A by yellow and H2B by red color. The nucleosome core particles are presented from two perspectives (modified from(Luger, 1997). **B.** Visualization of Drosophila chromatin under electron microscope showing the nucleosomes are beads connected by the DNA-string (modified by (Baldi et al., 2020).

Various DNA binding proteins and interactions between nucleosomes (primary chromatin structure) contribute to further genome organization into a 30nm chromatin fiber (secondary structure) and into higher-order chromosome compaction in the nucleus (Cremer and Cremer, 2001). The linker histone H1 appears to play a crucial role for the formation or stabilization of higher-order chromatin structures (Thoma and Keller,1977). Initially, two models had been proposed for the organization of the secondary structure: the solenoid model and the zigzag model (**Figure 2**). The solenoid model consists of nucleosomes assembled in a consecutive anti-clockwise manner that result in an interdigitated one-start helix in which a nucleosome in the fiber interacts with its fifth and sixth neighbor nucleosomes. A regular helical loop contains 6 nucleosomes per turn connected by a slightly bended linker DNA (Finch and Klug, 1976; Tremethick, 2007). In the zigzag model, which is the more accepted model consistent with available data, the chromatin fiber is a two-start helix in which nucleosomes are arranged in a zigzag fashion such that a nucleosome in the fiber binds to the second nucleosome in proximity (Bednar et al., 1998; Dorigo et al., 2004; Schalch et al., 2005) . The nucleosomes are organized face-to-face in two rows and the linker DNA is straight across the nucleosomes (Chen and Li, 2010). From zigzag formation, chromatin is further packaged and condensed into chromosomes which occupy discrete territories building the whole nuclear architecture. The radial organization of chromosome territories is evolutionally conserved and strongly correlated with gene density.

The nuclear periphery tends to harbor chromosomal domains containing less genes, whereas the central areas are gene-rich (Sun et al., 2000; Tanabe et al., 2002; Mayer et al., 2005).

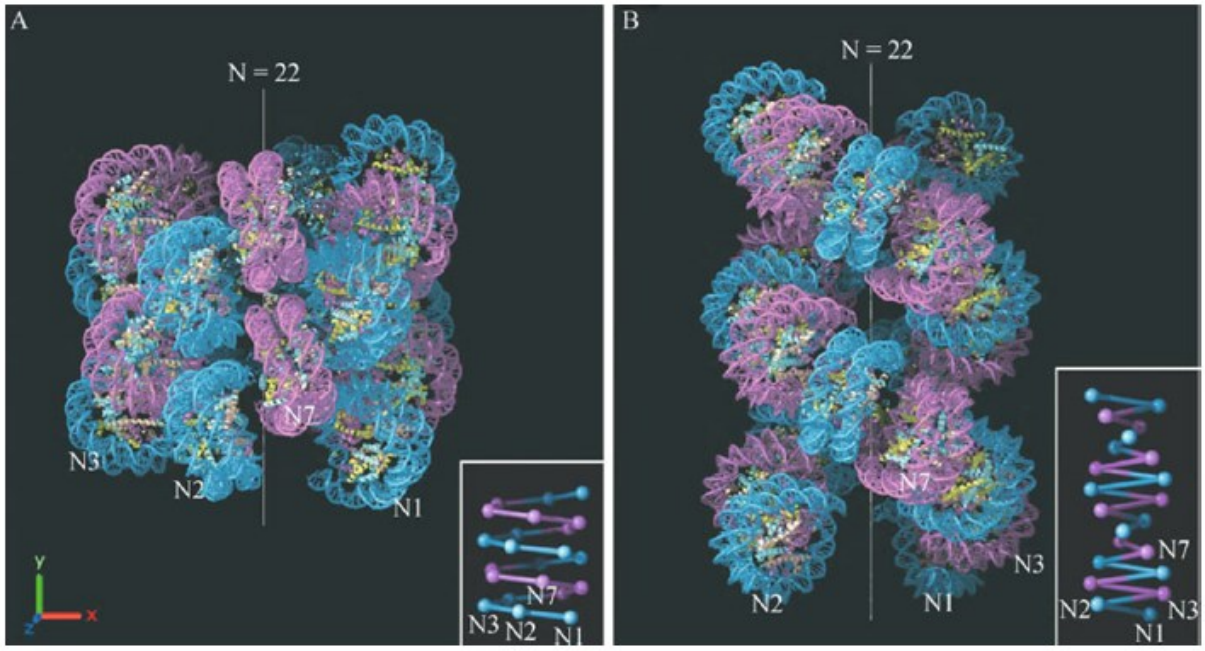


Figure 2. A. Solenoid and **B.** Zig-Zag chromatin organization model. The nucleosomal positions are indicated by N1, N2, N3 and N7 in both cases (adapted from Robinson et al., 2006).

2. DNA replication

A nuclear process that is fundamental and highly conserved in all cellular organisms is DNA replication. Eukaryotic cells initiate replication from many genomic locations known as replication origins that are scattered throughout chromosomes. This replication strategy allows cells to uncouple genome size from S-phase length, as the chromosomes are simultaneously replicated by many different DNA replication machines (replisomes), thereby reducing S-phase length. The size of the *Saccharomyces cerevisiae* (hereafter named yeast) is 14.4Mb (Karaoglu et al., 2005) and its duplication is achieved by the distribution of ~400 replication origins along the 16 chromosomes. For such large numbers of origins, however, the firing of individual origins must be tightly regulated and coordinated so that the genome is copied exactly once resulting in error-free transmission of genetic information to the next generation. To avoid re-replication of the DNA and, therefore multiple gene amplification, genomic instability and cancer (Lengauer et al., 1998), DNA replication is a strictly regulated process and is initiated by two timely separated, however tightly interdependent, steps: origin licensing and origin firing.

2.1 Origin licensing

A six-subunit complex, named as origin recognition complex (ORC) is the first molecular complex binding to the double-stranded DNA of replication origins (**Figure 3**). The 5 homologous subunits Orc1-5 form a crescent moon shape and together form a surface essential for the DNA recognition by ORC (Lee and Bell, 1997). Even though Orc6 is not required for the initial binding, it is crucial for DNA replication initiation and in general cell viability similar to the other ORC subunits (Li and Herskowitz, 1993). Orc6 is structurally related to the transcription factor TFIIB, which is composed of a tandem repeat of cyclin-box folds with the C-terminal end docked onto Orc3 (Chesnokov et al., 2003; Bleichert et al., 2015; Miller et al., 2019). The DNA-ORC interaction is mediated by the helicase-loader Cdc6. Like most ORC subunits, Cdc6 contains an AAA+ ATPase module and shows a highly similar structure as Orc1. Point mutations in the Walker A-motif of Cdc6, which is associated with ATP binding, lead to reduced or no function of Cdc6 *in vivo* (Elsasser et al., 1996; Perkins and Diffley, 1998; Weinreich et al., 1999). For later replication initiation, the loading of Mcm2-7 is required. Mcm2-7 shapes a discontinuous ring exhibiting a small gap between Mcm2 and Mcm5 assisting its loading around the DNA helix like an open bracelet. In yeast, the Cdt1 is stably connected with the Mcm2-7 in contrast to the higher eukaryotes. In particular Cdt1 is bound through its N- and C- terminal winged helix domains (WHD) to 3 MCM subunits (Mcm2/3 and 4) without hampering the entry to DNA. Older cryo-EM and biochemical data indicates the major role of Mcm3 and Mcm7 to the Mcm2-7 / Cdt1 loading to ORC / Cdc6 resulting in OCCM complex formation (**Figure 3**). Cdc6 is the first molecule dissociated from the OCCM followed by Cdt1. Both release steps require ATP hydrolysis. A third ATP hydrolysis step leads to ORC disassembly from the C-terminus of Mcm2-7 which is necessary for the loading of the second Mcm2-7 hexamer resulting in double-hexamer formation. Time-resolved EM analysis of MCM loading indicates that a second ORC recruitment is possible to occur before the complete eviction of the first one. The second ORC is bound to the N-terminus and not to the initial C-terminus of Mcm2-7 forming an inverted configuration, which helps the final double-hexamer formation (MO, **Figure 3**) (Miller et al., 2019). The ORC binding assists the loading of a second Cdc6 protein to ORC and thus a second Mcm2-7 hexamer together with Cdt1. Similar to the first ORC, the second ORC is released through the Mcm2-5 gate and the double Mcm2-7 hexamer is completed. Electron microscopy studies of purified *S. pombe* MCM complex confirm its doughnutlike structure with a central cavity (Adachi et al., 1997), which is similar to MCM purified from human cells (Sato et al., 2000). MCM proteins do not only participate in the pre-RC formation and DNA replication initiation, but serve an essential function as the replicative helicase to unwind the two daughter strands during DNA

replication elongation (see below). Thus, it has been found that inactivation of any of the six MCM2-7 subunits prevents progression of replication forks in *S. cerevisiae* (Tye et al., 1997).

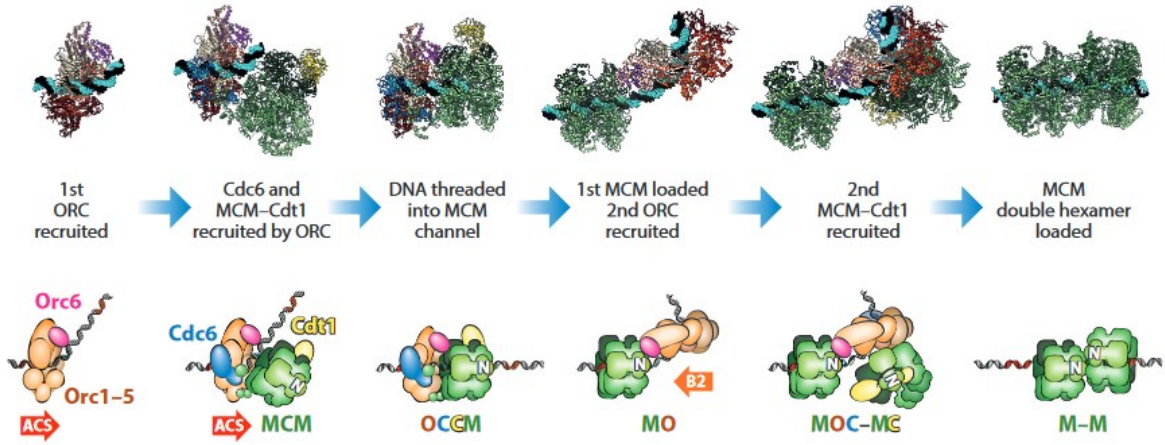


Figure 3. Structural analysis and schematic representation of the six steps required for a sequential and symmetric loading of MCM2-7 double hexamers and pre-RC formation (Adapted from (Costa and Diffley, 2022).

2.2 Origin firing

In the beginning of S-phase, the double hexamer (DH) undergoes a series of structural changes in order to activate the helicase. This includes several yet not completely understood conformational changes so that each MCM ring encircles one single DNA strand, thereby unwinding the DNA. Origin firing also includes the stable association of two additional protein complexes, Cdc45 and the GINS complex (go-ichi-nii-san, or 5-1-2-3 in Japanese), in yeast consisting of Sld5 (synthetic lethal with Dpb11), Psf1 (partner of Sld5 1), Psf2, and Psf3 (**Figure 4**) (Takayama et al., 2003; Moyer et al., 2006; Aparicio, 2013). The resulting Cdc45-MCM-GINS (CMG) complex forms the active replicative helicase that uses ATP hydrolysis to translocate and displace the complementary strand as the complex proceeds along the DNA strand (Gambus et al., 2006; Moyer et al., 2006). In particular, according to older studies origin firing is initiated by the loading and binding of Sld3 - Sld7 and Cdc45 (**Figure 4**). This recruitment is regulated by the pre-RC phosphorylation from DDK (Boos and Ferreira, 2019). In addition, CDK phosphorylates Sld3 and Sld2 and then both of them bind to distinct sequence-specific domains of Dpb11 forming the SDS complex (Sld3/7-Dpb11-Sld2) (Tanaka et al., 2007, 200). The phosphorylation-dependent formation of this complex is temporary, despite its crucial role in facilitating the loading

of Cdc45, GINS and DNA polymerase. Subsequently, Cdc45, Mcm2-7 and GINS form the stable CMG complex. CMG together with Mcm10 and DNA polymerase ε shape the active replicative helicase and initiate bi-directional DNA synthesis, while the SDS complex has dissociated from the origin (**Figure 4**) (Kunkel and Burgers, 2008; Boos and Ferreira, 2019).

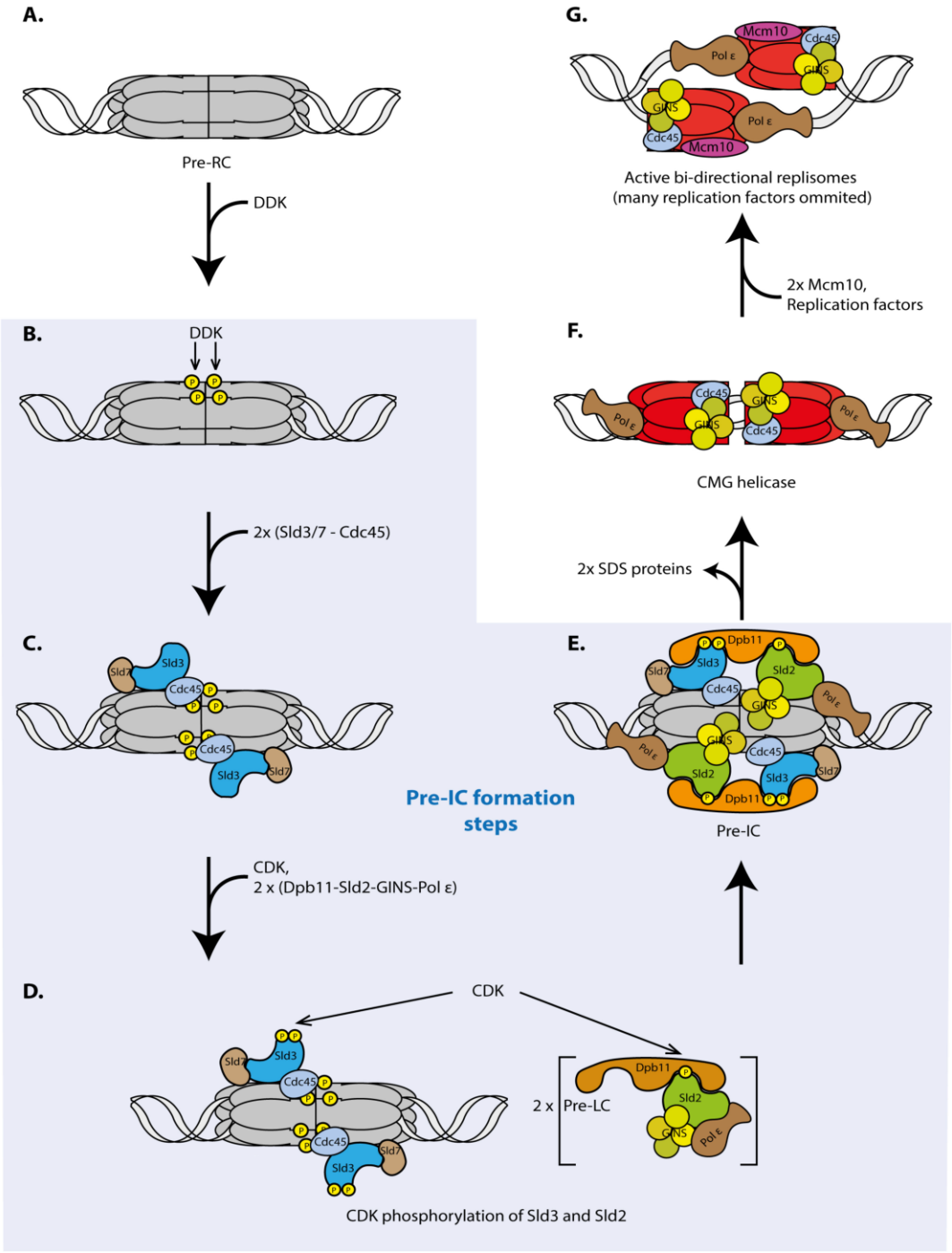


Figure 4. Schematic representation of origin firing process adapted from (Boos and Ferreira, 2019)

2.3 Cell Cycle Regulation of DNA replication

The two phases of origin licensing and origin firing are regulated during the cell cycle by the concerted action of two enzyme complexes: the Cyclin-Dependent Kinases (CDKs) and the Anaphase-Promoting Complex/Cyclosome (APC/C). APC/C is a multi-subunit E3 ubiquitin ligase containing adaptors such as Cdh1 and Cdc20, which polyubiquitinate other proteins resulting in their degradation by the proteasome complex (Peters, 2006). In general, three types of cyclin-CDKs can be distinguished:

1. G1-cyclin-CDKs which promote the entry to S-phase
2. S-cyclin-CDKs, which promote DNA replication in S-phase
3. M-cyclin-CDKs, which promote chromosome segregation though mitosis (Coudreuse and Nurse, 2010)

DNA replication is regulated during the cell cycle by biphasic oscillation of the CDK and APC/C enzymes that divide the cell cycle into two major states: After mitosis until the end of G1-phase, APC/C activity is high and degrades its targets including the S-CDK and M-CDK complexes. In contrast, G1-CDKs are resistant to this degradation, as they consist targets of a different E3 ligase named as Skp1–Cul1–F-box protein (SCF) complex (Ang and Wade Harper, 2005). Growth factors stimulate the expression and accumulation of G1-CDKs leading to APC/C-Cdh1 phosphorylation and hampering its activity. Subsequently, the second state extends from S-phase through mitosis and is characterized by low APC/C activity and high S-CDK and M-CDK activity (Jaspersen et al., 1999). MCM2-7 DH loading can only occur during the low S-CDK period in G1-phase, whereas helicase activation can occur only during the high S-CDK period. Thus, large numbers of origins can have helicase loaded before entry into S-phase but new helicases cannot be loaded during S-phase, thus ensuring origins can only fire once in a cell cycle. As a consequence, there is no urgency for all origins to fire immediately, and indeed, origins are activated throughout S-phase in a deterministic fashion resulting in a temporal program of origin firing (Blow et al., 2011).

2.2 Replication origins in yeast

As this thesis is focused on the characterization of replication origins in the budding yeast *Saccharomyces cerevisiae*, the following chapter will introduce and summarize our current

knowledge of replication origins, which are best understood in this unicellular eukaryotic organism.

2.2.1 Conserved ACS sites

Yeast replication origins are called autonomous replication sequences (ARS), since they confer the ability of autonomous replication on a circular plasmid molecule (Stinchcomb et al., 1981). Each ARS site contain an AT-rich 11-bp motif (T/A)TTTAT(A/G)TTT(T/A) which is known as A element and named ARS consensus sequence (ACS) (**Figure 5**). ACSs are necessary for the specific binding of the ATP-dependent origin recognition complex (ORC) (Bell and Stillman, 1992; Theis and Newlon, 1997; Li et al., 2018) but not sufficient for the function of the origin (Bolon and Bielinsky, 2006). Apart from the A element, the B1 element of ARS sites is partially involved in ORC-DNA interaction (Duderstadt and Berger, 2008; Li et al., 2018). In addition, the ACS and B1 element were recently shown to function as a binding site for an ORC-Cdc6 complex that facilitates the recruitment of the MCM2-7 helicase (Feng et al., 2021; Schmidt et al., 2022). The B2 element contains a common sequence of ANWWAAAN among 74% of annotated origins and is also suggested to support the binding of the MCM2-7 complex (Wilmes and Bell, 2002) (**Figure 5**). The B3 element is a transcription factor binding site for Abf1, which helps to position nucleosomes around a certain subset of origins (Miyake et al., 2002; Ganapathi et al., 2011). Interestingly, there is significant variability in the presence and abundance of these DNA elements and not every origin utilizes the complete set of B elements, suggesting that their modular presence at origins is auxiliary and promotes origin function, but is not an essential feature of an active chromosomal replication origin.

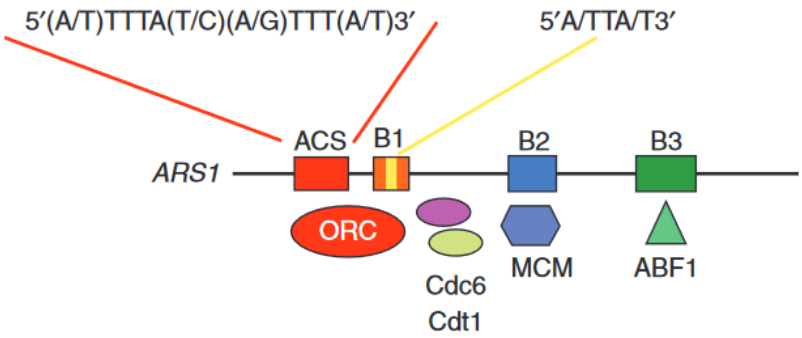


Figure 5. Schematic representation of ARS1 replication origin in *S. cerevisiae*. It is shown the regulatory DNA sequence elements existing on ARS1 locus together with their consensus sequences and the proteins binding to them (Leonard and Michaeli 2013).

Interestingly, the *Saccharomyces cerevisiae* genome contains more than 12.000 ACS sites but only 300 - 400 have been experimentally confirmed as replication origins and binding sites of ORC (Lipford and Bell, 2001), indicating that the ACS consensus motif alone cannot be used for the prediction of origin locations (Nieduszynski et al., 2006). However, different bioinformatic algorithms have been developed to attempt the prediction of replication origin locations *in vivo* based on the ACS motif. The Oriscan algorithm uses 268 bp sequence bins including the T-rich ACS and a 3′ A-rich region and has thereby recognized a large portion of the origins and ranked potential additional origin candidates by their likelihood of activity (Breier et al., 2004). Moreover, another computational study managed to identify the origins responsible for DNA replication in eukaryotic genomes based on phylogenetic conservation and microarray data (Nieduszynski et al., 2006). Additional bioinformatic analysis revealed that, despite the previous belief that replication origins in yeasts are AT-rich - while in plants and metazoan GC-rich, an industrially important methylotrophic budding yeast, *P.pastoris*, contained GC-rich ACS motifs. However, the phylogenetic tree based on the cytochrome c of *S. cerevisiae*, *K. lactis*, and *S. pombe* (AT-rich motifs) and *P.pastoris* (GC-rich) showed that the conserved motifs, found in the four yeast ARSs and analyzed by MEME-ChIP (Bailey et al., 2009; Machanick and Bailey, 2011), do not correlate significantly with their phylogenetic relationships (Peng et al., 2015). MEME-ChIP is a web tool which can be used for detection and motif analyses with meaningful biological function in DNA and RNA datasets (Ma et al., 2014).

Although it is clear that *in vivo*, budding yeast has origins with specific DNA sequence elements, and the ORC complex exhibits specific DNA binding activity (Bell and Stillman, 1992a), the ORC complex *in vitro* shows nonspecific DNA binding activity (Remus et al., 2009). Independently of the presence of a functional replication origin, this nonspecific ORC-DNA interaction can be functional leading to the loading of the Mcm2-7 hexameric ring (Remus et al., 2009). An additional study showed that for naked DNA replication *in vitro*, a functional replication origin is not a necessity. Yet, in the same system, a decrease in ORC concentration and the addition of competitor DNA was seen to enhance sequence-specific ORC binding, resulting in replication that showed origin dependency (On 2014).

2.2.2 Timing and efficiency

Despite that replication origins in yeast contain a conserved sequence, they present striking differences regarding their actual replication profile. Origins can initiate replication in a timely continuum from early to late stage during the S-phase (Reynolds and McCARROLL, 1989). In addition, replication origins can fire with different efficiencies within a cell population (Friedman et al., 1997). However, the probability of firing should not be directly interpreted as replication efficiency, as the firing efficiency of an origin is usually determined by two factors: the probability of origin firing and the proximity of origins with higher or lower firing probabilities. For example, an inactive or less active origin is frequently replicated in a passive manner by the earlier firing of a neighboring active replication origin (de Moura et al., 2010). As in metazoans, the late-replicating regions in yeast are telomeric regions as well as the small portion of heterochromatic areas, such as mating type regions (Grewal and Klar, 1996). Additionally, complex effects are revealed after interrupting the activity of specific chromatin modification enzymes. For example, a shift towards later replication initiation in the early replicating origins has been observed after the deletion of telomeric proteins Tac1 and Rif1 (Hayano et al., 2012; Tazumi et al., 2012). In addition, deletion of Rif1 in budding yeast shows earlier replication in telomeric region (Lian et al., 2011), whereas the disruption of histone acetyltransferase Rpd3L advances the firing time of approximately 100 origins (Knott et al., 2009). Given that it has been found that the global chromosome organization in G1 phase supports the enrichment of late replication origins in the nuclear periphery indicating the crucial role of G1-phase at replication timing (Heun et al., 2001), it was necessary for the levels of loaded MCM2-7 complexes, which initiates replication, on early and late replication origins to be analyzed. A first indication came from *in vitro* data from yeast nuclear extracts that showed that multiple MCMs can be loaded at individual ARS (Bowers et al., 2004), which was also confirmed by *in vivo* studies that showed that there are many more MCMs loaded on chromatin than the predicted number of origins necessary to complete DNA replication in S-phase (Donovan et al., 1997; Woodward et al., 2006). Based on these results, a mathematical model of replication kinetics analysis of budding yeast proposed that the number of loaded MCM complexes during the G1 phase in replication origins can determine the replication timing in S phase. According to this model, early origins have more MCMs and therefore show higher probability to initiate replication earlier (Yang et al., 2010). The assumption of this model was confirmed by several ChIP-seq results suggesting that MCM peaks are on average stronger at early origins than at late origins *in vivo* (Das et al., 2015a). Apart from that more recent study, measuring for the first time the MCM2-7 levels and the replication timing in the same experiment

proved that loading capacity of origins is the primary determinant of MCM stoichiometry in wild-type yeast cells, but the stoichiometry is influenced by the origins' ability to recruit ORC and compete for MCM2-7 loading when MCM2-7 complexes become limiting (Dukaj and Rhind, 2021a). In this study, a reduction of MCM2-7 complexes due to MCM4 degradation led to slower replication through S phase, increased sensitivity to replication stress and differential helicase loading in origins, overall revealing their heterogeneous sensitivity to limiting amounts of MCM2-7 complexes. Interestingly, despite the fact that MCM2-7 reduction results in slower replication, overexpression of MCM2-7 complexes has no effect on cell cycle progression and DNA replication rate, indicating that under normal conditions the loading of MCM2-7 complexes reaches a saturated level for efficient DNA replication (Dukaj and Rhind, 2021a).

2.2.3 Nucleosome positioning and chromatin environment

As mentioned, ARS sites in yeast are AT-rich regions and therefore not conductive to nucleosome assembly. Thus, a nucleosome-free region (NFR) is preferentially formed, thereby providing an ideal accessible environment for ORC binding and replication origin licensing (Lipford and Bell, 2001a; Eaton et al., 2010) (**Figure 6**). ChIP-seq data analysis identified 238 sequences in the yeast genome that intrinsically repel nucleosome formation, but lack ORC-binding and origin function despite their high compatibility. Functional origins with ORC-binding show an asymmetric position of the ORC complex in the NFR and flanking well-positioned nucleosomes. As mentioned, the additional B element sequences may maintain a larger open region to accommodate MCM2-7 loading and therefore the NFR is typically 90 base-pairs larger than the ORC footprint (Dorn and Cook, 2011; Eaton et al., 2010 n.d.). The ORC-binding is crucial for the nucleosome configuration around the origins and disruption of the ORC-directed nucleosome phasing interferes with the efficiency of origin firing (Lipford and Bell, 2001a). Using bulk assays correlating the replication timing profile with the local nucleosome occupancy, it has been found that early-origins tend to show higher nucleosome occupancy at the +1 and -1 position around the ARS sites compared to late origins (Soriano et al., 2014). Mutations in the ORC binding site can increase the nucleosome occupancy at ACS, which subsequently reduce origin function (Simpson, 1990). Additionally, mutations in the ORC binding site in both ARS1 and ARS307 allow nucleosomes to encroach into the origin, indicating a role for ORC in maintaining a NFR at origins (Lipford and Bell, 2001a). Interestingly, even though that disruption of nucleosome positioning adjacent to ORC binding sites can negatively affect the function of ARS1, the ORC binding

efficiency remains unaffected (Lipford and Bell, 2001a). These findings suggest both positive and negative roles of nucleosome occupancy in the origin replication profile.

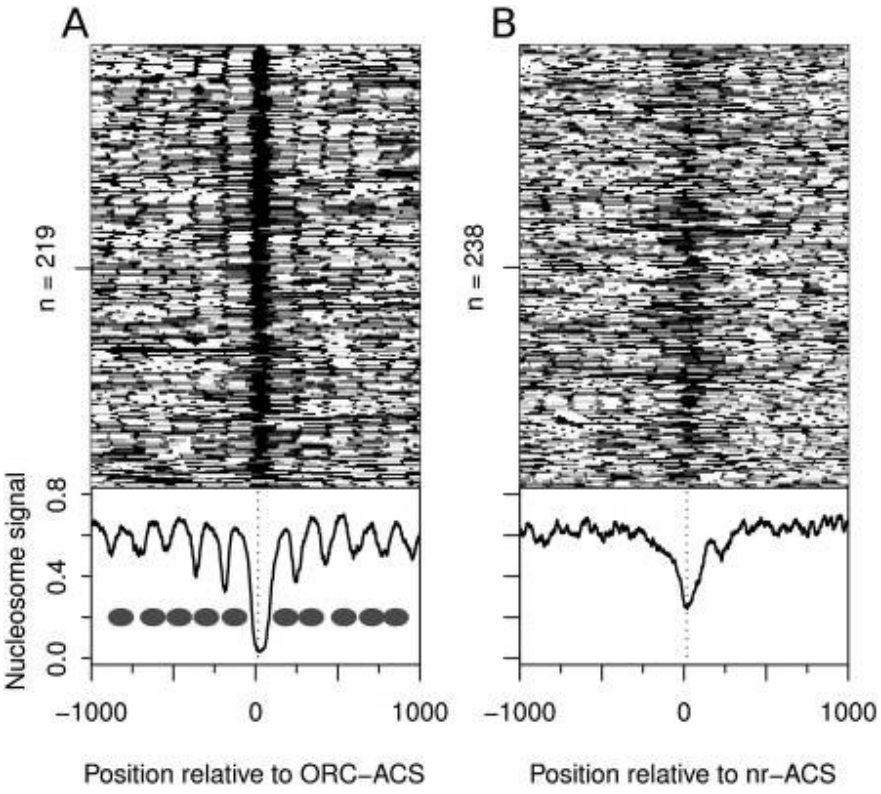


Figure 6. A. Heatmap of nucleosomal positions of 219 sites where ORC is associated with ACS sites. **B.** Heatmap of nucleosomal positions of 238 ACS sites without ORC association. The grey ovals on the left heatmap represent nucleosomes. Adapted from Eaton et al., (2010) .

For this reason, the further investigation of the local chromatin environment around DNA replication origins was necessary in order to uncover the detailed mechanism of replication timing and efficiency. Genome-wide nucleosome occupancy maps showed surprising diversity at replication origins (Berbenetz et al., 2010), emphasizing that the average and over-simplified view of population-averaged nucleosomal profiles does not reflect all different states that may co-exist at active or inactive replication origins. Neighboring genomic features, most notably TSS elements and gene ends, could also influence the nucleosome positioning and thus the replication identity of an origin. For example, it has been found that the presence of a TSS is possible to result in a second NFR in ACS, thereby improving the accessibility of ARS for ORC and increase the likelihood that these origins will be active in early S phase (Berbenetz et al., 2010). Late replication

origins were mostly correlated to either extremely narrow or wide NFR regions, but in all cases, ORC functioned as barrier element to position and phase nucleosomes on both sides of the origin (Berbenetz et al., 2010).

Together with nucleosomes, posttranslational modifications of histones and transcription factors can change the activity of an origin. In budding yeast, the global increase of histone acetylation in rpd3Δ cells shifts origins to fire earlier (Aparicio et al., 2004, 200; Knott et al., 2009), whereas the enrichment of H3K36me1 in ARS is correlated to Cdc45 loading to origins (Pryde et al., 2009). Additionally, it has been found that Forkhead transcription factor-dependent clustering of replication origins is crucial for early firing of some origins, suggesting that the replication profile and the activation of replication origins is correlated with their spatial organization (Knott et al., 2009b, 2012). Direct interaction studies of Dbf4 and forkhead transcription factors, Fkh1 and Fkh2 has revealed that Fkh1/Fkh2 loading determines the preferential binding of Dbf4 to early replication origins. However, its role can be overridden by fusing the DNA-binding domain of Fkh to Dbf4 forcing their interaction. Additionally, the association of origins with Dbf4 results in direct interaction with Sld3 and thus boosts the loading of Sld3–Cdc45, characterized as limiting replication factors and supports early replication of non-centromeric regions (Fang et al., 2017; Lynch et al., 2019). Therefore, it is clear that the replication profile of one origin could be controlled by an interplay of multiple chromatin and transcription factors. In this context, two distinct classes of origins have been defined: some origins firing very early during the S phase in their native sites are able to retain their early activation pattern even if they are placed in a different chromatin context (chromatin-independent). On the other hand, the activation time of other origins can change accordingly to their local chromatin environment (chromatin-dependent). It has been found that the early chromatin-independent origins contain two binding-sites for Forkhead transcription factors and can transform a late-replicating origin to an early one if they are inserted next to each other. However, only the insertion of two Forkhead-binding sites at chromatin-dependent origins is not sufficient for conversion (Lõoke et al., 2013).

2.2.4 Conclusion

Overall, important steps in our understanding of the molecular mechanism behind the DNA replication initiation have been made. However, most assays used so far represent bulk analysis and therefore fail to detect rare events that may occur in a cell population and result in limited or

biased conclusions. Thus, the usage of single-molecule methods with higher resolution is critical (Chanou and Hamperl, 2021a). Single-molecule analysis showed that in individual cells, the profile of origins firing is heterogeneous and there are no two cells with exactly identical replication profiles, indicating that the origin firing is rather a stochastic event (Czajkowsky et al., 2008; Rhind and Gilbert, 2013). This makes it difficult to correlate the specific chromatin structure with a specific replication behavior over the large amount of contaminating chromatin with inactive origins not selected by the replication machinery. Therefore, single molecule analysis of specifically selected origins with distinct properties regarding the timing and efficiency is necessary for a more complete understanding of this complicated molecular mechanism.

3. Chromatin remodelers

Chromatin remodeling factors are critical components involved in chromatin regulation. Using ATP hydrolysis, they access the packaged DNA, alter the nucleosome landscape by ejection, sliding and reassembly and thereby participate in many biological processes such as DNA replication, repair, and transcription (Vincent et al., 2008a; Conaway and Conaway, 2009; Tyagi et al., 2016; Patel et al., 2019). Currently, four families of chromatin remodelers are distinguished (SWI/SNF, ISWI, CHD and INO80 family) that show large structural differences and specialize on different biological functions. Despite this variability, all four families share five similar characteristics:

1. Nucleosome affinity domain,
2. Histone modification recognition domain,
3. ATPase domain,
4. ATPase regulatory domain or proteins
5. Domains interacting with other chromatin remodelers or transcription factors (Clapier and Cairns, 2009a).

The ATPase domain is conserved in all of the four families and each individual family is conserved from yeast to human with some variation in their detailed protein/complex composition (Flaus, 2006).

The SWI/SNF (switching defective/sucrose nonfermenting) family was initially found in budding yeast and is composed of 8 to 14 subunits. It contains an active ATPase separated into two parts by a short insertion, an HSA (helicase-SANT and a C-terminal bromodomain (**Figure 7**). SWI/SNF

family is known for nucleosome sliding and ejection at various functional and regulatory sites (Tang et al., 2010).

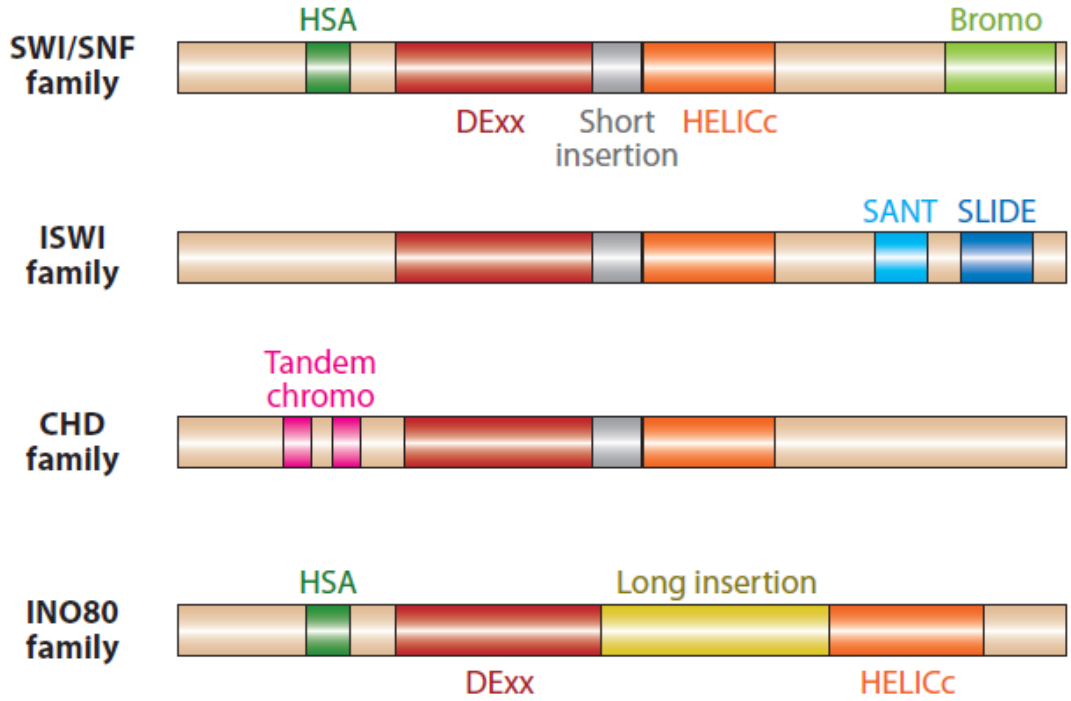


Figure 7. Organization of the four different chromatin remodeler families. Helicase-SANT (HSA – dark green) domain can be found on SWI/SNF and INO80 families, whereas the ISWI has a SANT SLIDE domain (blue) and the CHD family a Tandem chromo domain (pink). The SWI/SNF family has also a Bromo domain (light green). All families contain an ATPase subunit which is interrupted into two parts, the DExx (red) and the HELICc (orange) domains. Remodelers of INO80 family are the only ones having a long insertion between the two parts of ATPase. The members of the other three families show only a short insertion (adapted from (Clapier and Cairns, 2009b).

The CHD (chromodomain, helicase and DNA binding) family is known for facilitating transcription through sliding and ejection of nucleosomes. It contains one active ATPase separated by a short insertion and two N-terminal chromodomains (**Figure 7**) (Marfella and Imbalzano, 2007).

The ISWI (imitation switch) family consists of an active ATPase separated by a short insertion, one SANT and one SLIDE domain in the C-terminus (**Figure 7**). These two domains are responsible for the recognition of nucleosome core particles. SANT domain binds to the unmodified histone tail and DNA (Boyer et al., 2002, n.d.; Grüne et al., 2003) whereas SLIDE domain binds to nucleosome entry site and the linker DNA (Dang and Bartholomew, 2007). The

majority of ISWI members play important roles in nucleosome spacing during chromatin assembly and transcription repression.

In yeast, two ATPases, Isw1 and Isw2, have been identified as members of the ISWI family (**Figure 8**) (Tsukiyama et al., 1999). Isw1 can be found in two distinct complexes named Isw1a and Isw1b (Vary et al., 2003). Isw1a contains Ioc3, while Isw1b contains Ioc2 and Ioc4 proteins. In contrast to Ioc3, which contains no recognizable domains, Ioc2 has a PHD finger domain which is common between many chromatin-associated proteins. Ioc4 contains a PWWP motif, a binding domain for H3K36me3 (Vary et al., 2003).

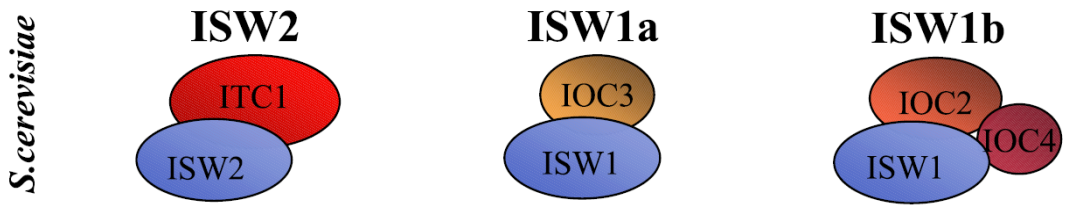


Figure 8. Members of the ISW family adapted from (Corona and Tamkun, 2004)

The INO80 (inositol requiring 80) family was initially purified in budding yeast and contains one N-terminal HSA domain and an active ATPase split by a long insertion where the Rvb1 and Arp (actin related proteins) can bind (**Figure 7**). The INO80 group includes the INO80 and SWR1 complexes.

Both SWR1 and INO80 remodelers also contain the Arp4-actin dimer and other actin-related proteins (Arp) (Gerhold and Gasser, 2014). INO80 complexes contain Arp5 and Arp8, as well as specialized subunits, like the Ino80 subunits 2 and 6 (Ies2 and Ies6) (**Figure 9**). The SWR1 complex promotes nucleosomal H2A exchange with the histone variant H2A.Z (Mizuguchi et al., 2004; Luk et al., 2010). In contrast, the INO80 complex has antagonistic functions and replaces the nucleosomal H2A.Z with H2A (Papamichos-Chronakis et al., 2012; Brahma et al., 2017).

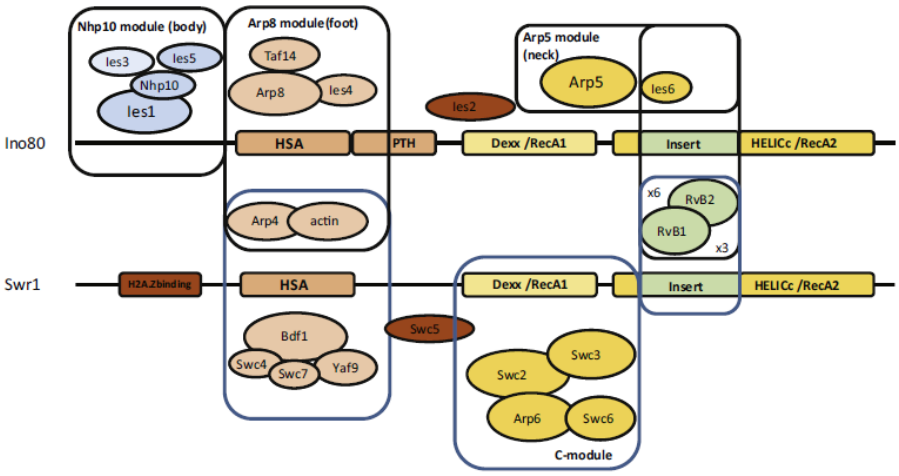


Figure 9. Structural similarities in the architecture of INO80 (upper) and SWR-C (lower)- both members of the INO80 family modified from (Gerhold and Gasser, 2014).

3.1 How does DNA translocation lead to nucleosome sliding or exchange of histone octamers?

The mechanism behind the nucleosome sliding and DNA translocation by ISWI has been broadly investigated in various organisms (Saha et al., 2002, 2006; Fitzgerald et al., 2004; Kagalwala et al., 2004; Strohner et al., 2005) suggesting a “loop recapture” model (**Figure 10**). According to this model, chromatin remodeler and ATPase are bound to the nucleosome. ATP hydrolysis is responsible for the DNA translocation and its binding site is located two turns from the dyad. As a next step, the ATPase, which is stably attached at a fixed position, draws DNA from the linker into the nucleosomal dyad via the DNA Binding Domain (DBD) (**Figure 10** step 1), thereby forming a DNA loop on the nucleosome. (**Figure 10 step 2**). Subsequently, the loop on the nucleosomal surface next to the dyad is spread to the second half of the nucleosome core particle (**Figure 10 step 3**), disrupting histone-DNA contacts at the leading edge of the loop and replacing them at the lagging edge of the loop and resulting in nucleosomal sliding (**Figure 10 step 4**). Large DNA loops have been successfully detected by single molecule methods (Zhang et al., 2005; Lia et al., 2006). In case that this loop is restricted to a position on the nucleosome surface where H2A-H2B reside, this can facilitate INO80-mediated replacement by an H2A.Z-H2B dimer (Brahma et al., 2017), which applies similar mechanism for nucleosome translocation (Kobayashi and Kurumizaka, 2019).

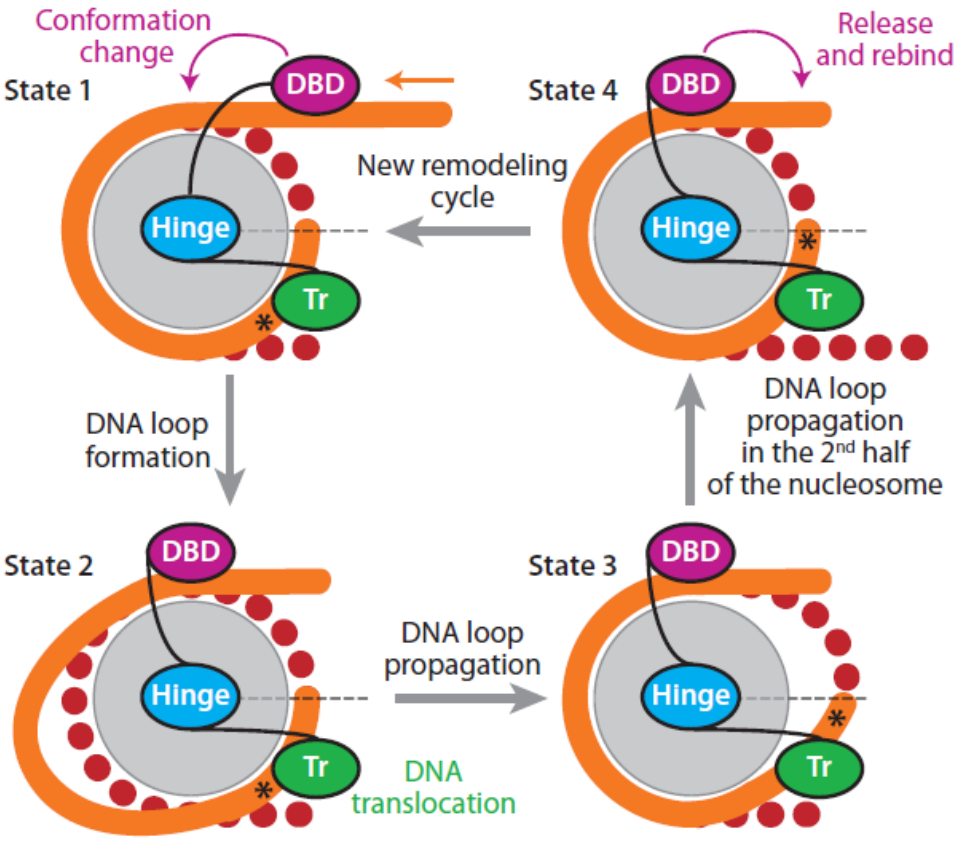


Figure 10. 1-4 steps describing nucleosomal sliding. The nucleosome is presented as a grey cycle wrapped by two turns of DNA (orange and red dots). The DNA Binding Domain (DBD) is shown as pink element on a fixed position on DNA helix and the ATPase/Translocase (Tr) as a green oval (modified from Clapier and Cairns, 2009).

3.2 Role of ISWI and INO80 in biological processes

3.2.1 Transcription and gene activation/repression

The biochemical characterization of ISWI and genetic studies in various organisms have revealed roles for this chromatin-remodeling factor in a wide variety of nuclear processes such as transcription, DNA replication and gene repression or activation. The non-catalytical subunits of the ISWI complex (Ioc2, Ioc3 and Ioc4) and a Sir2-interacting protein showing strong homology

to Ioc3, Esc8, influence silencing at telomeres and the HM mating type loci (Cuperus and Shore, 2002). Severe reduction in silencing at HMR has been observed in the absence of Isw1. Moreover, Isw2 complex inhibits transcription in a pathway that is parallel with the Sin3–Rpd3 complex. By moving nucleosomes into repressive positions, the Isw2 complex can establish nuclease-inaccessible chromatin structures at specific sites *in vivo* and repress early meiotic genes in parallel with the Sin3–Rpd3 histone deacetylase complex (Goldmark et al., 2000). These findings are consistent with a report showing that Isw2 mutants exhibit defects during the early stages of sporulation (Trachtulcov et al., 2000). In addition to meiotic target genes, other sets of genes are also derepressed in mutants lacking both Isw2 and the Sin3–Rpd3 complexes (Fazzio et al., 2001). Additional data showed the Isw1 may repress PHO8 and PHO84 gene expression in response to glucose deletion (Moreau et al., 2003). In rich medium (YPD), a short region spanning the MET16 promoter and the early part of the coding sequences, are enriched for Isw1 which is also responsible for +/- 1 nucleosome positioning in this locus (Morillon et al., 2005). Disruption of nucleosome positioning at +/- 1 sites in Isw1 ATPase defective strains has been correlated to repressed expression of MET16 gene. Upon activation of MET16 gene Isw1/Ioc3 is absent from the promoter region, whereas Isw1/ Ioc2 and Ioc4 can be detected in the coding region. These results suggest the diverse function of ISWI in gene expression.

INO80 chromatin remodeling has broad effects on promoters, facilitating both transcriptional activation and repression by modulating the position and composition of nucleosomes at promoters. An *in vitro* nucleosome reconstitution system proved that INO80 was the only remodeler able to recognize and establish NFRs by itself, and to position correctly the +/-1 nucleosomes relative to the transcription start sites (TSS). In addition, INO80 in cooperation with ISWIa was responsible to correctly space nucleosomes downstream of the TSS (Krietenstein et al., 2016). *In vivo* studies showed that generally at promoter regions INO80 does not evict nucleosomes upon activation but together with SWI/SNF family is responsible for nucleosome eviction at the PHO5 promoter during gene activation ensuring the accessibility of this region to transcription factors and supporting the complete activation of the PHO5 gene (Steger et al., 2003; Barbaric et al., 2007). Genome-wide studies indicated potential INO80 binding to over 90% of budding yeast's gene promoters (Yen et al., 2013). Misregulation of approximately 15% of yeast genes upon loss of either the catalytic Ino80 or the Arp5-Ies6 core subcomplex has been shown by RNA-seq. The effect of INO80 deletion on these genes was not uniform, instead similar numbers of genes were up- and downregulated, respectively (Yao et al., 2016). These broad effects on transcription could be correlated with the occupancy of Ino80, Arp5 and Ies6 at the +1

nucleosome of the TSS of the affected genes, suggesting strongly a direct role in transcriptional regulation. Related to its proximity to promoter regions, INO80 prevents bidirectional transcription at functional promoters (Marquardt et al., 2014), and loss of INO80 leads to increased transcription of noncoding RNAs (Alcid and Tsukiyama, 2014; Xue et al., 2015).

3.2.2 DNA replication

Despite that the role of chromatin remodelers on transcription has been investigated thoroughly, the available information regarding their function on DNA replication is limited. It has been found that upon stress conditions, DNA replication of late-replicating regions could be promoted by Isw2 and In80 reducing the checkpoint response during S-phase (Vincent et al., 2008a; Au et al., 2011), while the Isw2 alone plays crucial role at the DNA replication in pericentromeric heterochromatin regions in mammalian cells (Collins et al., 2002). On the other hand, yeast cells with disrupted function of INO80 chromatin remodelers reveal minimized replication fork progression and low ability of re-activation of stalled forks (Papamichos-Chronakis and Peterson, 2008; Shimada et al., 2008). It has been shown that presence of INO80 at replication origins supports genome stability by preventing transcription, thereby eliminating potential transcription – replication conflicts (TRCs) (Topal et al., 2020).

4. Single molecule techniques to study chromatin

To understand these complex processes in more mechanistic details, it is critical to obtain a comprehensive overview of the composition and nucleosome configuration in correlation with the functional state of a locus of interest. Various population-based techniques have been developed to investigate the architecture of chromatin. The genome can be surveyed for exposed regions accessible for nuclease (DNAse-seq, MNase-seq, FAIRE-seq) (Schones et al., 2008a; Bianco et al., 2015a; Ishii et al., 2015a) or transposase attack (ATAC-seq) to distinguish between open regulatory regions of chromatin vs. those protected by nucleosomes (Buenrostro et al., 2015).

Despite their usefulness and great insights into the chromatin structure and spatial organization of chromosomes, it is important to realize that all of these genomic methods describe the properties of enormous numbers of molecules, averaging the measured parameters over a large population of molecules and cells. Thus, the behavior of individual molecules with distinct conformations and properties cannot be observed over time. In addition, bulk methods fail to detect rare events that occur only in a small subpopulation of molecules. Therefore, single-molecule methods provide the only possible solution to detect the functional differences and uncover intermolecular heterogeneities, particularly important in the context of chromatin.

4.1.1 Electron microscopy after psoralen crosslinking

Electron microscopy was the first method used for nucleosome detection as beads on the string structure (Olins and Olins, 1974; Kornberg, n.d.). Two years later nucleosomal structures were treated with psoralen derivates and visualized under electron microscope (EM) (Hanson et al., 1976; Sogo et al., 1984). Psoralen has the ability to bind only to non-nucleosomal double strand DNA and thus unprotected, accessible DNA and create covalent crosslinks between pyrimidines upon ultraviolet irradiation at 366nm (Cimino et al., 1985). The crosslinked DNA has been visualized in native agarose gels exhibiting a typical shift towards bigger DNA sizes (Lucchini, 2001; Wellinger and Sogo, 1998, Comino et al., 1985) revealing the presence of two distinct chromatin states of ribosomal DNA (rDNA), one nucleosomal and fully protected population and one nucleosome-free accessible state (Wittner et al., 2011). Despite the fact that this population-based approach could successfully define major qualitative differences, it does not uncover individual nucleosome configurations among the whole cell population.

However, single molecule analysis on nucleosome positioning can be achieved by combining the psoralen-crosslinking with electron microscopy. After denaturation of DNA strands, individual DNA bubbles corresponding to nucleosomes appear along the DNA fragments connected by undisrupted linker DNA (Brown and Boeger, 2014a) (**Figure 11**). In this way, two clusters of rRNA genes were determined by the combination of psoralen crosslinking and electron microscopy technique (Dammann et al., 1993). More recently, this approach was used in yeast to correlate the local chromatin landscape of single gene molecules adjacent to PHO5 promoter in its active or inactive transcription state. Interestingly, high level of heterogeneity was observed, since the actively induced promoter is surrounded by every combinatorial possibility of occupying three promoter nucleosome positions (2^3 = 8). This way gene expression noise can be potentially derived from the observed heterogeneous chromatin structure of the gene promoter (Brown et al., 2013). Psoralen crosslinking revealed 12 heterogenous groups of molecules with distinct nucleosome positioning at 5S rDNA molecules (Hamperl et al., 2014a). Interestingly, the averaging number of molecules containing an open 5S locus matches with older studies and is correlated with its expected transcription status (Hamperl et al., 2014a). Even though that the frequent use of this approach is rather challenging because of the specialized equipment and the time-consuming experimental set-up, EM is one of the most pioneering tools in the field of single molecule nucleosome positioning analysis.

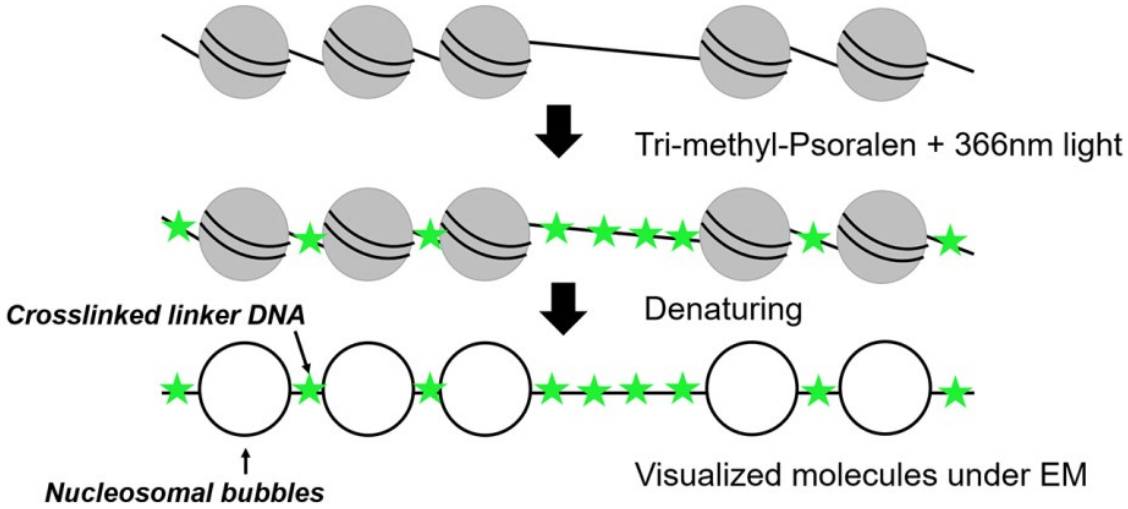


Figure 11. Cartoon describing the psoralen crosslinking process. The green stars represent the unprotected crosslinked DNA, whereas the white bubbles indicate previously occupied nucleosomal positions which are shown as single-stranded DNA after denaturation process. (Adapted from Chanou and Hamperl, (2021b)

4.1.2 Methylation footprinting

Traditional Sanger or short-read Illumina sequencing are also suited for chromatin accessibility studies. The chromatin is treated by DNA methyltransferases which detect and modify the unprotected CpG dinucleotides, whereas the sites covered by nucleosomes or transcription factors remained unmethylated (**Figure 12**) (Miranda et al., 2010). Next, the methylated cytosine bases are deaminated to uracil after bisulfite conversion. Each DNA fragment is amplified by PCR, cloned into plasmid vectors and transformed into *E. coli* for further amplification. The amplified DNA is sequenced and upon comparison with a reference genome, the chromatin accessibility landscape is revealed in a single molecule level (Miranda et al., 2010; Li and Tollefsbol, 2011).

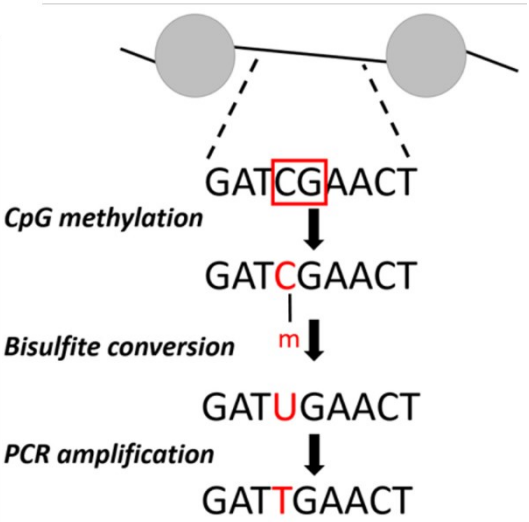


Figure 12. Schematic representation of bisulfite conversion. The red square shows non-nucleosomal CG sites recognized by CpG DNA methyltransferases. After C-U conversion, the uracil is converted into thymidine after PCR amplification. The nucleosomal regions will be distinguished from the non-nucleosomal by comparing the PCR products with the reference DNA sequence. Modified from Chanou and Hamperl, (2021)

In more recent years, two different third generation sequencing approaches have been developed, the Single Molecule Real-Time sequencing (SMRT-seq) and Nanopore sequencing (Clarke et al., 2009; Eid et al., 2009). Both methods are able to sequence native DNA avoiding the PCR amplification steps and thus minimizing potential errors during this processing step. In

this way, direct recognition of modified DNA bases is achieved via simpler and time-efficient experimental set-up. Importantly, these methods are the first providing long DNA reads up to 4.6 million base-pair (Zook et al., 2020), enabling the accurate assembly of repetitive elements and the monitoring of epigenetic modifications on multiple co-existing chromatin states (Jain et al., 2016; Ardui et al., 2018).

The SMRT-seq method emerged by PacBio uses fluorescently labeled nucleotides which are bound one by one to the tested complementary DNA strand by DNA polymerase. Polymerase kinetics and the direct recognition of modified nucleotides are being defined by the nucleotide's arrival times and the duration of fluorescent signal, since each modification has a unique kinetic profile (Flusberg et al., 2010) (**Figure 13**). In 2015 a method combining the bisulfite conversion with the SMRT-seq, named SMRT-BS, was reported including multiple optimization steps for CpG detection. In particular, regions up to 1.5Kb upon bisulfite conversion were sequenced by PacBio platforms resulting in higher accuracy (Yang et al., 2015).

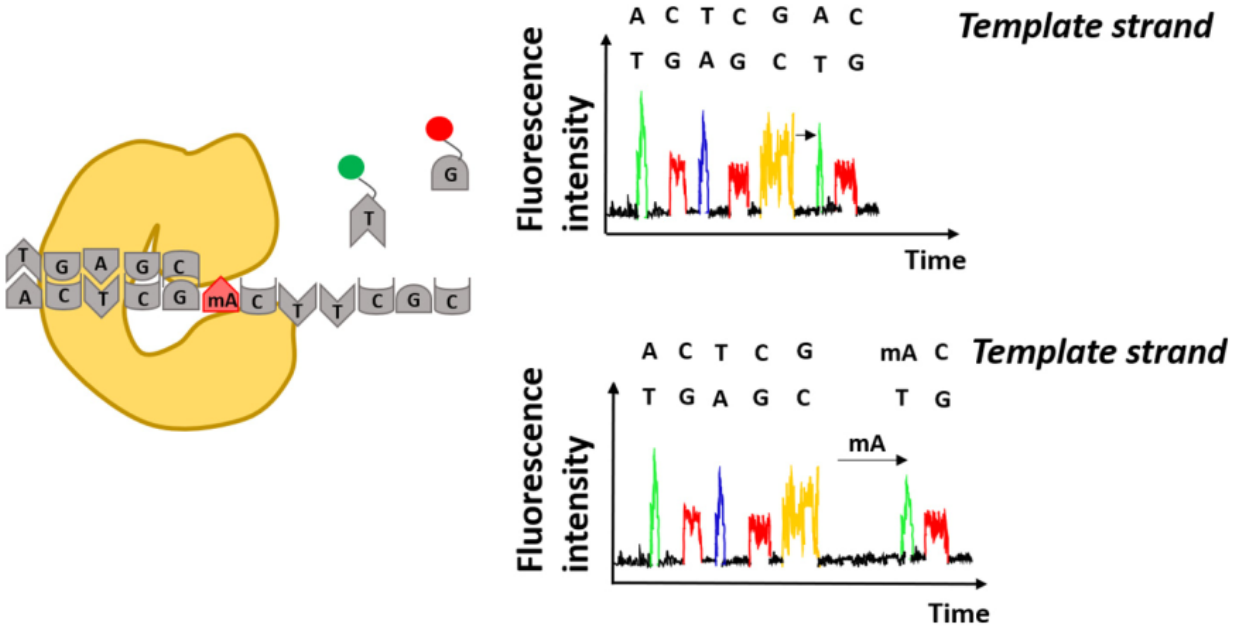


Figure 13. Single Molecule Real-Time sequencing for tracking modified adenine residues. Specific fluorescence underlies incorporation of modified nucleotides by DNA polymerase. Changes at the expected arrival times of a certain nucleotide (e.g. thymidine) suggest the presence of a modified nucleotide (e.g. adenine). Adapted from (Chanou and Hamperl, 2021)

Even though that adenine is frequently present along the DNA, the eukaryotic genome is lacking endogenous N6-Methyladenine (m6A). Using this characteristic as advantage, two recent methods called Fiber-seq and Single-Molecule-Adenine Methylated Oligonucleosome Sequencing Assay (SAMOSA) use DNA methyltransferases for m6A to the unprotected chromatin region or nucleosome positioning detection. This approach provided high coverage of methylation sites and increased resolution compared to CpG methylation (Stergachis et al., 2020; Abdulhay et al,. 2020).

Recently, the primary landscape of chromatin fibers at single nucleotide resolution was recreated by the Fiber-seq approach (Stergachis et al., 2020). After using as control *Drosophila melanogaster* S2 cells to confirm that protected nucleosomal DNA is unable to be methylated by DNA methyltransferases, the authors investigated whether the unique and most important factor for actuation of the regulatory elements is presence or absence of nucleosome core particles. The findings indicate that co-actuation is tightly connected to clustered elements, meaning that actuation of regulatory DNA accessibility at one distal element appears to potentiate accessibility at neighboring elements in a distance-dependent manner. Additionally, given that well-positioned nucleosomes were closer to regulatory elements being in an actuated state, it was shown that DNA sequence alone cannot determine the nucleosome positioning (Stergachis et al., 2020). Similar to Fiber-seq, *in vitro* assembled nucleosomal arrays were used to prove the ability of SAMOSA to detect presence or absence of nucleosomes. Interestingly, the results of SAMOSA suggested that euchromatin and heterochromatin regions, which are actively transcribed, show high level of heterogeneity, even though the latter one is typically considered as a more static epigenomic domain.

Independently from SMRT-seq, Oxford Nanopore Technologies (ONT) also offers innovative technology for detecting the chromatin accessibility landscape in a single molecule level. The idea of nanopore sequencing has emerged from the fact that recognition of individual DNA bases could be achieved by the distinct ionic current of DNA bases while passing through a biological pore (Kasianowicz et al., 1996). Methylated DNA bases also result in different signals compared to non-methylated bases, which assists the direct detection of free vs protected nucleosomal regions **(Figure 14).** A proof-of principle analysis was provided in 2019 for transcription start sites (TSSs) analysis (Wang et al., 2019a). The authors developed an approach called **Me**thyltransferase treatment at GpC sites (by M.CviPI) on haploid *Saccharomyces cerevisiae* followed by **S**ingle-**M**olecule **L**ong-**R**ead sequencing (MeSMLR-seq) and revealed the elevated heterogeneity of the silent genes compared to transcriptionally active genes. The long reads provided by MeSMLR-

seq enabled the detection of heterogeneity between two neighboring regions along the same molecule and therefore a much more detailed and quantitative analysis compared to population-based methods. In particular MeSMLR-seq uncovered three different profiles of chromatin accessibility on CLN2 promoters: closed, narrowly opened and widely opened and expanded our current view on the different chromatin states and their heterogeneity at this regulatory DNA element (Wang et al., 2019).

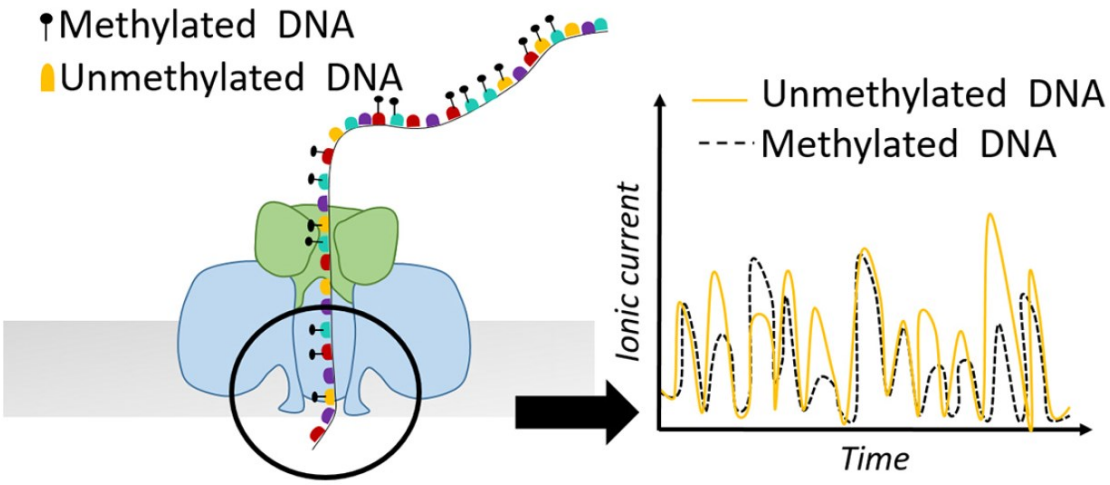


Figure 14. Cartoon presenting a nanopore detecting the different ionic current of an unmethylated DNA sequence compared to a methylated one. Adapted from (Chanou and Hamperl, 2021)

An additional approach using a similar concept is the **S**ingle-**M**olecule long-read **A**ccessible **C**hromatin mapping sequencing (SMAC-seq) (Shipony et al., 2020). The authors advanced the previous protocols by using a combination of three methyltransferases M.CviPI (CpG), M.SssI (GpC) and EcoGII (m6A) to map the nucleosomal pattern on yeast. For the analysis, three different bioinformatic algorithms were used: Albacore for raw base calling, Tombo for m6A and Nanopolish for CpG and GpC detection. The results showed well-positioned nucleosomes on centrometric DNA regions on yeast and the co-existence of two chromatin states on rDNA. In particular, the actively transcribed 35S area has fully accessible molecules in contrast to the inactive non-transcribed intergenic sequence (Shipony et al., 2020).

4.2 Single molecule approaches to monitor DNA replication

4.2.1 DNA Fiber and Molecular Combing for mapping the DNA replication progress

One of the most typical methods used for DNA replication analysis is based on detection of newly synthesized DNA by flow cytometry or immunofluorescence (Pozarowski and Darzynkiewicz, 2004; Harris et al., 2018) after its labelling with thymidine analogues like BrdU (5-bromo-2'-deoxyuridine) or EdU (5-ethynyl-2'-deoxyuridine). Despite their broad usage, these techniques lack single molecule resolution and therefore the ability to uncover heterogeneity among multiple replication forks in a cell population.

Molecular combing and DNA spreading/fiber analyses can monitor the dynamics of DNA replication on a single-molecule level (Bensimon et al., 1994; Michalet, 1997; Tuduri et al., 2010). Similar to the previous bulk methods, molecular combing and DNA spreading/fiber analysis use thymidine analogues iododeoxyuridine IdU and chlorodeoxyuridine CldU incorporating into the newly synthesized DNA (Técher et al., 2013). Subsequent antibody staining against these halogenated nucleotides enables the visualization of the DNA fiber under the fluorescence microscope as tracts of red and green labeled DNA (Nieminuszczy et al., 2016) (**Figure 15**). Even up to 12Mb of individual DNA molecules can be analyzed by molecular combing (Kaykov et al., 2016). The advantages of using two subsequent pulses of labeled nucleotides are expanding further from the simple detection of active replication forks at a specific time-point (Czajkowsky et al., 2008) to the analysis of replisome speed, and frequency of DNA replication initiation and termination events (Bialic et al., 2015; Nieminuszczy et al., 2016; Ray Chaudhuri et al., 2016; Vujanovic et al., 2017).

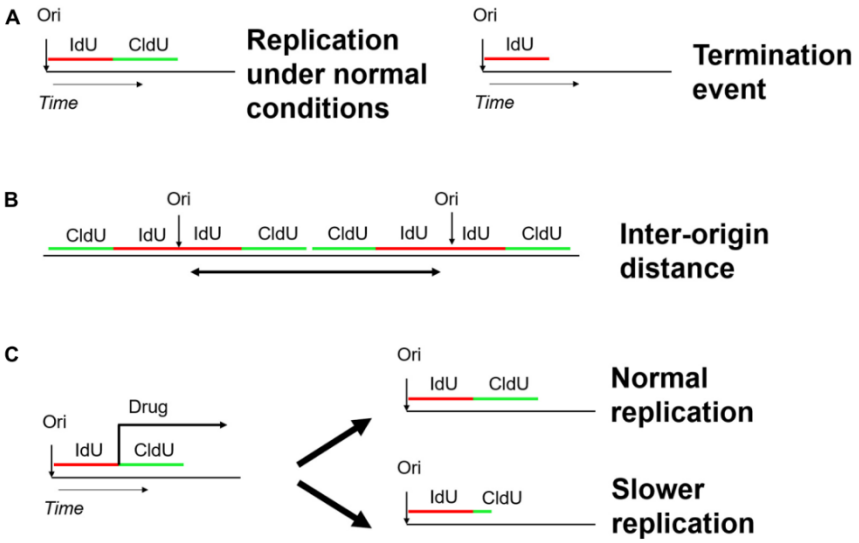


Figure 15. Cartoon presenting the DNA fiber technique to track different replication events (initiation / elongation/ termination (**A**), distance between origin (**B**), DNA replication elongation under stress conditions (**C**). Replicating DNA is labelled by Idu and CIdU subsequently (adapted from (Chanou and Hamperl, 2021).

In different species, e.g. chicken DT40 cells, *Xenopus*, mouse and human cells, the progression of DNA replication forks has been measured by the DNA fiber approach (Blow et al., 2001; Schwab et al., 2010; Guilbaud et al., 2011). Additionally, it has shed light to major interaction of crucial replication initiation proteins such as MCM10 at DNA replication origins. When combined with known replication inhibitors (hydroxyurea or aphidicolin), this method has revealed that the human RIF1 and Protein Phosphatase 1 can eliminate the DNA breaks after replication stalling (Garzón et al., 2019). Additionally, DNA fiber analysis has shown that under stress conditions the overall speed of replication forks is reduced, while the frequency of replication stalling - defined by calculating the ratio of two sister forks - increased. The general slower rate is depicted by ratio = 1, since the two sisters have the same length, whereas the stalling leads to asymmetric progression and thus ratio <1 (**Figure 15**).

To overcome the challenging low throughput of DNA fiber technique, DNA stretching has been also achieved by nanochannels in order to monitor the DNA replication in *Xenopus laevis* egg extracts and mammalian systems (Lacroix et al., 2016; De Carli et al., 2018; Wang et al., 2020). In particular, in *Xenopus laevis* egg extracts DNA replication was visualized under an optical mapping device and bacteriophage λ DNA replicated inside the egg extracts was used as proof of principle. The bacteriophage λ was stained with fluorescent deoxyuridine triphosphate (dUTP), whereas the DNA was stained at distinct restriction sites by a fluorescently labelled nicking endonuclease. Due to the higher throughput, this technique let the tested DNA being aligned with the reference genome with higher resolution and allowed robust statistical analysis of the replication profile (De Carli et al., 2018).

4.2.2 Psoralen crosslinking -EM to visualize replication forks

Given that psoralen crosslinking in combination with EM visualization provides unprecedented resolution, direct single molecule detection has also been used to monitor DNA replication forks under normal or stressed conditions (Vindigni and Lopes, 2017; Zellweger and Lopes, 2018). It has been observed that single DNA stranded gaps and formations of Holiday junction could be derived by the absence of RAD53 kinase (Sogo, 2002), while the absence of DNA2 and RECQ1

could result in accumulated fork reversal events suggesting that undisrupted function of these factor is necessary to DNA replication progression (Thangavel et al., 2015).

4.2.3 Fluorescent and force manipulation methods for DNA replication monitoring

Force manipulation approaches like magnetic or optical tweezers have also been used for single molecule analysis of DNA replication. In particular, during the usual magnetic tweezer set-up, one end of single-strand DNA is stably attached on a glass surface with a DNA primer supporting the loading of DNA polymerase and the other end is attached to a magnetic bead through streptavidin. Using this set-up, the role and efficiency of two different DNA polymerase (Sequenase VS Klenow) has been determined confirming that pausing events are possible to occur with both enzymes and there is no dependency of sequence specificity but rather on the applied stretching force to the DNA template (Maier et al., 2000). Replication rate is eliminated by forces stronger than 4pN and stops upon forces higher than 20pN. Similarly, this higher probability of pausing events has been observed after maximum tension of the DNA template using optical tweezers (Wuite et al., 2000).

4.2.4. Third generation sequencing to detect DNA replication

The precise mapping of early replication origins during the S-phase has been achieved by the development of a new approach combining the nanopore sequencing technology with the thymidine analogue BrdU (Hennion et al., 2018). For this analysis initially a basecalling algorithmus was generated to detect the BrdU-shift in the electric current using an *in vitro* system and the pipeline was trained on a genetically modified yeast strain dependent on exogenous thymidine or BrdU for its replication (Hennion et al., 2018). An improved version of this method named D-Nascent was published one year later and provided extremely long nanopore reads in combination with a Hidden Markov Model for tracking the BrdU incorporation (Müller et al., 2019).

5. Aim of study

Despite the fact that the position of DNA replication origins in yeast is well defined and their DNA sequence has been analyzed, the main factors regulating efficient DNA replication remain unknown. The aim of this study is to investigate the role of the local nucleosome positioning at ARS on DNA replication efficiency on wildtype and chromatin remodeling enzymes mutant strains and most importantly to provide a comprehensive analysis of the heterogenous chromatin landscapes existing among a cell population. In the CRE mutant strains deletion of ISW2 and IES6 subunits of ISW2 and INO80 complexes respectively has been performed. To this end, we developed a new tool based on single-molecule Nanopore sequencing named Methylation Accessibility of a Targeted Chromatin locus (MATAC-seq) sequencing approach and applied it to specific targeted DNA replication origins with different replication profiles. Applying this tool to wildtype and mutant yeast strains defective in the chromatin remodeling complexes ISW2 and INO80 showing changes of the replication profile and nucleosome positioning at the selected origins, we aim to understand the chromatin-based rules of origin selection and to further our understanding of the molecular basis of the replication timing program in eukaryotic cells, deregulated in a myriad of human disease states.

6. Results

6.1. Disruption of the function of major chromatin remodelers

The goal of this work was to test the hypothesis that the local chromatin structure, and especially the positioning of nucleosomes, is a predictive feature for efficient or inefficient firing of specific DNA replication origins. To this end, we disrupted the function of two major chromatin remodeling enzymes (CRE), ISW2 and INO80, by deletion of the *ISW2* or *IES6* genes, respectively. Importantly, previous studies showed that knockout of these subunits in yeast completely abolishes the enzymatic function of these complexes (Vincent et al., 2008) The genes were replaced by homologous recombination with auxotrophic markers and individual positive clones were selected and verified by genotyping PCR and Sanger sequencing (data not shown).

6.1.2. Deletion of ISW2 or IES6 does not affect generation time and cell cycle progression

As the investigated CRE have multiple functions in the cell including transcription and DNA replication (see also Introduction), I first wanted to exclude the possibility that any differences observed between wildtype (Control) strain and the obtained *ISW2* or *IES6* mutant strains are caused by a global defect regarding their growth rate or cell cycle kinetics.

To this end, the wildtype (Control), *isw2*Δ and *ies6*Δ yeast strains were subjected to growth analyses in liquid cultures by determining the change in optical density (OD600) over time (**Figure 16**). In the logarithmic growth phase, the CRE mutants do not display any statistically significant growth defect compared to the control (**Figure 16**).

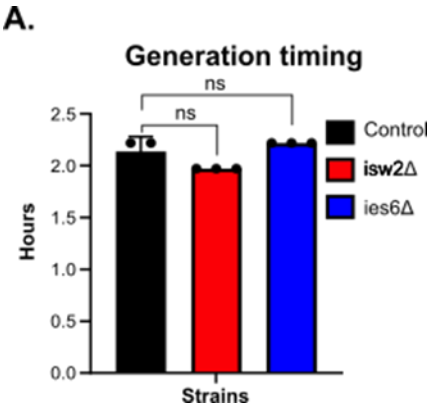


Figure 16. Generation timing of Y66 (control), Y104 (*isw2*Δ) and Y136 (*ies6*Δ). The strains were grown in YPD medium and OD600 measurements were taken every 30 sec by a multi-plate reader.

In order to profile global cell cycle kinetics, wildtype (Control), *isw2*Δ and *ies6*Δ cells were synchronized in G1 by addition of alpha factor and subsequently released into S-phase by addition of Pronase enzyme degrading the alpha factor. Samples during this arrest and release experiment were taken at the indicated timepoints (**Figure 17**) and the G1, S or G2 cell cycle populations quantified by flow cytometry. In this analysis, all strains showed comparable arrest kinetics and the large majority of cells was arrested in G1 phase after 120 min. Importantly, all strains initiate replication and move into S-phase at ~16 min after addition of Pronease and progress through the cell cycle with highly similar kinetics (**Figure 17**).

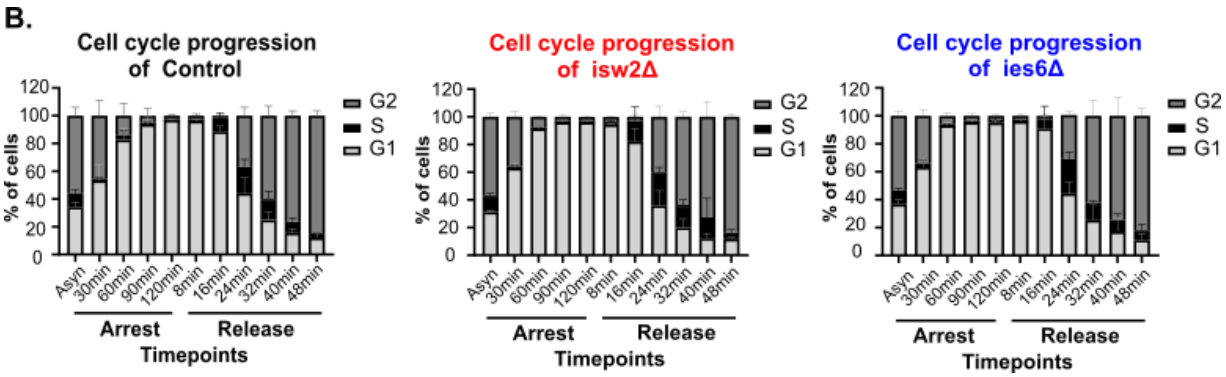


Figure 17. Cell cycle progression of control Y66 (control) and mutant strains Y104 (*isw2*Δ) and Y136 (*ies6*Δ). The cells were arrested in G1-phase by addition of 50 mg/ml α-factor and released into replication by the addition of 125 U Pronase. Samples for flow cytometry were taken at the indicated timepoints and the relative population of cells in G1, S or G2 were quantified.

In summary, all strains show similar growth rates and cell cycle progression kinetics, suggesting that any observed difference in DNA replication rates in CRE mutants is unlikely to be connected to growth or cell cycle perturbations.

6.1.3. Deletion of *ISW2* and *IES6* leads to changes of replication efficiency at selected replication origins

I first investigated whether the genetic knockout of the ISW2 and INO80 chromatin remodeling enzymes changes the replication timing and/or efficiency properties of selected early-efficient (EE) ARS305/ ARS315 or late-inefficient (LI) ARS313/ ARS316) origins on chromosome III compared to a wildtype strain (**Figure 18**).

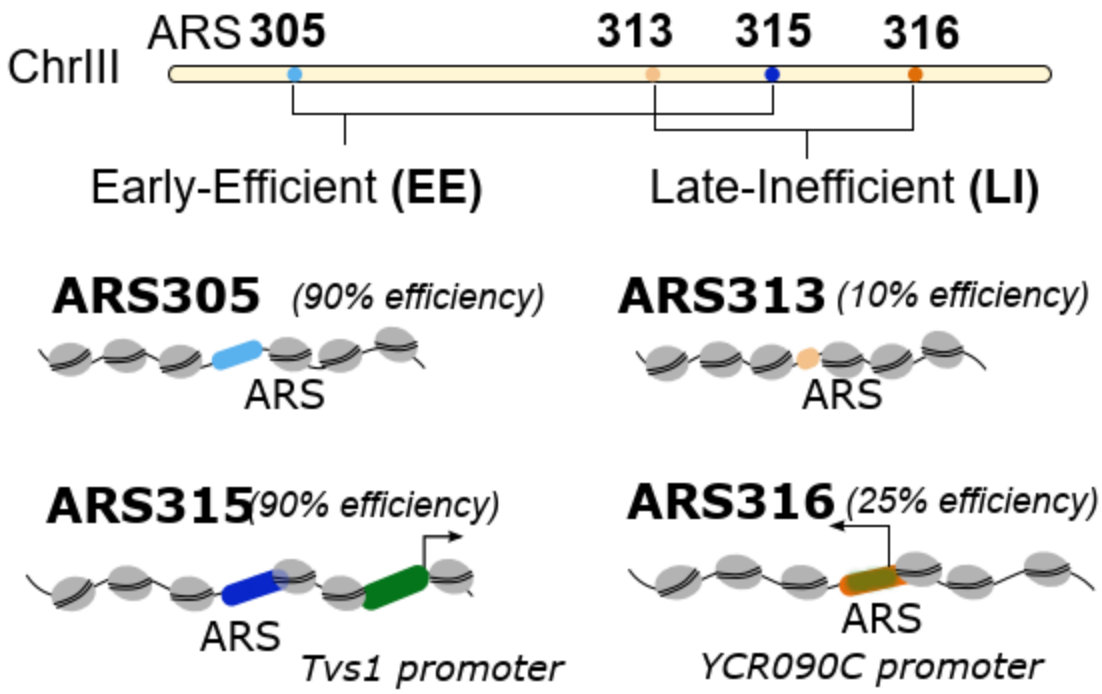


Figure 18. A. Schematic representation of Early-Efficient (EE) and Late-Inefficient (LI) replication origins on Chromosome III of *Saccharomyces cerevisiae*. The two EE origins (ARS305/ARS315 – ARS blue) show 90% efficiency and two LI origins (ARS313/ARS316 – ARS orange) show 10% and 25% efficiency, respectively. Distinct chromatin features are indicated around the ARS loci such as nucleosomes (grey) and gene promoters (green) at or in vicinity to the replication origins.

To this end, wildtype and CRE mutant cells deleted for subunits of the ISW2 (*isw2*Δ) and INO80 (*ies6*Δ) remodeling complexes were synchronized in G1-phase by addition of α-factor and subsequently released into S-phase by addition of 125 U of Pronase. Samples were collected at the indicated timepoints and subjected to genomic DNA extraction and copy number analysis by qPCR (**Figure 19**). In wildtype cells (Control), when comparing to a described late replicating reference region on ChrIV (Batrakou et al., 2018), the relative copy number of the EE regions ARS305 and ARS315 increased from 0 to 32 min after release, reflecting the earlier replication status compared to a described late replicating region of ChrIV. After this timepoint, replication is also initiated from the reference and copy number ratios decrease until S-phase is completed ~ 48 min after release into S-phase (**Figure 20 A-B**, Control). In contrast, the LI origins ARS313 and ARS316 initiate replication with similar kinetics compared to the ChrIV reference, resulting in a constant copy number ratio of ~ 1 over the complete timecourse of the experiment (**Figure 20C-D**, Control).

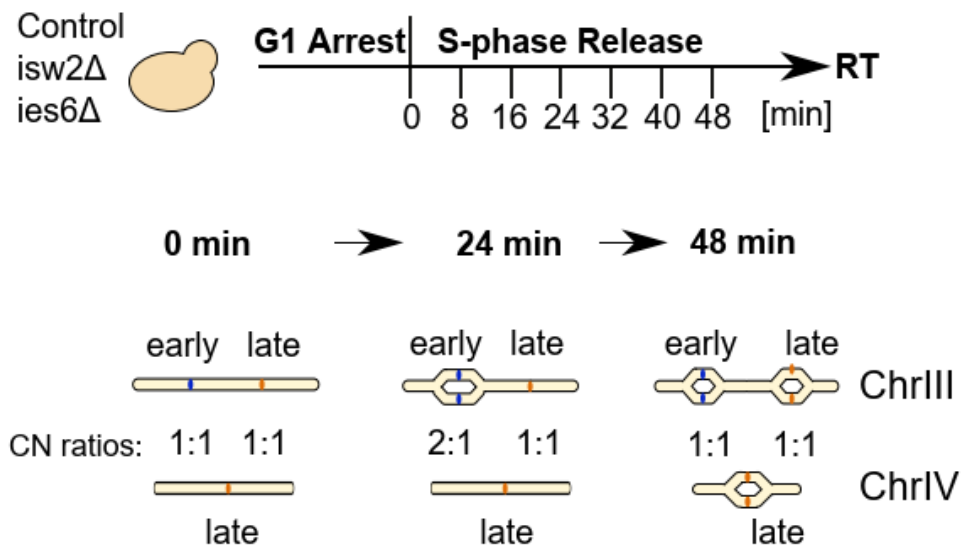


Figure 19. Experimental outline for measuring replication timing of each origin (Control, *isw2*Δ, *ies6*Δ). After 2 h arrest, the cells were released into S-phase and samples were collected every 8 min for FACS and qPCR analysis.

In *isw2*Δ and *ies6*Δ cells, the replication timing profiles of ARS305 showed significantly lower copy numbers when referenced to the late ChrIV region, particularly at 24, 32 and 40 min after release (**Figure 20A**, *ies6*Δ and *isw2*Δ). There was no difference between the profiles of the two CRE mutant strains, suggesting that knockout of ISW2 or INO80 decrease replication efficiency of ARS305 similarly by ~ 10% across multiple timepoints from early to mid/late S-phase. Similar results were obtained for the replication profile of the ARS315 origin, where both CRE mutant strains displayed lower efficiencies at 32 and 40 minutes after release (**Figure 20B**, *ies6*Δ and *isw2*Δ). These results suggested that deletion of the ISW2 and INO80 chromatin remodelers decreases the replication efficiency of EE origins, supporting the notion that the enzymatic activity of these two CREs is critical for replication timing control as demonstrated here for the two EE origins.

Interestingly, we observed the opposite result for the INO80 mutant strain at the ARS313 replication origin (**Figure 20C**, *ies6*Δ). At this origin, the INO80 mutant displayed significantly higher copy numbers in early S-phase (~16 min after release) and mid S-phase (~32 min after release), suggesting that replication timing or efficiency of ARS313 over the ChrIV reference was increased. A similar trend could be observed in case of the ISW2 mutant, but the effect was less pronounced and statistically not significant (**Figure 20C**, *isw2*Δ). However, the second LI origin

ARS316 showed identical DNA copy numbers for wildtype and CRE mutant strains across the timecourse (**Figure 20D**), suggesting that ISW2 and INO80 chromatin remodelers have no measurable effect on the DNA replication profile at this particular origin.

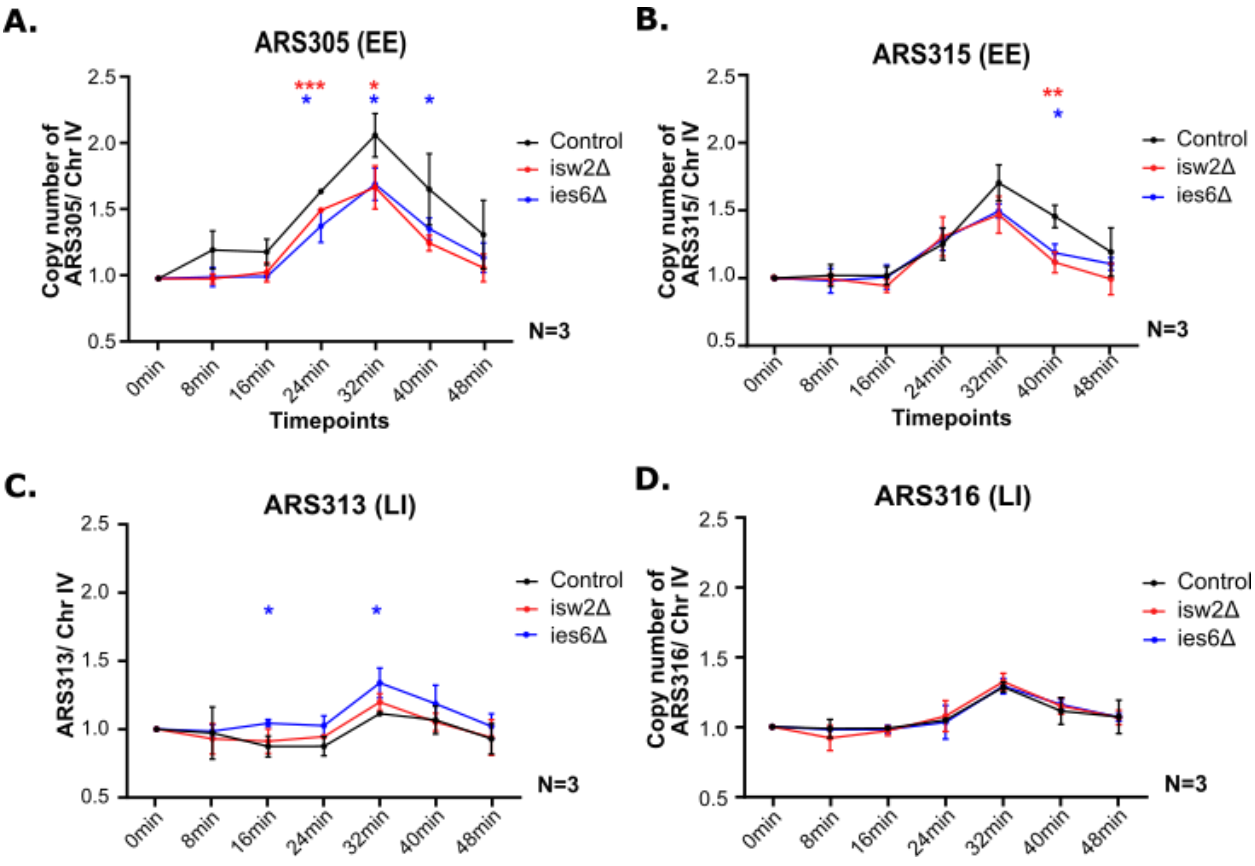


Figure 20. The replication timing plots show the average copy number ratios of ARS305, ARS315, ARS313, ARS316 compared to the late replicating region of ChrIV with standard deviation from n = 3 biological replicates (*, **, *** indicates statistical significance P < 0.05, or P < 0.01, or P < 0.001, unpaired t-test). They were used Y66 (control), Y104 (*isw2*Δ) and Y136 (*ies6*Δ).

Together, these data suggested that ISW2 and INO80 chromatin remodelers are affecting the DNA replication landscape at selected replication origins. This occurs to a varying extent and effect size, depending on which origin was investigated. In order to better understand the underlying molecular cause of these changes, I further profiled the chromatin landscape and in particular how the nucleosome positioning around these four replication origins was affected in the wildtype and CRE mutant strains.

6.2 Restriction enzyme accessibility assay reveals nucleosome occupancy heterogeneity at ARS regions.

As a first approach to measure bulk nucleosome occupancy at our selected replication origins, we applied classical restriction enzyme (RE) accessibility assays (**Figure 21**). The presence of nucleosomes at restriction site sequences blocks the accessibility of the restriction enzymes and, thus, halts the chromatin digestion. For this study, we define the +1 nucleosome of an origin as the nucleosome that is closest to the annotated ACS within the ARS where the initial binding of ORC occurs. Accordingly, we used enzymes that digest specifically at individual positions around the ARS expected to be protected by NS-2, NS-1, NS+1 or NS+2 nucleosomes around the ARS (**Figure 22A**). For each restriction enzyme, I used two different amounts of the enzyme to accomplish saturation of the digestion reaction, thereby ensuring that incomplete chromatin digestion is exclusively caused by the presence of a nucleosome. Nuclei isolated from logarithmically growing wildtype cells were digested with two different amounts of RE and analyzed by Southern blot analysis with specific probes directed against the selected EE and LI origins. After the initial enzymatic digestion on chromatin level, DNA extraction was performed and each locus was digested with a second set of restriction enzymes. This strategy allows unambiguous detection of both digested and undigested DNA fragments by their different sizes using indirect end-labeling Southern blot analysis (**Figure 21**). Importantly, increasing the amount of REs did not increase the efficiency of the digestion reaction, supporting the notion that the observed incomplete digestion is exclusively caused by the protective role of chromatin-bound factors, most likely caused by nucleosomes.

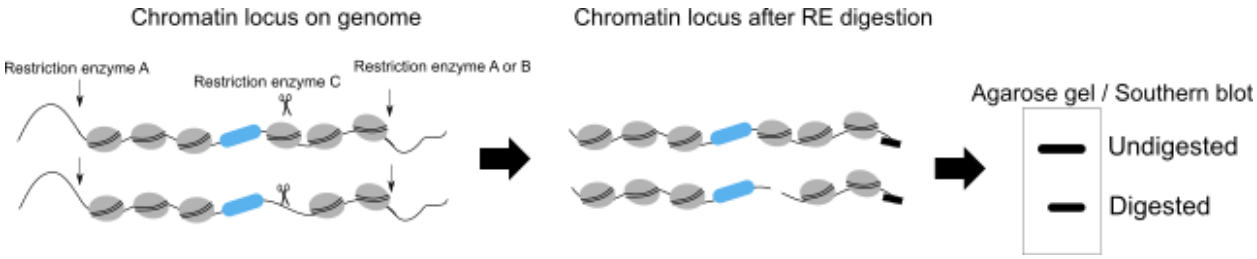


Figure 21. Cartoon showing the principal of Restriction Enzyme Accessibility (REA) assay.

In the wildtype strain, we observed similar percentages of digested molecules (~30-40%) at the NS+1 and NS+2 restriction sites among all 4 origins (**Figure 22B**), indicating that around 30-40% of all molecules were not protected by nucleosomes at these two positions. Interestingly, we

observed major differences in RE accessibility between origins at the NS-2 and NS-1 positions. For example, the NS-2 HpaI site only showed 20% digestion at ARS305, whereas more than 60% of molecules were digested by HpaI in the same relative NS-2 position of ARS316 (**Figure 22B**). The ARS305 NS-1 position also showed strong protection (~10% digestion), whereas more than 50% of the molecules were accessible at this site in ARS316 (**Figure 22B**), consistent with a strongly positioned NS-1 nucleosome at the EE origin ARS305.

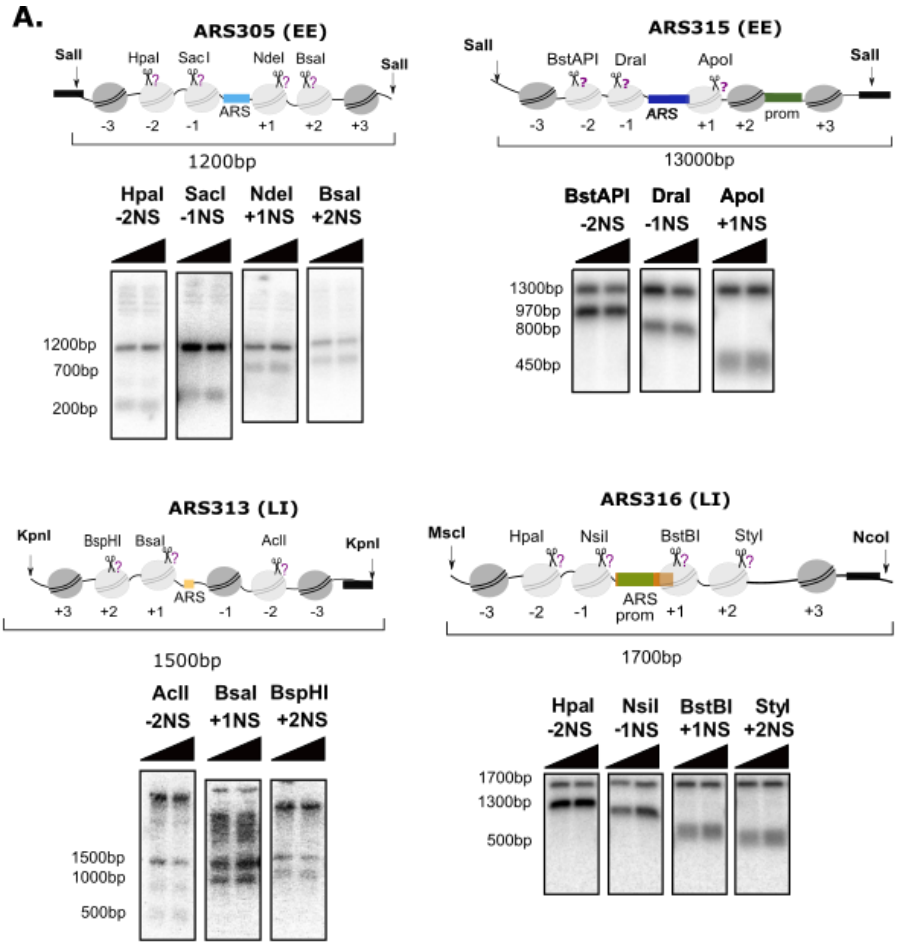


Figure 22. A. Restriction endonuclease accessibilities in chromosomal ARS305, ARS315, ARS313 and ARS316 locus. Nuclei from yeast strains Y65 (ARS305), Y91 (ARS315), Y94 (ARS313) and Y69 (ARS316) were isolated and digested with increasing amounts (10 U, 50 U or 100 U) of the indicated restriction enzymes (triangle on top of each pair of panels / scissors). DNA was isolated, digested with SalI in ARS305 and ARS315, KpnI in ARS313 and MscI/NcoI in ARS316 and subjected to indirect end-labeling Southern blot analysis with a specific radioactively labeled probe. **B.** The histogram shows the results of Southern blot quantification as a percentage of digested chromatin locus derived from n=3 biological replicates.

6.3 Restriction enzyme accessibility assay reveals changes in nucleosome occupancy in INO80 mutant strain at ARS315

In order to test whether nucleosome occupancy or positioning has been altered between wildtype (Control) and the INO80 mutant strain, I applied the REA assay in both strains and analyzed changes in nucleosome occupancy at ARS315 using the same restriction enzymes as above to profile the level of nucleosomal protection of the -2NS, -1NS and +2NS sites. In this experiment, an additional naked spike-in plasmid (K273) was added to the digestion reaction of the INO80 mutant strain. This plasmid contains restriction sites for all 3 RE and was used to confirm that each enzyme (BstAPI, DraI, ApoI) can digest unprotected naked DNA with 100% efficiency under the used chromatin digestion conditions (**Figure 23**). After chromatin digestion and DNA extraction, the ratio of undigested and digested DNA fragments derived from the ARS315 locus was analyzed by indirect-end labeling Southern blot analysis. For the resulting Southern blot in the INO80 mutant strain, the expected additional DNA fragments that are derived from the spike-in plasmid are indicated by asterisks and confirmed 100% digestion efficiency of the BstAPI, DraI, and ApoI enzymes (**Figure 24A**). The bar plot shows percentage of digested, and therefore non-nucleosomal, chromatin fragments over the total population. The DraI and ApoI sites flanking the ARS in closer distance showed an 18% and 24% increase in accessibility (P=0.036 and P=0.047), respectively (**Figure 24B**), suggesting that upon loss of INO80, the typically well positioned nucleosomes in close vicinity to the ARS are evicted or repositioned, thereby making this site more accessible in the mutant strain. However, there is no significant difference between the wildtype and the CRE mutant strain on nucleosome occupancy at BstAPI site located 222 bp upstream the ARS locus.

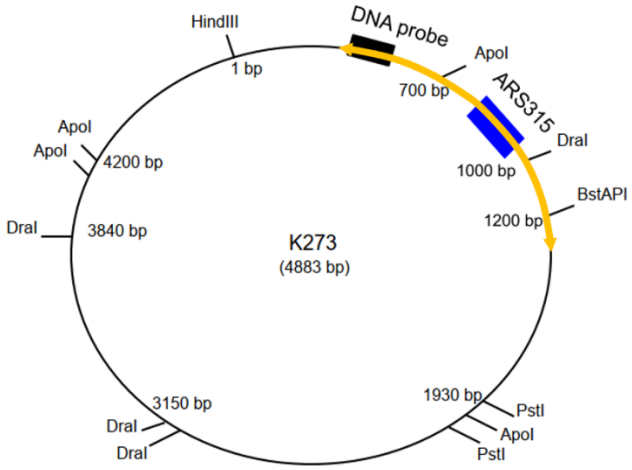


Figure 23. Schematic presentation of plasmid K273 containing the DNA sequence of ARS315 locus used as spike in control of the REA analysis of *ies6*Δ-ARS315 strains. The shared DNA sequence between the plasmid and the chromatin locus is highlighted in yellow. The black rectangle indicates the position of the DNA probe used for the Southern blot and the blue rectangle indicates the position of the ARS315 locus. On the map it is shown the distance between the restriction sites using the HindIII site as the starting point. Similar to the native chromatin locus, the plasmid was initially digested by BstAPI, DraI and ApoI, separately, and later all the samples were digested by HindIII and PstI (secondary digestion). The expected detectable DNA fragments according to the two digestion steps and the position of the probe are 1200 bp (for BstAPI and HindIII), 1000bp (for DraI and HindIII) and 700bp (for ApoI and HindIII). The larger fragments 4110 bp (in BstAPI) and 1300bp (in ApoI) indicate inefficient digestion of HindIII. All the fragments deriving from the plasmid are noted by asterisks in **Figure 24**.

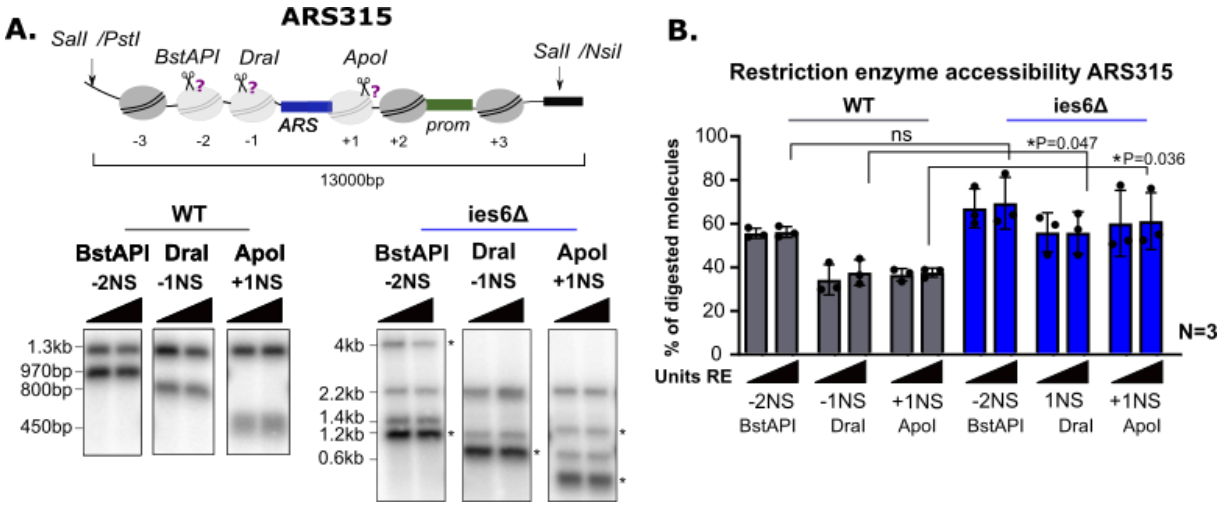


Figure 24. A. Restriction endonuclease accessibilities (REAs) in chromosomal ARS315 locus. Nuclei from yeast strains Y91 (control-ARS315), and Y130 (*ies6*Δ-ARS315) were isolated and digested with increasing amounts (10 U, 50 U and 100 U) of the indicated restriction enzymes (scissors). DNA was isolated, digested with SalI (wildetype) or PstI/NsiI (ies6Δ) and subjected to indirect end-labeling Southern blot analysis with a specific radioactively labeled probe. Top: schematic representation of the ARS315 locus with restriction sites used to probe chromatin structure (scissors) and secondary restriction sites to isolate the locus. Asterisks in the mutant strain indicate restriction bands derived from a spike-in naked plasmid control to verify complete digestion of the enzymes (**Figure 23**). **B.** The histogram shows the results of Southern blot quantification as a percentage of digested versus undigested chromatin fragments. Average and standard deviations are from n= 3 biological replicates (*, **, *** indicates statistical significance with $P < 0.05$, or $P < 0.01$, or $P < 0.001$, unpaired t-test).

To investigate further the changes in the chromatin accessibility landscape around the replication origins among the cell population, the establishment of a single molecule approach was necessary.

6.4. Purification of defined replication origins by site-specific recombination *in vivo*

In order to analyze the precise nucleosomal arrangement of the selected replication origins, I took advantage of a site-specific recombination approach to isolate specific ARS chromatin domains derived from four replication origins with distinct replication profiles (ARS305/ARS315 -Early and Efficient whereas ARS313/ARS316- Late and Inefficient) in their native chromatin context. This approach was previously established by the groups of Roger Kornberg and Joachim Griesenbeck and applied to the multi-copy ribosomal DNA (rDNA) locus as well as the single-copy *PHO5* gene locus (Griesenbeck et al., 2003). In brief, the genomic region of interest is flanked by two recognition sites for the site-specific R-recombinase (RS) from *Zygosaccharomyces rouxii* (Araki et al., 1992). Next to one of the RS sites, a cluster of three DNA binding sites for the bacterial LexA protein is additionally inserted (**Figure 25**). The R-recombinase recognizes the RS sites and does not require any co-factor or energy source (e.g., ATP hydrolysis) for the recombination reaction (Ansari et al., 1999). The enzyme mediates staggered DNA cleavage at specific bases of each strand and subsequently, excises the intervening area between the RS sites – in this case containing the four replication origins of our interest – from the chromosome as a closed chromatin ring. This ring includes one RS site, the area of interest and the cluster of 3x*LEXA* binding sites. Expression of a LexA protein fused to a C-terminal TAP-tag is controlled by TEF2 promoter whereas R-recombinase is controlled by an inducible *GAL1-10* promoter. Thus, addition of galactose to the medium results in the expression of the R-recombinase that excises the area of

interest as a ring and subsequently allows for affinity purification of the native chromatin domains (**Figure 25**)

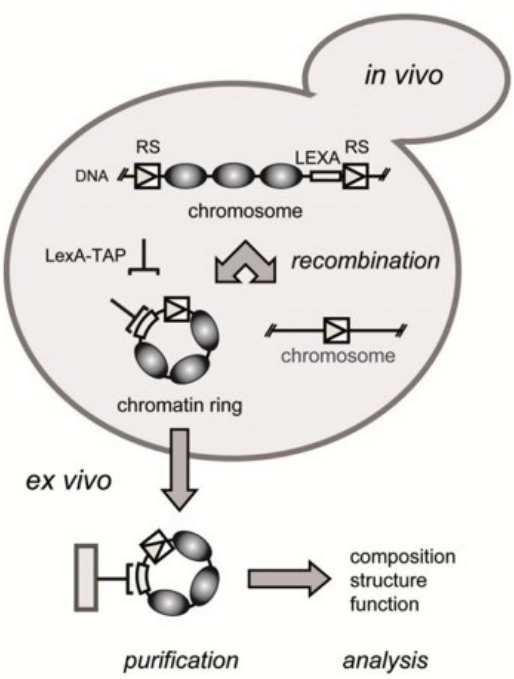


Figure 25. Cartoon showing the purification strategy of specific chromosomal domains. The area of interest is flanked by recognition sites of R-recombinase and a cluster of *LexA* sites. After expression of R-recombinase the enzyme extracts the specific locus as a ring (RS, boxed arrowheads). Additionally, the expressed *LexA* protein binds to the *LexA* binding sites. After cell lysis, the soluble chromatin rings can be purified via a recombinant LexA-TAP fusion protein (LexA-TAP, bracket connected to a line) as well as to an affinity support (magnetic beads - filled rectangle). Filled ovals represent nucleosomes and other chromatin components. (Adapted from Hamperl 2012)

6.4.1 Establishment of yeast strains with a modified ARS locus competent for excision of distinct ARS chromatin domains

To generate yeast strains competent for the site-specific recombination and affinity purification of the selected origins, the origins were first deleted by homologous recombination (HR) and replaced by an *URA3* selectable marker. The survival clones containing the *URA* marker were subjected to homologous recombination with plasmids containing the area of interest (ARS with the 3 surrounding nucleosomal pairs) flanked by RS/*LexA* sites and transferred on Fluoroorotic Acid Monohydrate (FOA) plates. 5-FOA is toxic for yeast containing *URA3*- gene and therefore yeast clones without successful homologous recombination are unable to survive (**Figure 26**). In

this way, a library of yeast strains was created where each of the four origins was tagged with RS/*LEXA* sites at increasing distances after the first, second or third pair of nucleosomes around the annotated ARS region. In this work, I focused on the purification of chromatin domains containing three pairs of nucleosomes in order to obtain a more comprehensive view on the nucleosomal landscape surrounding the origins of interest. At the latest stages of the project (see results 3.5 and 4 part) I took advantage of a previously established strain that contains RS/*LEXA* sites integrated into the multicopy ribosomal ARS (rARS) locus allowing the purification of a domain containing two pairs of nucleosomes around the annotated rARS (Hamperl et al., 2014).

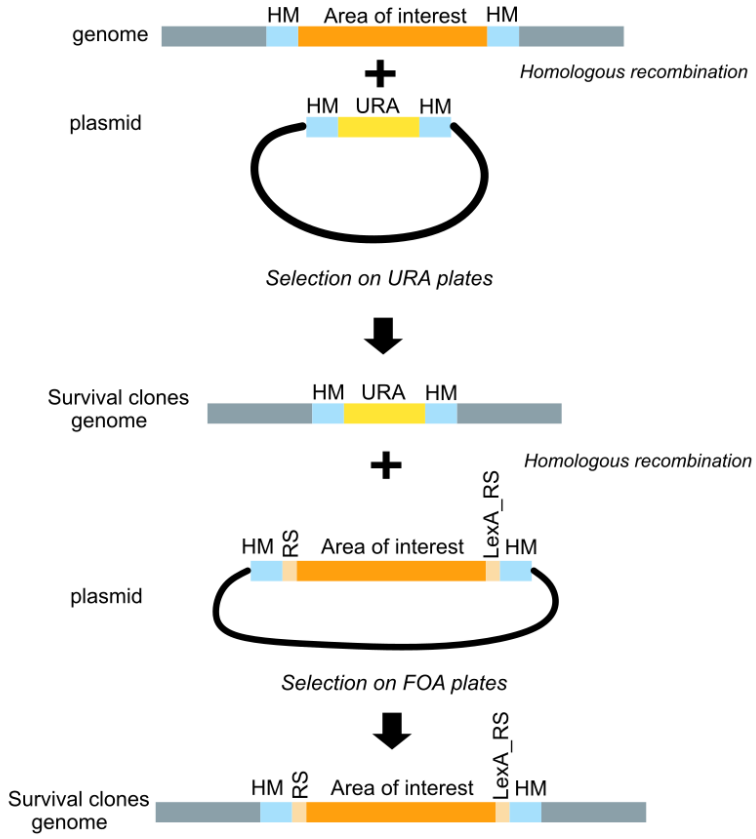
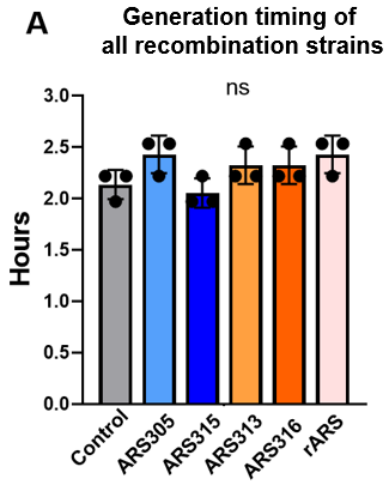


Figure 26. Schematic representation of the cloning strategy used for establishment of yeast strains competent for recombination. The orange rectangles represent the ARS locus with the surrounding nucleosomes which will be subjected to purification. The light blue rectangles represent the neighboring regions with identical DNA sequence used for the homologous recombination. The yellow rectangles correspond to uracil selection marker gene. Yeast strains containing this *URA* gene survive on –URA plates, but not on FOA plates. The light orange rectangles represent the recognition sites of R-recombinase and *LexA* protein.

6.4.2 Genetic modifications of the replication origins do not interfere with cell growth and origin chromatin structure and function

6.4.2.1 The recombination strains show normal cell growth and cell cycle dynamics

After confirming the successful integration of RS/LEXA sites at the selected replication origins by genotyping PCR and Sanger sequencing, I excluded the possibility that the inserted genetic elements affect cell growth or show other defects in cell cycle distribution and dynamics. As a control, an isogenic strain was used that also expresses the R recombinase and LexA-TAP protein, but does not contain any insertion of RS or LEXA binding sites in the genome. The cells were grown in YPD medium and OD600 measurements were taken every 30 sec in a multi-plate reader. No significant difference in generation timing was observed between the control and the five recombination strains (**Figure 27A**). To test the cell cycle dynamics, strains were cultured in YPR medium and arrested in G1 by α-factor treatment for 2 h. Samples for FACS analysis were collected from asynchronous cells and at 30 min time intervals (**Figure 27B**). Asynchronous samples of control, ARS313 and ARS315 cells showed similar levels of G1, S or G2 populations as expected (**Figure 27B**). Upon addition of alpha factor, the percentage of G1 cells gradually increased over the timecourse and there was no difference in the individual arrest kinetics between the three strains (**Figure 27B**), indicating that the insertion of RS/LEXA sites does not interfere with the global cell cycle kinetics.



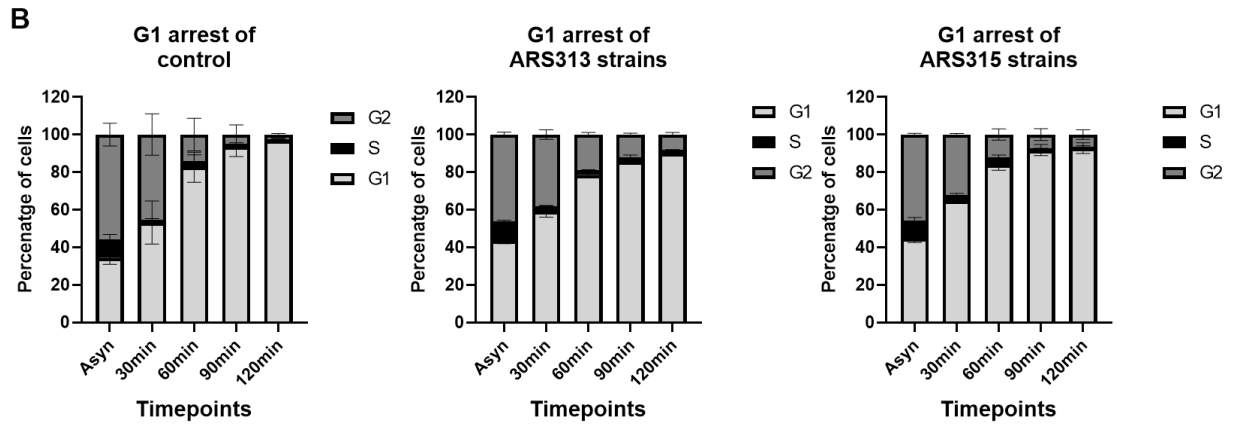


Figure 27. A. Generation timing of yeast strains competent for recombination (Y65-ARS305, Y91-ARS315, Y94-ARS313, Y69-ARS316, Y84-rARS) and control (Y66-without RS/*LEXA*). The strains were grown in YPD medium and OD600 measurements were taken every 360 sec by a multi-plate reader. The bar graph depicts the doubling time of each strain during the exponential growth phase of the cells **B.** Cell arrest in G1-phase. The strains Y91 (ARS315) and Y94 (ARS313) were grown in YPR medium and arrested with 50 mg/ml α-factor. Samples for FACS analysis were taken at the indicated timepoints and stained by Sytox green to monitor the distribution of G1, S and G2-phase in both profiles.

6.4.2.2 R Recombinase excises the targeted replication origin domains with high efficiency

Next, I wanted to investigate whether the inserted RS sites are functional and lead to the expected recombination of the targeted replication origins. To this end, control, ARS315 and ARS313 cells were grown in YPR medium to exponential phase and recombination induced by the addition of galactose to the medium. The recombination kinetics and efficiency under these conditions were monitored in a time-course experiment by Southern blotting of extracted genomic DNA. In both ARS315 and ARS313 strains, near complete recombination of the targeted loci was observed within ~60-90 min after recombinase induction (**Figure 28**), whereas the genomic locus of ARS315 in the control strain remained intact across all timepoints (**Figure 28A**).

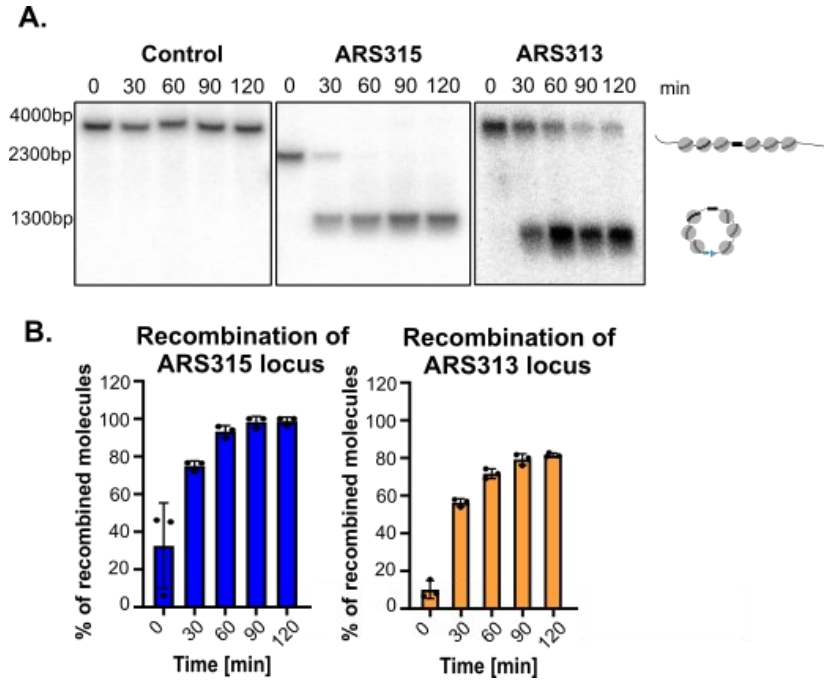


Figure 28. A. Recombination kinetics of ARS315 and ARS313 loci. The strains Y91-ARS315, Y94-ARS313 and Y66-control (w/o RS or *LexA* sites) were grown in YPR medium to logarithmic phase and arrested in G1 phase by addition of α-factor (50 ng/ml) and recombination was induced by the addition of 2 % galactose. Samples were taken at indicated timepoints. DNA was isolated and linearized by BsrGI (Y66 and Y91) and by NcoI (Y94) and subjected to Southern blot analysis. The positions of unrecombined and recombined molecules are shown on the right. **B.** The histogram shows the results of Southern blot (n=3) as percentage of recombined chromatin locus.

Generally, these experiments indicated that the recombination reaction occurs with high specificity and yield and the necessary genetic modifications of the strains did not affect the physiological cell growth and cell cycle dynamics of the established yeast strains in comparison to the parental control strain. Therefore, further use of the strains and investigation of the nucleosome positioning could be performed with confidence.

6.4.2.3. Integration of the RS/*LEXA* binding sites does not interfere with the local nucleosome positioning of the replication origins

Next, I wanted to ensure that the integration of RS/*LEXA* sites directly after the 1st, 2nd and 3rd pair of nucleosomes does not disrupt their positioning around the ARS site. To investigate the nucleosomal pattern of the recombination and parental control strains, I performed Micrococcal nuclease (Mnase) assay in combination with Southern blot analysis. Mnase has the ability to digest only the non-nucleosomal DNA in a sequence-independent manner, whereas its activity is blocked by the presence of a nucleosomal core particle. After indirect end-labeling Southern blot analysis, the digested accessible DNA regions are depicted as individual bands, separated by the undigested DNA, together revealing a pattern that represents the local nucleosomal profile of the locus. For this experiment, a control strain without inserted RS/*LEXA* biding sites was used to show the native nucleosomal profile of the ARS305 replication origin in the absence of genetic modifications. Nuclei isolated from all strains (control, ARS305 +/-1NS, +/-2NS and +/-3NS) were incubated with 10 U of Mnase and samples were collected at different timepoints (0, 5, 10, 35, 60 min) (**Figure 29**). After Mnase digestion, DNA extraction was performed and the region around the ARS305 locus was isolated by a restriction enzyme digestion using BamHI and PstI restriction enzymes. The total length of this DNA fragment is ~2 kb and the nucleosome-free ARS region is located at the center of the fragment. The Southern blot probe was designed next to the PstI recognition site (**Figure 29**, green box). In control samples, the ARS is displayed as highly accessible region (see timepoint 10 min), as known from literature (Eaton et al., 2010). The nucleosomes around the ARS are well positioned and separated especially in the lower part of the agarose gel. The BamHI/PstI fragment of ARS305 derived from the recombination strains is ~2.2 kb due to the additional RS/*LEXA* sites inserted at different positions along the locus. Importantly, in all strains, the ARS region was highly accessible similar to the control, indicating that the nucleosome-free region was formed as expected and not affected by the RS/*LEXA* sites. Additionally, the cluster of 3 *LEXA* sites was highly accessible, consistent with previous results that this sequence adopts a nucleosome-free region and therefore allows efficient binding of the LexA-TAP protein (Hamperl et al., 2014b). In addition, the recombination strains revealed footprints of protected DNA, consistent with a pattern of one, two or three nucleosomal pairs surrounding the ARS, suggesting that the genetically modified strains with additional RS and *LEXA* binding sites adopt a similar nucleosomal profile as the endogenous control strain without genetic modifications.

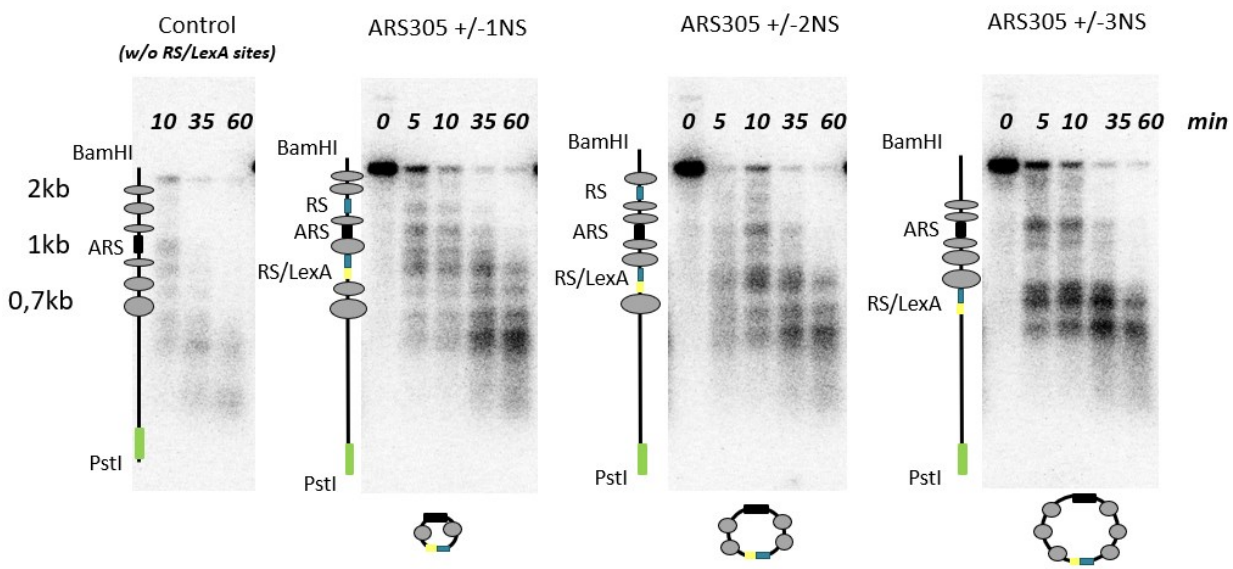


Figure 29. Comparison of nucleosomal pattern at ARS305 replication origin between Control (no RS/*LexA* binding sites) and recombined strains. On the left of each membrane is shown the schematic representation of each chromatin locus. The grey ovals represent nucleosomes, the black, blue and yellow rectangles represent ARS, RS and *LexA* sites, respectively. The enzymes used for secondary digestion of DNA are indicated on the two ends of the cartoons. The light green rectangles represent the position of the DNA probe used for Southern blot analysis. The circular formations bellow the membranes show how each locus looks after recombination.

6.4.2.4 Replication origins remain functional after site-specific recombination

Given that the R-recombinase excise the origins from its endogenous location, I wanted to investigate whether the origins retain their replication properties regarding the timing and efficiency even after uncoupling from its genomic chromatin context. To test this, I compared the replication timing of ARS315 in the control strain without RS/*LEXA* sites with the ARS315 strain competent for excision of this origin. The strains were arrested and in parallel were supplemented with Galactose to induce recombination. After synchronous release into S-phase, samples for FACS and qPCR analysis were obtained at the indicated timepoints (**Figure 30**). According to the FACS profiles, both strains showed highly similar arrest and release kinetics. After 2 h of alpha-factor treatment, the majority of cells (~95 %) was arrested in G1 phase (**Figure 30A-B**). Addition of 125 U Pronease resulted in synchronous release into S-phase after 24 min and both strains replicated their DNA and progressed through S-phase within the residual timepoints of the experiment (**Figure 30C**, 32 to 56 min release). Given that the S-phase progression was highly

similar between the two strains, I performed DNA copy number analysis at the indicated timepoints after release by qPCR comparing the replication timing of ARS315 in its genomic location (Control, black) versus in its excised state as a chromatin circle (ARS315, blue) (**Figure 30C**). When comparing the ratio of ARS315 replication with a late replicating region on chromosome IV (Chr4, **Figure 20**), the two replication profiles showed similar kinetics where ARS315 started replicating between 24 to 32 min and DNA the copy number ratio increased. After ~40 min release, the region on Chr4 also started to replicate and DNA the copy number ratio decreased in both strains, consistent with earlier replication of ARS315 compared to Chr4. However, the excised ARS315 chromatin circles did not reach the same maximum DNA copy number ratio at 40 min as the Control genomic ARS315, suggesting that a fraction of chromatin circles did not replicate as efficiently as in its genomic context which could stem from the lack of stochastic firing of neighboring origins and passive replication when isolated from its genomic context.

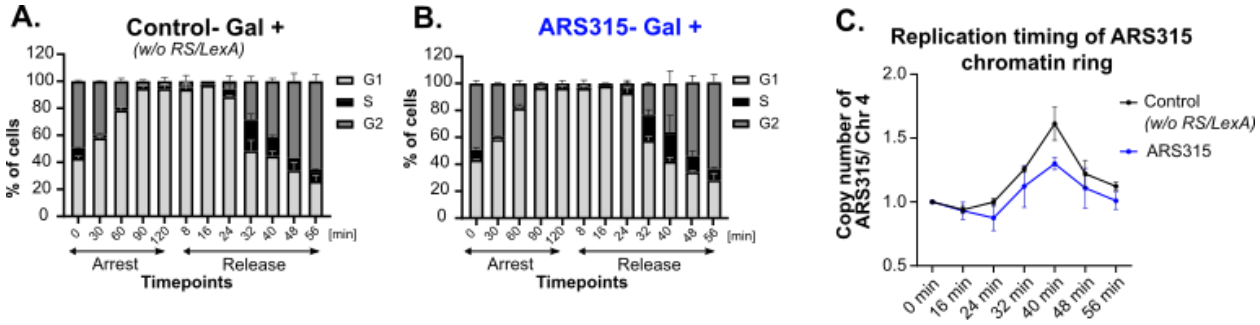


Figure 30. A-B. Cell cycle progression of control (Y66) (w/o RS/LexA sites) and ARS315 (Y91-with RS/*LexA* sites) strains after addition of 2 % galactose. The cells were arrested by 50 ng/ml α-factor and released into replication by addition of 125 U Pronease. The FACS samples and qPCR samples were taken side-by-side at the indicated timepoints. **C.** The replication timing plots show the average copy number ratios after addition of 2 % galactose of ARS315 locus, either in the Control (Y66) strain containing no RS/*LexA* sites for recombination or in the ARS315 (Y91) containing RS/*LexA* sites, to the late replicating region of Chr4 with standard deviation from n = 2 biological replicates. The cells have been arrested for 2 h by α- factor and released to S-phase by 125 U Pronease.

6.5.1 Establishment of a purification technique for selected chromosomal domains

Having confirmed that the site-specific recombination efficiently releases the targeted replication origins in a native chromatin state that accurately reflects the nucleosome configuration and functional state of the endogenous locus, I proceeded with affinity purification of the domains according to a published protocol (Hamperl et al., 2014b; Weiβ et al., 2023) with minor

modifications. Briefly, magnetic IgG beads, and not IgG-sepharose or agarose, were used as affinity matrix for purification of the LexA-TAP- fusion protein from cell extracts. The advantage of magnetic beads is the increased number of affinity binding sites as well as their higher binding specificity resulting in lower contamination and higher purification yield. Initial attempts to retrieve the target chromatin using the original constructs and protocols (Hamperl et al., 2014b) suffered from low yields and high levels of contamination from genomic DNA and non-specific proteins. This could be overcome by replacing the original weak *CYC1* promoter driving LexA-TAP expression by the stronger, constitutive *TEF2* promoter. This modification increased LexA-TAP levels ~ 5 to 6-fold and markedly improved retention and purity of the chromatin domains upon purification (Weiβ et al., 2023).

After gentle cell lysis using dry ice and a coffee grinder, the soluble chromatin rings are separated from the cellular extract by centrifugation. IgG coated magnetic beads are added to the supernatant (Input) and subjected to 2 h incubation at 4 °C allowing the beads to bind onto the Protein A and therefore allow retention of the chromatin rings. After removing the flowthrough, (FT), beads are washed extensively prior to elution of the chromatin rings by cleavage of tobacco etch virus (TEV) protease. TEV protease specifically recognizes and digests the linker region between the C-terminal protein A moiety and the calmodulin binding peptide of the TAP tag (**Figure 31**). After cleavage for 24 h, the chromatin ARS domains are separated from the beads (B) and released into the final elution (E) (**Figure 31**). At each step of the purification process DNA and protein samples were collected to track the recovered chromatin ARS domain and LexA-TAP proteins. By convention, the nucleosomes located closer or further away from the ACS site of the ARS was termed as (+) or (-), respectively.

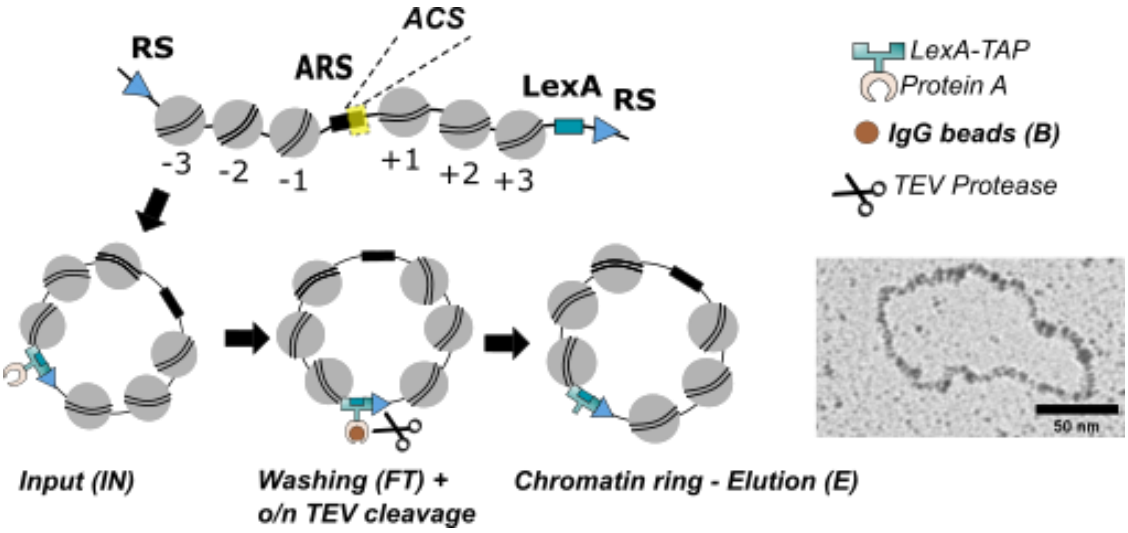


Figure 31. Schematic representation of LexA affinity purification. Recombination sites (RS) were integrated after the +/- 3 nucleosomes around the ARS locus of single-copy origins and after the +/-2 nucleosomes of the multi-copy rARS origin. Electron micrograph shows individual ribosomal ARS molecule after native spreading.

Western blot analysis of LexA-TAP in the different fractions showed near-complete depletion of LexA-TAP in the flowthrough and recovery in the final elution in all strains as expected (**Figure 32**). Using qPCR analysis, I quantified the relative enrichment of origin DNA (ARS315 or rARS) in the different fractions (**Figure 33**). The final yield derived from the multi-copy rARS locus ranged from 20 - 40 % of the starting amount present in the input fraction (**Figure 33A**), whereas the final yield derived from the single-copy ARS315 locus ranged between 10 - 20 % of the Input amount (**Figure 33B**). Measuring the DNA concentration of these final eluates, I could typically recover ~200-300 ng of total DNA from the rARS purifications and ~10-20 ng of total DNA from the single-copy ARS315 purifications. As purification control, I used an isogenic yeast strain which expresses the LexA-TAP fusion protein, but contains no RS or LEXA binding sites around the replication origins. An unrelated housekeeping gene (*PDC1*), which is not flanked by RS or *LEXA* binding sites, was quantitatively lost in the flowthrough fractions in both control and recombination strains, and no recovery of this locus was observed in both beads (B) and final elution (E) fractions, indicating a specific enrichment of the targeted loci. Indeed, the final eluates showed a very high enrichment of the ARS315 and rARS over the PDC1 locus and by factoring in the size of the yeast genome, it can be estimated that 80 % (ARS315) up to almost 100 % (rARS) of all chromatin molecules are derived from the specific loci of interest (**Figure 33**).

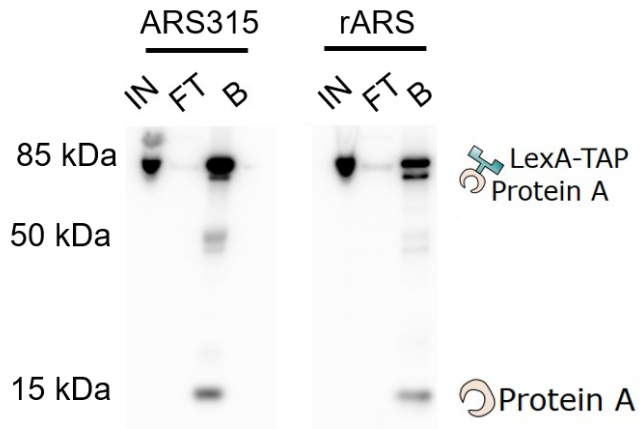


Figure 32. Protein analysis of LexA affinity purification process. The protein samples (0,5% IN, FT, B, E) were collected in parallel with DNA samples and loaded on an SDS-PAGE gel in order to monitor the LexA protein during the purification by Western blot. Samples were taken from Y91, where the single-copy locus

ARS315 is targeted by RS/LEXA after the third nucleosomal pair and Y84, where the multi-copy rARS locus is targeted by RS after the second nucleosomal pair.

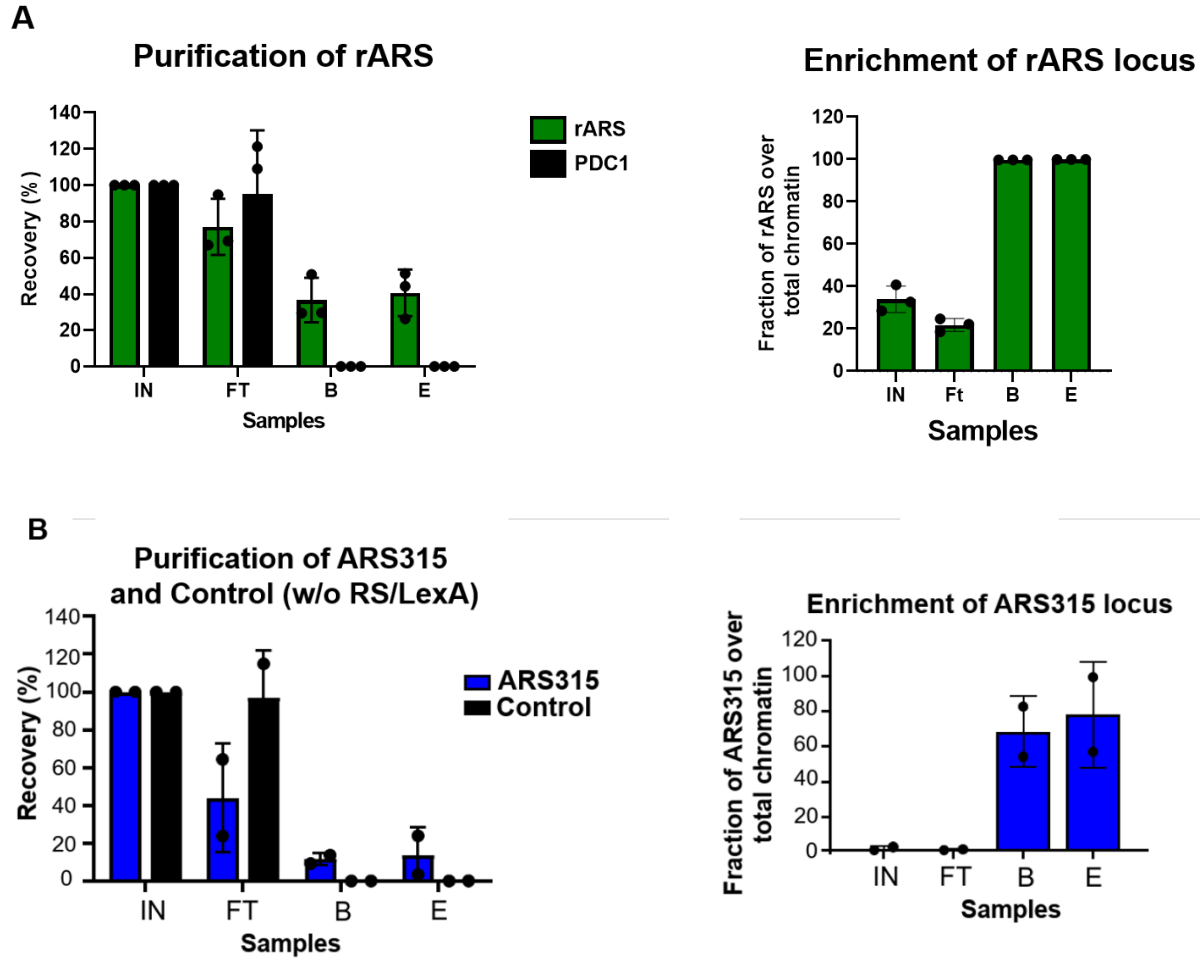


Figure 33 A. (Left panel) LexA affinity purification was performed for strain Y84, where the multi-copy rARS locus is targeted by RS. The PDC1 gene was used as a negative control for qPCR analysis. In both purifications, DNA samples were taken (0.1% Input (IN), Flowthrough (FT), Beads (B), Eluate (E) from n=2 (ARS315) and n=3 (rARS) biological replicates and analyzed by qPCR. (Right panel) Enrichment of rARS locus in the final eluate. Given that the size of yeast genome is ~12,000 kb, and the length of a chromatin ring is ~1 kb, the fold enrichment ratio of the specific origin to PDC1 was used to define the enrichment of an origin in the DNA samples. **B.** (Left panel) LexA affinity purification was performed for strain Y91 (ARS315), where the single-copy locus ARS315 is targeted by RS/LEXA, and strain Y66 (control), lacking RS/LEXA recognition sites but containing the LexA expression cassette, therefore no recombination or purified chromatin is expected. (Right panel) Enrichment of ARS315 locus in the final eluate, similar to rARS enrichment.

In summary, the established single-locus purification approach of the targeted chromatin ARS domains shows sufficient yield and specificity to pursue further single-molecule assays and characterize the nucleosomal landscape of the isolated domains.

6.5.2 Purification process and induction of recombination do not affect the local nucleosome positioning

Given that the recombination and the purification processes represent major manipulations that could change the composition and structure of the native chromatin locus, it is important to verify that the procedure does not alter the local nucleosome arrangement and the structure of the final purified chromatin ring is preserved similar to the *in vivo* situation. Therefore, I decided to compare the nucleosome occupancy at specific chromatin sites around ARS between the native locus, the recombined locus and the purified chromatin rings using Restriction Enzyme Accessibility (REA) assay.

For this analysis I picked two different strains competent for recombination and purification of ARS305 and ARS313 locus, respectively. The cells were grown to exponential phase, arrested in G1 phase by 50ng/ml α-factor and in parallel recombination was induced by addition of galactose. Samples were obtained before / after recombination and after purification as depicted in the **Figure 34** above the Southern membranes. For each origin, I selected four restriction enzymes allowing cleavage at specific sites occupied by +/- 2 nucleosome (NS) and/or +/- 1 NS as indicated (**Figure 34**). The selected restriction enzymes only digest chromatin when present in the non-nucleosomal state, whereas the presence of nucleosomes blocks the accessibility and prevents their activity. To ensure that saturation of the restriction enzyme reaction is achieved and the limited accessibility is indeed dependent on the protection by chromatin, I used increasing amounts of the enzyme, (indicated by a triangle on top of the gel, **Figure 34**). After DNA extraction, the genomic loci around the ARS305 and ARS313 were digested by secondary restriction enzymes (**Figure 34**), whereas the recombined and purified rings were linearized by a single restriction enzyme digestion. All the samples were loaded onto a 1% agarose gel and analyzed by indirect end-labeling Southern blot assay. The undigested ARS305 and ARS313 genomic loci have a total length of 1.2 kb and 1.5 kb, respectively, whereas the recombined and/or purified ARS305 and ARS313 rings have a total length of 1.25 kb and 1.2 kb, respectively. The size of the digested fragments is dependent on the distance between the enzymatic cut site and the designed probe (**Figure 34**).

The resulting Southern blot signals derived from digested and undigested molecules were quantified. In this way, I obtained the percentage of molecules that was not protected by a nucleosome. First, rising amounts of the used restriction enzymes did not significantly increase the percentage of digested molecules, suggesting that the incomplete chromatin digestion is exclusively caused by the presence of a nucleosome. Second, comparing the percentage of digested molecules derived from the genomically integrated locus versus the recombined or purified ring chromatin did not show statistically significant changes for any of the chosen restriction sites (**Figure 34**- bar plots). Therefore, I conclude that major structural properties of ARS305 and ARS313 chromatin are preserved upon excision from the chromosomal location and biochemical purification of the locus. Therefore, the final purified chromatin domains represent suitable substrates for the investigation of the nucleosomal landscape and the local ARS chromatin structure.

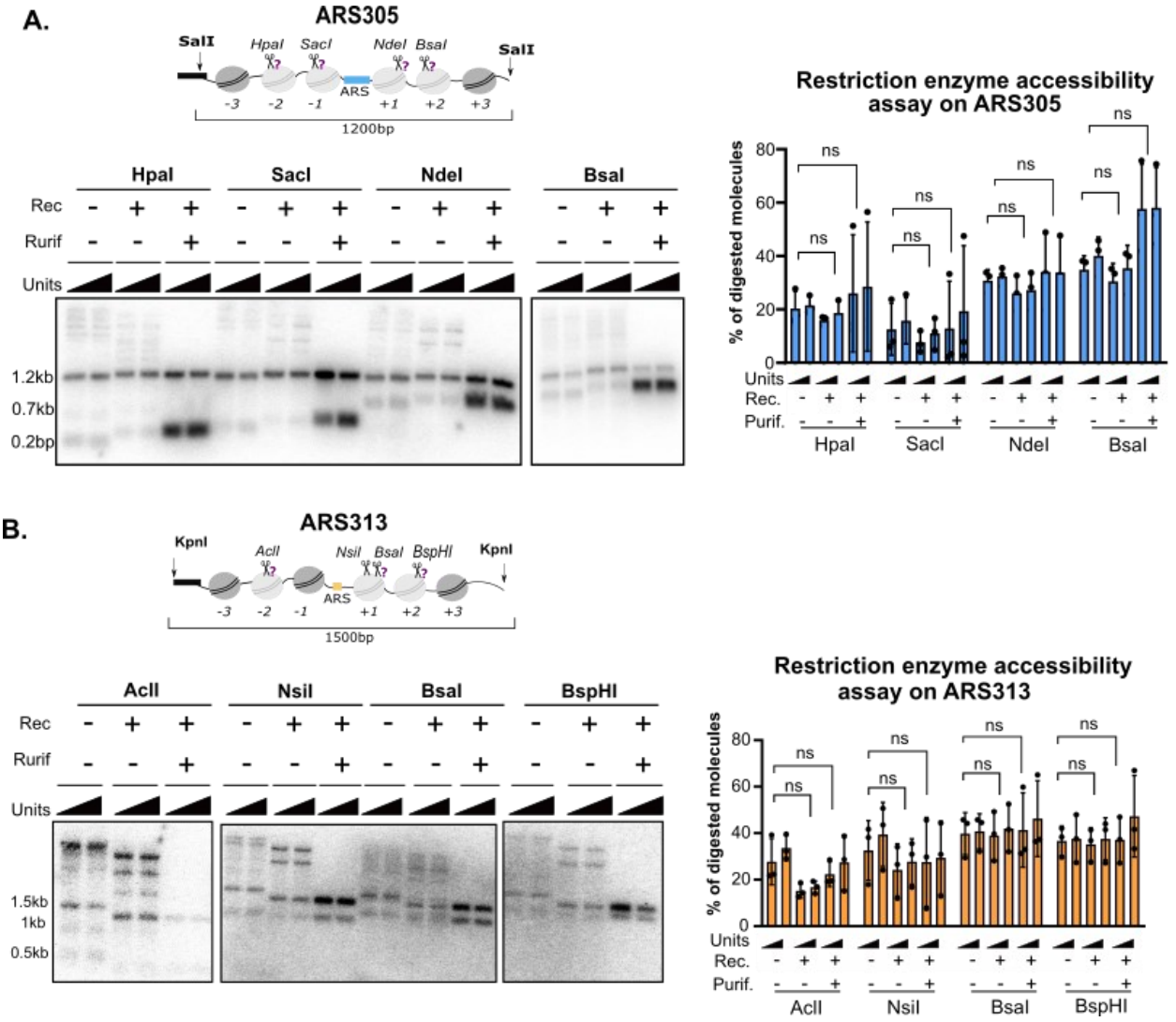


Figure 34. A-B. Comparative REA analysis in chromosomal, recombined and purified ARS305 and ARS313 loci. Nuclei (before and after recombination) and chromatin rings (after purification) from yeast strains Y65 (ARS305), Y94 (ARS313) were isolated and digested with increasing amounts of the indicated restriction enzymes (scissors). DNA was isolated, digested with SalI in ARS305, KpnI in ARS313 and subjected to indirect end-labeling Southern blot analysis with specific radioactively labeled probes. Top of each origin: schematic representation of each ARS locus with restriction sites used to probe chromatin structure (black rectangles) and restriction sites for isolation of DNA fragments for subsequent Southern blot analysis. The histogram shows the results of Southern blot quantification as a percentage of digested chromatin locus. Mean and standard deviations of all plots are from n= 3 biological replicates (ns, indicates no significant statistical difference P > 0.05, unpaired t-test).

6.5.3 Recombination and purification do not disrupt the loading of MCM2-7 complex

Besides nucleosomes, one of the major structural components of licensed replication origins is the pre-replicative complex consisting of a double hexamer of the MCM2-7 complex. To test whether MCM2-7 complex remains associated with the replication origin chromatin after recombination and purification, I established a yeast strain expressing Mcm2 protein as a fusion protein with an HA epitope tag in the genetic background that allows Mcm2 detection after purification of the ARS315 origin domain. To test expression of Mcm2-Mnase-HA in the resulting yeast strain, I analyzed whole cell lysates by Western blot using an HA antibody. In contrast to the parental ARS315 strain, a specific band at the expected molecular weight of Mcm2-Mnase-HA at 121kDa (upper red rectangle) appeared in the Mcm2-Mnase-HA strain (**Figure 35**).

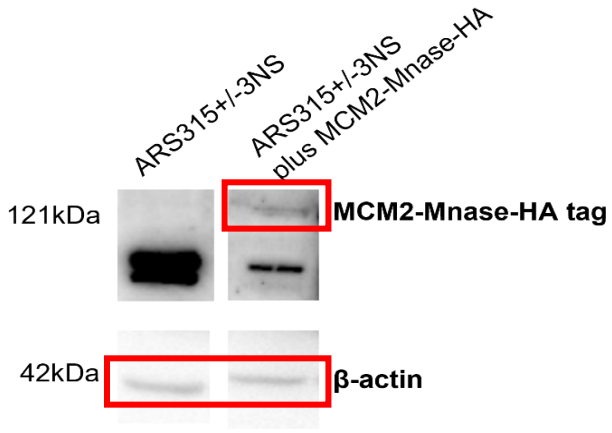


Figure 35. Western blot analysis showing the expression of Mcm2-MNase-HA tag with 121 kDa molecular weight in the recombined ARS315+/-3 NS yeast strain compared to negative control. The modified yeast strain expressing Mcm2-Mnase-HA shows overall lower protein expression level which could indicate an effect of the MNase-HA tag on protein or complex stability. β-acting was used as loading control. Note that the shown gel lanes are derived from different membranes and exposure times that explain the slight shift in the position of the β-actin bands.

In this strain background, the ARS315 origin was affinity-purified according to the established single-locus purification protocol. qPCR results showed similar recovery, yield and enrichment of ARS315 loci compared to the unrelated *PDC1* gene locus (**Figure 36**) as previously observed in the parental ARS315 strain (**Figure 33B**).

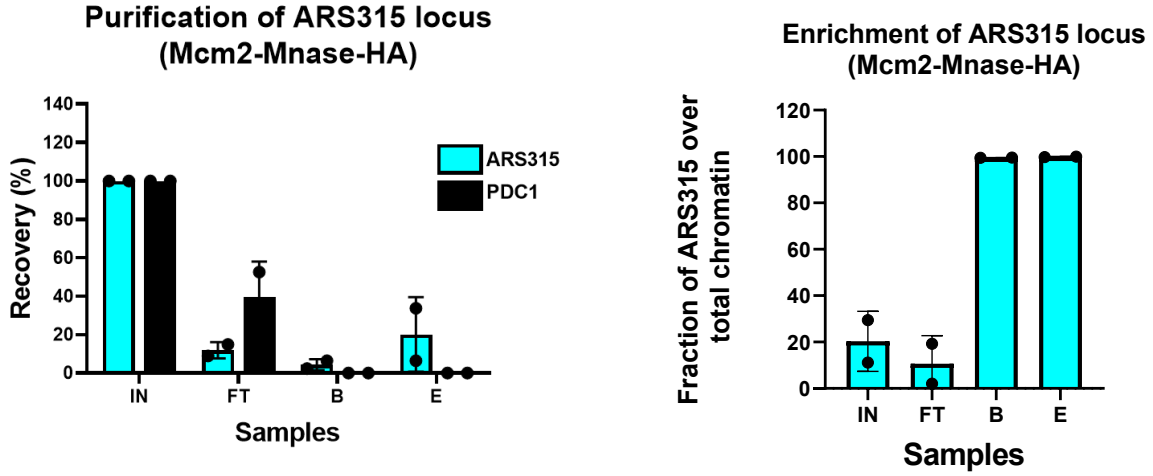


Figure 36. (Left panel) LexA affinity purification was performed for ARS315 strain where the single-copy locus is targeted by RS and Mcm2 protein is fused to Mnase and HA epitope. *PDC1* gene primers were used as a control for an unrelated genomic locus. DNA samples were taken (0.1 % Input (IN), Flowthrough (FT), Beads (B), Eluate (E) from n=2 biological replicates and analyzed by qPCR. (Right panel) Enrichment of ARS315 locus in the final elute. Given that the size of yeast genome is 12,000 kb, and the length of a chromatin ring is ~1 kb, the fold enrichment ratio of the specific origin to *PDC1* was used to define the enrichment of an origin in the DNA samples.

To test whether Mcm2 remains associated with the purified chromatin rings after recombination and purification, I subjected the protein samples to Western blot analysis using an HA antibody. Protein samples were obtained at each individual step. In this analysis, it is important to note that both strains also express the LexA-TAP protein with a molecular weight of 85 kDa (**Figure 37,**

lower red rectangle). The protein A moiety will therefore also interact with the HA primary and secondary antibodies and therefore explain the appearance of the LexA-TAP protein on the blots. Besides a strong band representing the enrichment of LexA-TAP protein in the final eluate (85kDa), an additional higher molecular weight band at 121kDa was observed in all fractions of the purification (**Figure 37,** upper red rectangle). Although unspecific binding of the Mcm2-Mnase-HA protein cannot be completely excluded, the fact that this band was also clearly visible in the final eluate supported the notion that Mcm2 was still bound to the purified replication origin chromatin domain. Together, these results suggest that the purification system allows the isolation of individual replication origins in their native chromatin context and major structural components such as the nucleosomal core particles and major replicative complexes such as MCM2-7 complexes are *bona fide* components of the purified origin domains.

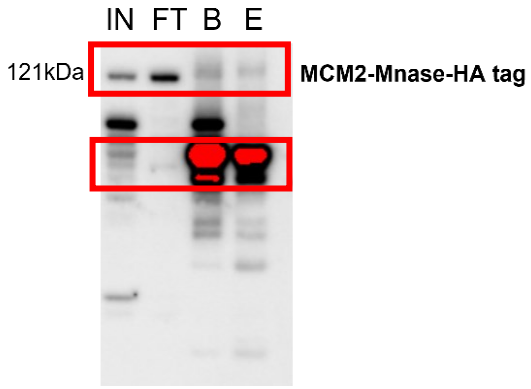


Figure 37. Western blot analysis of Mcm2-Mnase-HA tagged protein during the ARS315 purification.

6.6 Single-molecule DNA methylation footprinting analysis of DNA replication origins

6.6.1 Establishment of MATAC-seq as a tool to analyze single-molecule chromatin accessibility

Traditional methods to study chromatin structure and accessibility have been designed as bulk methods and provided great insights into the chromatin structure, epigenetic landscape as well as spatial organization of chromosomes. However, it is important to realize that the majority of these genomic methods describe the properties of enormous numbers of molecules, averaging the measured parameters over a large population of molecules and cells. Most of these assays also enrich for a single feature of interest such as transcription factor binding or chromatin marks. Thus, the behavior of individual molecules with distinct conformations and potential co-occurrence of multiple binding events cannot be observed over time. Therefore, single-molecule methods

have the great advantage to detect the structural differences of individual chromatin fibers and uncover intermolecular heterogeneities, particularly important in the context of replication origin chromatin.

To determine the local nucleosome arrangement at ARS sites, I took advantage of the Oxford Nanopore sequencing platform, where the different ionic current of DNA bases passing through a biological pore allows discrimination of the different DNA bases (Kasianowicz et al., 1996). Apart from the canonical four DNA bases, different ionic current is also observed for methylated vs. non-methylated bases. Thus, *in situ* treatment of nuclei or isolated chromatin with DNA methyltransferases in combination with Nanopore-sequencing can provide insights into chromatin accessibility by DNA methylation footprinting analysis. Initial approaches were limited to commercially available cytosine methyltransferases and therefore restricted the resolution of this assay to the availability of CpG or GpC residues (Miranda et al., 2010). In 2020, Shipony et al improved the resolution of this method by using a combination of three different methyltransferases M.CviPI (GpC-5mC), M.SssI (CpG-5mC), and EcoGII (A at any sequence context-m6A) and establishing the **S**ingle-**M**olecule long-read **A**ccessible **C**hromatin mapping sequencing (SMAC-seq) assay (Shipony et al., 2020). Methyltransferases have the ability to methylate only the non-nucleosomal, open regions and, hence, after basecalling provide an imprint of open (methylated regions) vs protected (non-methylated regions). To date, this approach has only been performed on whole genome samples, thereby limiting the number of high-quality reads per single gene locus. Thus, determining the frequency distribution of alternative and rare chromatin states at a given locus remained a challenge. During my PhD project, I managed to advance the SMAC-seq approach and overcome this limitation by establishing the Methylation Accessibility of a TArgeted Chromatin domain Sequencing (MATAC-seq), a single molecule approach to profile the chromatin structure of a targeted locus of interest at near basepair resolution (**Figure 38**).

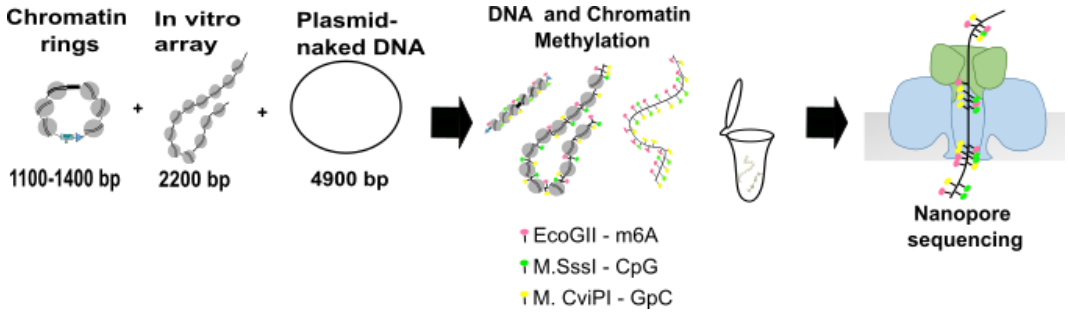


Figure 38. Experimental outline of single molecule Methylation Accessibility of TArgeted Chromatin loci sequencing assay (MATAC-seq). The purified chromatin rings and two spike-in controls; an *in vitro* assembled nucleosomal array and a naked plasmid were treated with m6A, CpG and GpC methyltransferases, which preferentially methylate accessible DNA bases. DNA was isolated and the circular molecules were linearized with BsrGI prior to nanopore sequencing.

6.6.2 Chromatin ring purification of ARS chromatin domains

In order to analyze the chromatin states of all selected replication origins (ARS305, ARS313, ARS315, ARS316 **Figure 39**) within a single Nanopore sequencing experiment, it was necessary to combine the cell lysates from all recombination strains and purify the origins of interest simultaneously by binding to IgG beads and stringent washing steps followed by elution with TEV cleavage. DNA and protein samples were collected again during the purification process to control the efficiency of the purification for all origins. In the final eluate, I typically recovered a total amount of 150-200 ng, which was likely dominated by the higher recovery of rARS molecules derived from a multicopy gene locus (**Figure 33A**). Even though that the results of qPCR analysis from the different purification fractions suggest that all four single-copy origins were depleted in the flowthrough and enriched on the magnetic beads in contrast to an unrelated housekeeping gene locus (*PDC1*) (**Figure 39A**), the elution from beads by TEV cleavage showed reduced efficiency (5-10 %) in comparison to the previous individual origin chromatin purifications (20-40 %) (**Figure 33**). However, the enrichment of each locus ranges between 60- 100 % of total chromatin, indicating the high purity and low levels of contaminating genomic chromatin in the final elution sample (**Figure 39B**). Thus, the recovered yield and number of chromatin molecules derived from the specific origin domains was expected to be sufficient for the intended chromatin accessibility analysis.

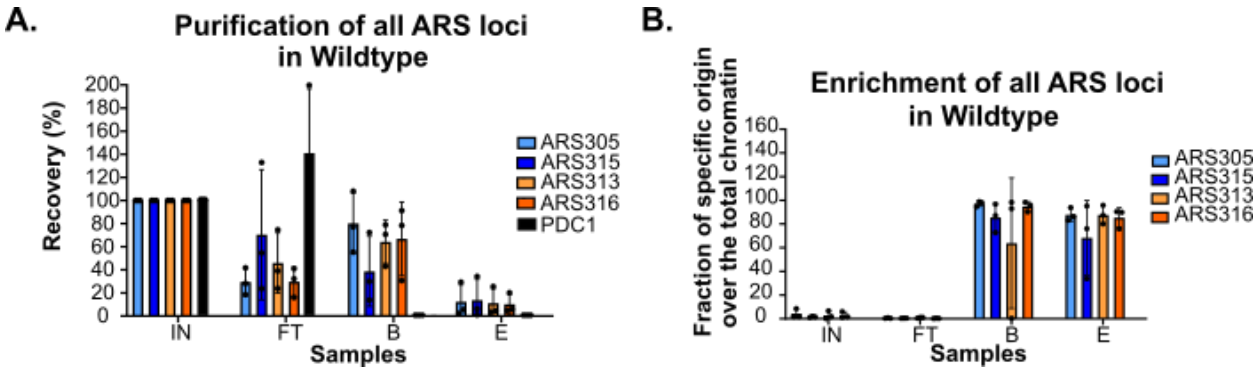


Figure 39. A. LexA affinity purification process was performed for all the WT strains together ((Y65 (ARS305), Y69 (ARS316). Y91 (ARS315), Y94 (ARS313)), which were then subjected to MATAC-Seq assay. DNA samples were taken (0.1 % IN, FT, B, E) from n=2 biological replicates and analyzed by qPCR. The enrichment of an unrelated region (PDC1) was tracked side-by-side with the regions of interest. **B.** The plot shows how much of the total DNA is originated from the chromatin locus of interest. Given that the size of yeast genome is 12,000 kb, and the length of a chromatin ring is ~1 kb, the fold enrichment ratio of the specific origin to PDC1 was used to define the enrichment of an origin in the DNA samples.

Similarly, all origins were purified in separate experiments from two other sets of yeast strains competent for excision of the selected replication origins and contain a genetic deletion of the *ISW2* or *IES6* gene, respectively. qPCR analysis of these CRE mutant purification samples showed similar levels of enrichment and yield for all four replication origins compared to the purification from wildtype cells (**Figure 40A-B**). However, it is important to note that the ARS316 *ies6*Δ and ARS313 *isw2*Δ origins were purified less efficiently (**Figure 40**). In addition, Western blot analysis showed sufficient reduction of LexA protein in the flowthrough (FT) (**Figure 41**) and >50 % TEV cleavage (B) for wildtype as well as CRE mutant origin purifications.

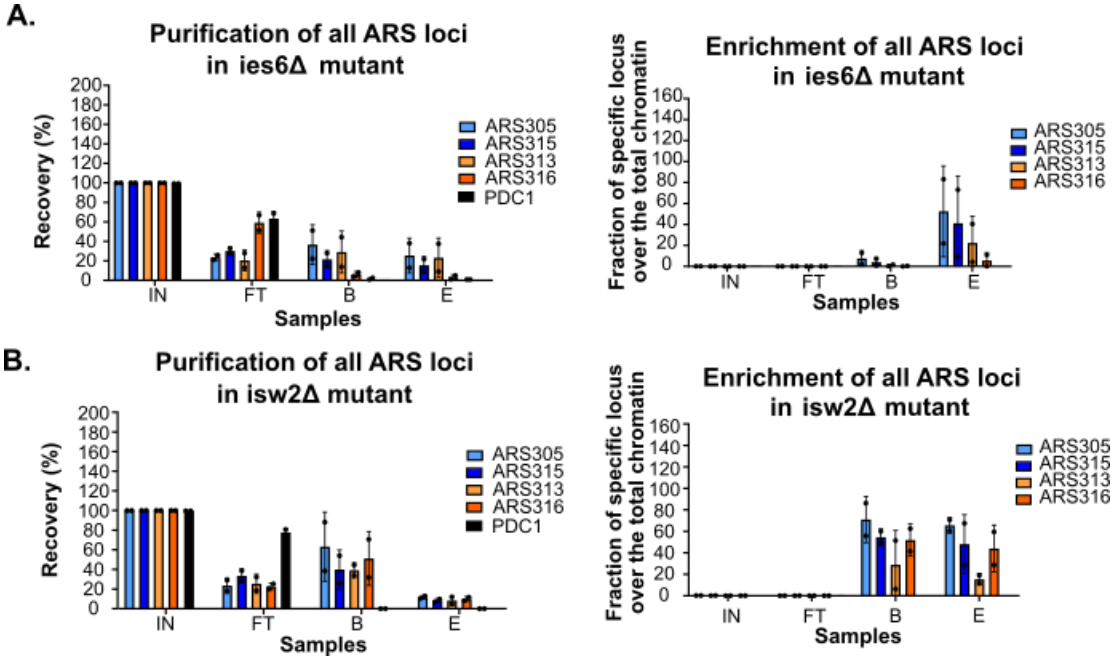


Figure 40. A-B. LexA affinity purification process was performed for all the CRE strains ((isw2Δ: Y105 (ARS305), Y106 (ARS316), Y102 (ARS315), Y107 (ARS313), ies6Δ: Y127 (ARS305), Y129 (ARS316), Y130 (ARS315), Y128 (ARS313)) together, which were then subjected to MATAC-Seq assay. DNA samples were taken (0.1 % IN, FT, B, E) from n=2 biological replicates and analyzed by qPCR. The enrichment of an unrelated region (*PDC1*) was tracked side-by-side with the regions of our interest. The plots on the left show how much of the total DNA is originated from the chromatin locus of interest. Given

that the size of yeast genome is 12,000 kb, and the length of a chromatin ring is ~1 kb, the fold enrichment ratio of the specific origin to PDC1 was used to define the enrichment of an origin in the DNA samples.

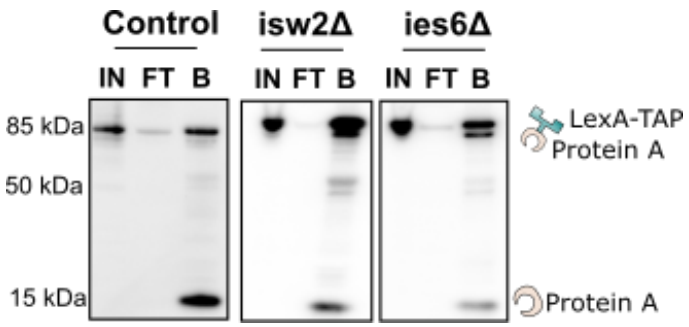


Figure 41. Protein analysis of LexA affinity purification process. The protein samples (0,5 % IN, FT, B, E) were collected in parallel with the previous DNA samples and loaded to an SDS-PAGE gel in order to monitor the LexA protein during the purification by Western blot. Samples were taken from WT ((Y65 (ARS305), Y69 (ARS316). Y91 (ARS315), Y94 (ARS313)), *ies6*Δ ((Y127 (ARS305), Y129 (ARS316), Y130 (ARS315), Y128 (ARS313)) and *isw2*Δ ((Y105 (ARS305), Y106 (ARS316), Y102 (ARS315), Y107 (ARS313)) strains.

6.6.3 Analysis of raw Nanopore data

For the analysis of sequencing data, we used the Megalodon algorithm (Liu et al., 2021), which is able to extract the CpG and m6A methylated bases from the raw nanopore reads. We took equal number of reads of forward and reverse DNA strand and obtained a very high depth of reads for each locus and condition. In order to allow a fair comparison of methylation states across origins and replicates, I considered for the analysis the same number of 1560 high-quality full-length reads, which could be obtained for the large majority of samples and replicates, with the exception of ARS316 *ies6*Δ, for which we recovered fewer reads (548 and 166) per replicate (see an overview in Table 1). This low number of reads is likely caused by the reduced purification efficiency of this origin (**Figure 40A**).

Table 1. The number of reads per replicate and condition that have been used for MATAC-Seq analysis.

	Wildtype		isw2Δ		ies6Δ	
	rep1	**rep2**	**rep1**	**rep2**	**rep1**	**rep2**
ARS305	1430	1560	1560	1560	1560	1560
ARS315	1394	1560	1560	1560	1560	1560
ARS313	984	1560	1560	1560	1560	1560
ARS316	1560	1560	1560	1560	166	548
In vitro array	1000	1560	1560	1560	1560	1560
Naked plasmid	1560	1560	1560	1560	1560	1560

An overview of the enriched ARS replication origins after MATAC-Seq is shown in **Figure 41** which shows robust enrichment of our targeted loci with this method.

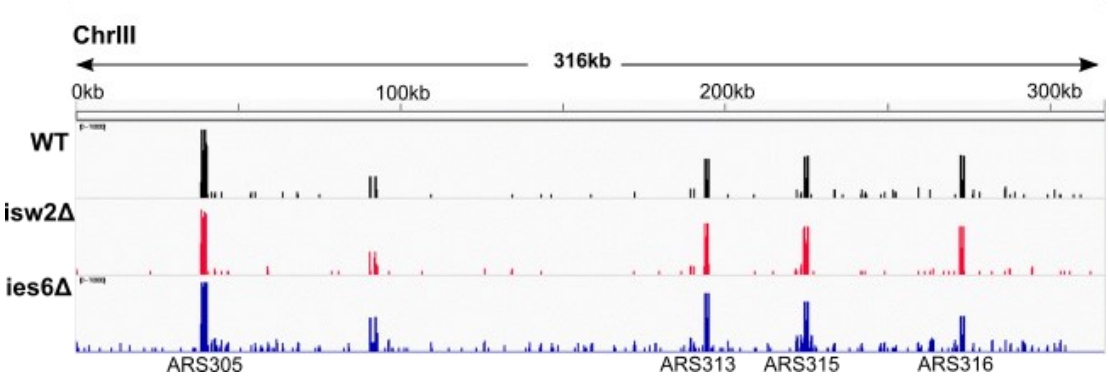


Figure 41. Enrichment of our targeted purified loci on ChrIII after MATAC-seq performance.

To reduce the noise from false positive detection of methylation, we compared the methylation density scores of an unmethylated to fully methylated naked plasmid and set specific thresholds for CpG (Ths = 2) and m6A (Ths = 1) methylation (**Figure 42**). Using this threshold approach, we then generated binary matrices setting as 1 the DNA bases as methylated that exceeded the threshold and as 0 the DNA bases as unmethylated below the thresholds.

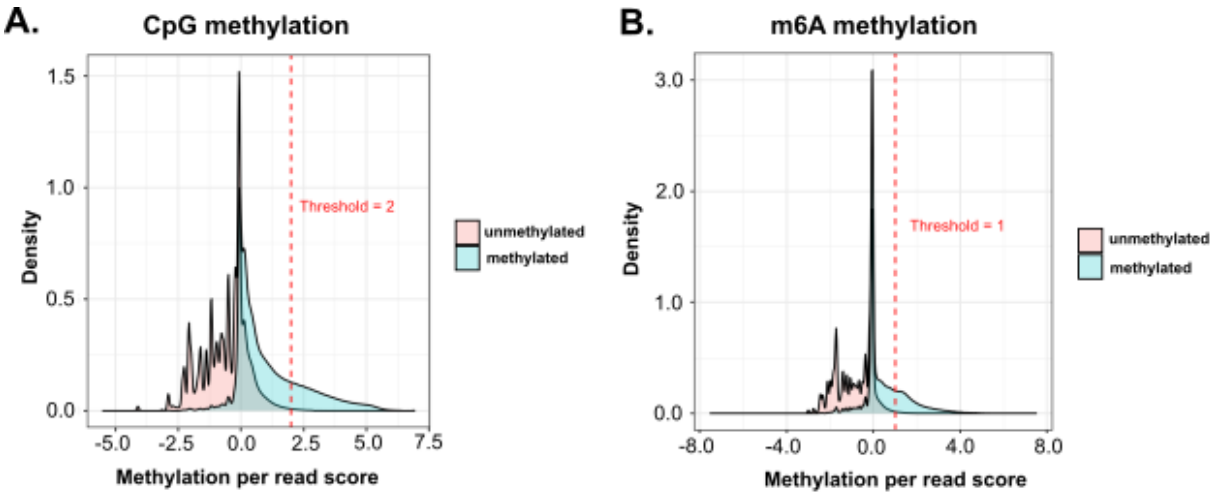


Figure 42. Methylation density comparison between a CpG/m6A methylated and an unmethylated plasmid. The dashed line indicates the threshold that was used to distinguish between signal and noise of methylation on the Nanopore sequencing data.

6.6.4 Validation and benchmarking MATAC-seq with complementary approaches

Given that MATAC-seq represents a new approach to study single-molecule chromatin accessibility, validation of the method and benchmarking with complementary approaches is an essential step. To this end, we followed two lines of investigation:

1. Comparing the MATAC-seq profile of multi-copy ribosomal ARS locus with the psoralen crosslinking and electron microscopy method
2. Comparing the MATAC-seq profile of single-copy replication origin sites with published ChIP-Exo data (Rossi et al., 2021).

6.6.4.1. MATAC-seq and Psoralen crosslinking Electron microscopy

The profile of binarized nanopore data on the *in vitro* assembled nucleosomal array present a repetitive methylation pattern which is consistent with the well-positioned nucleosomes on this synthetic construct. The short linker (10 bp) sites show high level of accessibility in comparison to the strongly protected 601 sequences corresponding to positioned nucleosome core particles. The MATAC-seq single molecule data are derived from both forward and reverse strands that can result in DNA sequence skews, such as an enrichment of methylatable A or C nucleotides on one

of the DNA strands (**Figure 43A**). This explains the slight shift in DNA methylation patterns of the linker regions between the forward and reverse strand reads (**Figure 43B**). In addition, individual molecules show a difference in the total DNA methylation levels as can be seen when arranged from the maximum to the minimum level. Nevertheless, all molecules from the nucleosomal array showed high level of homogeneity with a regular pattern of 12 nucleosomal regions interspersed with accessible DNA linkers for all molecules as expected (**Figure 43B**). Thus, the overall profile proves the sensitivity and accuracy of MATAC-seq regarding the detection of regular nucleosome positioning along the spike-in array (**Figure 43A-B**).

In contrast to the homogeneous profile of the *in vitro* assembled nucleosomal array, the local chromatin landscape around the ribosomal ARS locus reveals high heterogeneity (**Figure 43C-D**). 30% of the molecules have a defined nucleosome free ARS locus which is surrounded by well positioned nucleosomes (**Figure 43C**). Additionally, the RS/*LEXA* sites at the beginning and end of the molecules show high accessibility levels similar to the ARS, confirming the results from bulk Mnase assays (**Figure 29**) and previous studies demonstrating that the LEXA cluster adopts a nucleosome-free region (Hamperl et al., 2014b). The single molecule MATAC-seq analysis reveals different groups of rARS molecules, which either are overall less/more protected or have different level of nucleosome occupancy at the left or right side of ARS locus (**Figure 43D**).

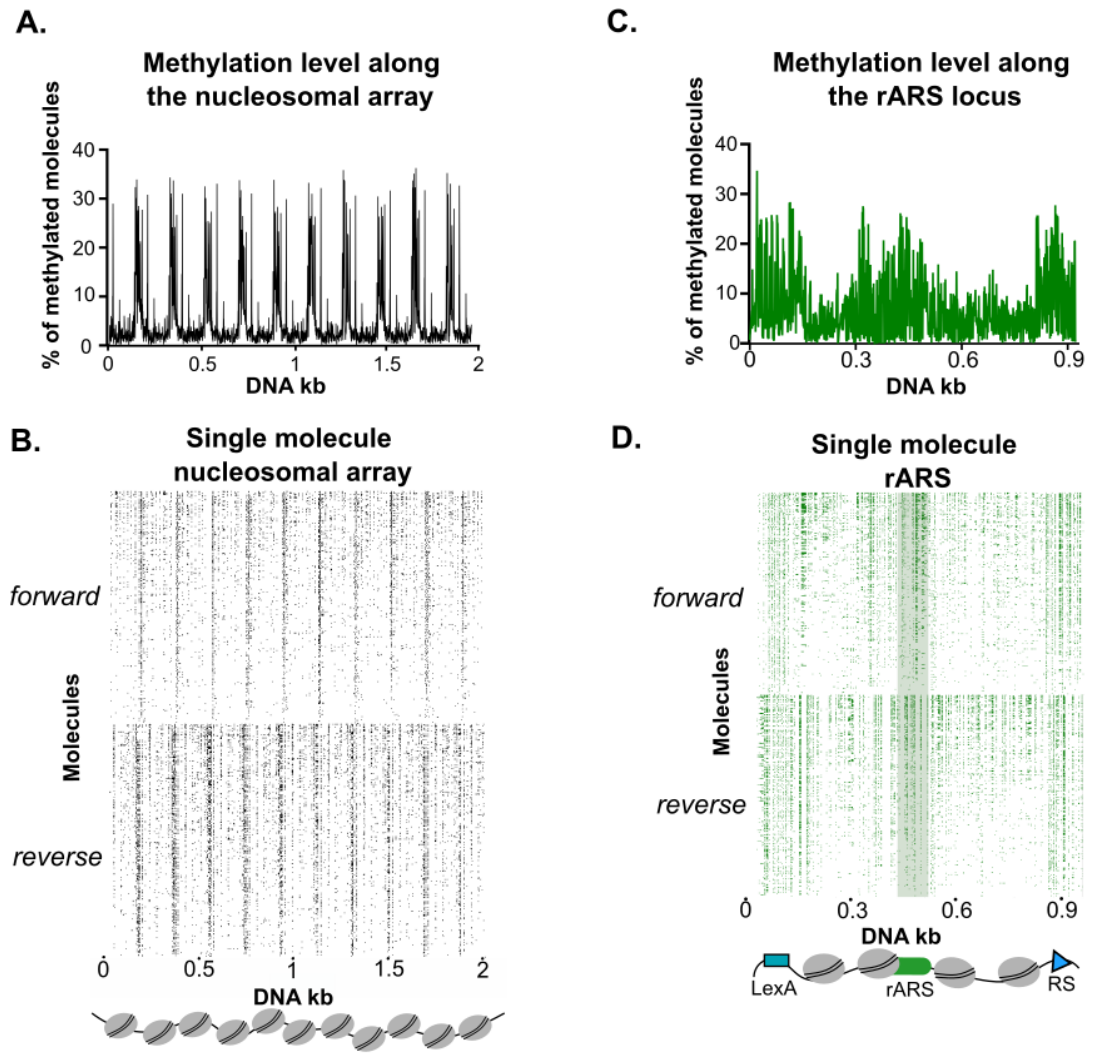


Figure 43 A. Average methylation profile of *in vitro* assembled nucleosomal array. MATAC-seq reveals the repetitive pattern of protected nucleosomal regions separated by narrow hyper-methylated linker sites. **B.** MATAC-Seq reads of the *in vitro* assembled nucleosomal array. Each row represents a single molecule displaying methylation events at near basepair (bp) resolution. The reads have been ranked according to their methylation level (higher on the top and lower on the bottom) on both DNA strands (forward and reverse). **C.** Average methylation profile on the native rARS locus reveals high accessibility at the centrally positioned ARS region. **D.** MATAC-Seq reads of the multi-copy rARS (Y81) chromatin locus. Each row represents a single molecule displaying methylation events at near bp resolution. The reads have been ranked according to their methylation level (higher on the top and lower on the bottom) on both DNA strands (forward and reverse). Identical number of reads were used per DNA strand and per sample (1560 reads) in MATAC-seq plots.

To ensure that MATAC-seq captures not only the highly repetitive profile derived from nucleosomal arrays but also more complex chromatin landscapes such as the ribosomal ARS are

accurately captured, we compared our rARS MATAC-seq results with an independent approach using psoralen crosslinking and electron microscopy, considered as a gold standard approach to analyze single-molecule nucleosome configurations (Brown et al., 2013). As discussed in the introduction, psoralen has the ability to intercalate into double-stranded DNA. After irradiation with ultraviolet (UV)-light at 366 nm, psoralen creates covalent crosslinks between pyrimidines of opposite DNA strands (Cimino et al., 1985). The crosslinking occurs in linker DNA, whereas the nucleosomal DNA is protected, which allows to distinguish whether a DNA region had been occupied by a nucleosome or not. The nucleosomal regions are depicted as single-strand DNA bubble connected by double-strand linker DNA. Initial attempts to investigate the nucleosome positioning of our single copy loci replication origins of ChrIII under EM were not successful. The reason for this was likely the low amount of origin-derived DNA from the single-copy gene loci (10 -20 ng). Even though we were able to detect these circular DNA molecules with simple native spreading (**Figure 44**), the denaturation and extensive washing steps as necessary parts of the psoralen-crosslinking protocol resulted in low recovery rates and no DNA molecules could be visualized on the EM grids under denaturing conditions required for psoralen crosslinking. The multi-copy ribosomal ARS purification instead provided at least 10-fold more DNA molecules (200-300 ng of total DNA), which allowed me to visualize rARS molecules in EM after psoralen crosslinking and denaturing spreading (**Figure 46A**).

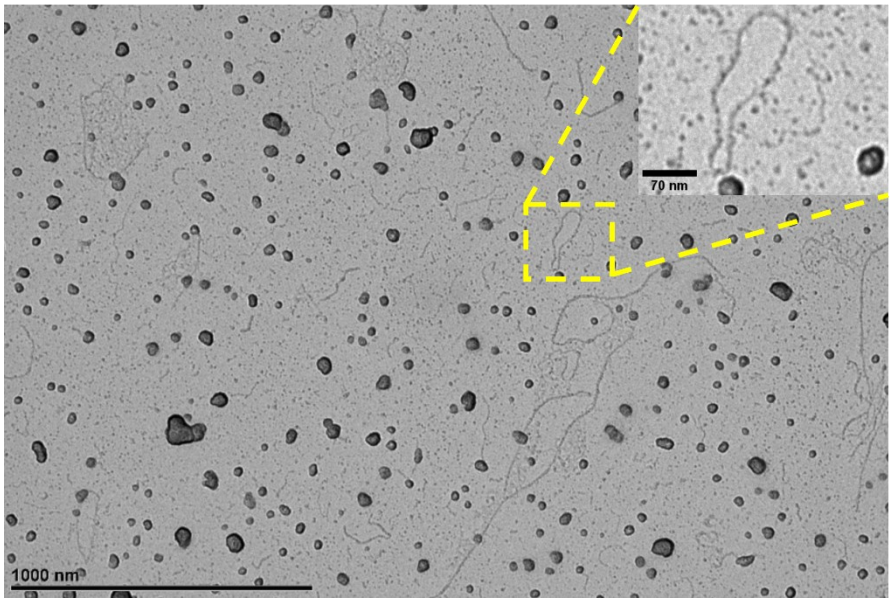


Figure 44. Detection of the single copy locus ARS305 (Y65) under electron microscope after native spreading.

The circular rARS chromatin rings were digested with NcoI, which recognizes a site in the nucleosome-free LEXA region. In this way, the NFR of the rARS locus is located in the middle region of the linearized molecules. The expected size of the rARS fragments is 1046 bp (**Figure 45**).

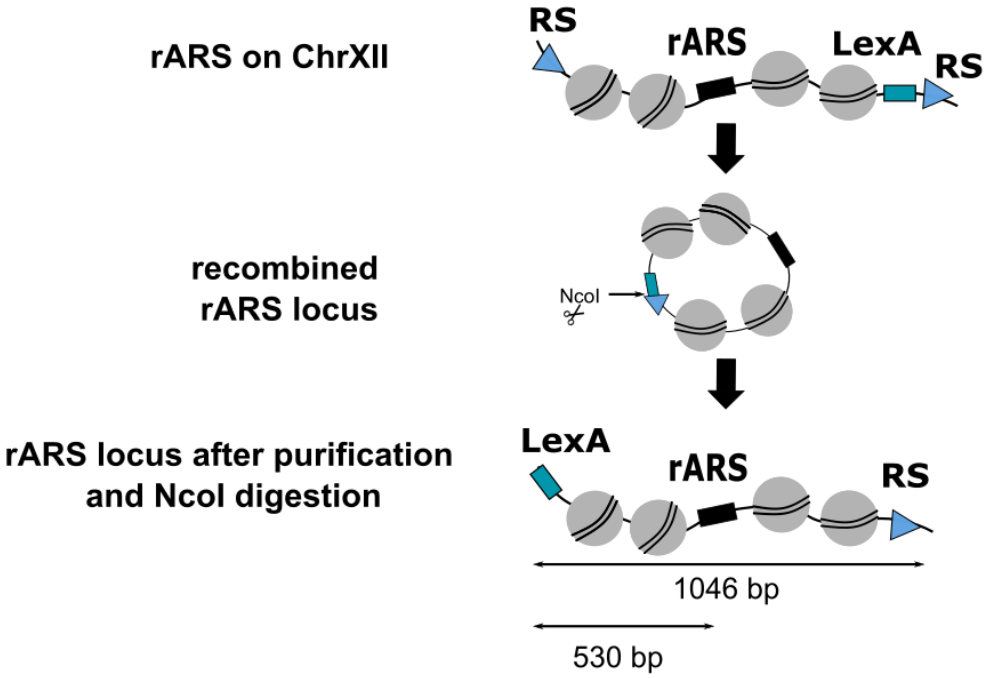


Figure 45. Schematic representation of rARS recombination and digestion by NcoI. The enzyme digests the molecules at a unique RE site located between LexA and RS elements. The linearized rARS molecules are then subjected to psoralen crosslinking and denaturation process.

For the analysis, 143 molecules could be detected that strictly conferred to the expected size range of +/- 10% (940 – 1160 bp) (**Figure 46B**). For each molecule, the relative position, number and size of the individual nucleosomal bubbles was measured. The majority of the detected bubbles (70 %) showed the expected nucleosomal size (135 -165 bp) (**Figure 46C**), suggesting that the majority of detected footprints were indeed derived from nucleosomal particles. However, 30 % of the detected bubbles were smaller (<135 bp), intermediate (165 –300 bp) or showed di-nucleosomal sizes (>300 bp), which could be a consequence of incomplete crosslinking of naked

DNA due to sequence preferences for psoralen-crosslinking, protection by chromatin components other than nucleosomes or inefficient linker DNA crosslinking between two nucleosomes.

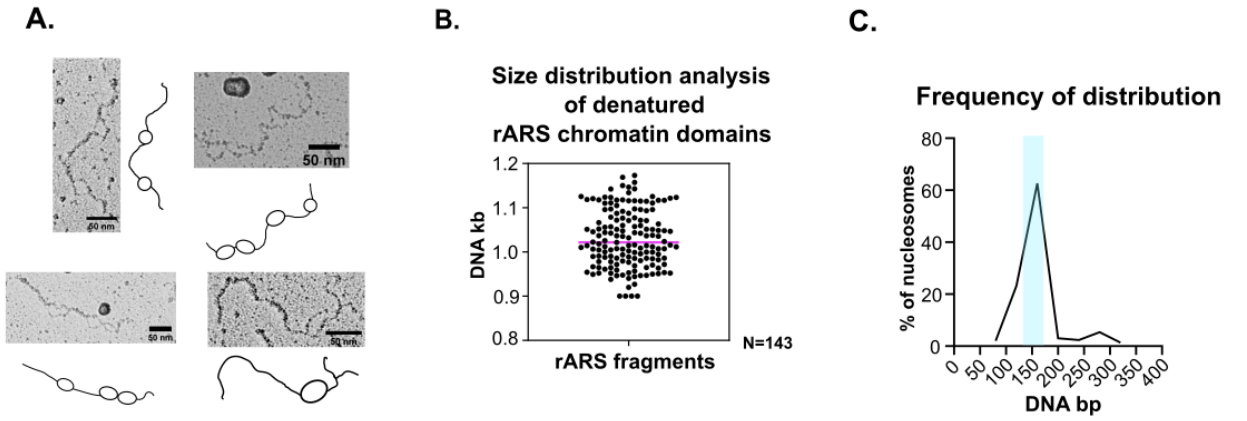


Figure 46. Determination of nucleosome positions at the rARS locus by single molecule EM analysis. **A.** Purified rARS (Y81) chromatin rings were psoralen-crosslinked. After DNA isolation, the rings were digested with NcoI and then subjected to denaturing spreading followed by EM analysis. 143 molecules were analyzed by measuring the size and the number of each nucleosomal bubble. Representative electron micrographs for molecules containing different numbers of nucleosomes are shown on the left. **B.** Size distribution of the purified and denatured rARS fragments as analyzed by electron microscopy. The expected length of the rARS locus is 1046 bp +/-10 % The pink line indicates the median. **C.** The plot shows the size distribution of the nucleosomal bubbles. The expected nucleosomal size is indicated in light blue. For this analysis data from n=2 biological replicates were pooled.

After grouping the rARS molecules according to the number of nucleosomal bubbles, we revealed high level of heterogeneity not only regarding different number of nucleosomes but also regarding their relative positions (**Figure 47A**). Unfortunately, this EM approach does not provide clear orientation of the rARS molecules and therefore, asymmetric distribution for example in the group of molecules with one bubble could not be strictly assigned to the underlying DNA sequence and we had to assign these molecules in a randomized fashion. However, many molecules also showed symmetric distributions as for example in the group with 2 bubbles, where the majority of the molecules reveals a profile where the nucleosome-free rARS is surrounded by nucleosomes on both sides. In contrast, only a small portion of the molecules showed both nucleosomal bubbles next to each other on one side of the ARS. Moreover, even though that the ARS sites are known as nucleosome-free regions, approx. 30 % of the molecules amongst all the groups present a protected ARS site. Overall, these EM data suggest the co-existence of a large variety of

chromatin states which could partially explain the different replication profiles of the ribosomal replication origins, where less than one third of the rARS origins initiate replication at a specific S-phase despite their identical DNA sequence (Fangman et al.,1991).

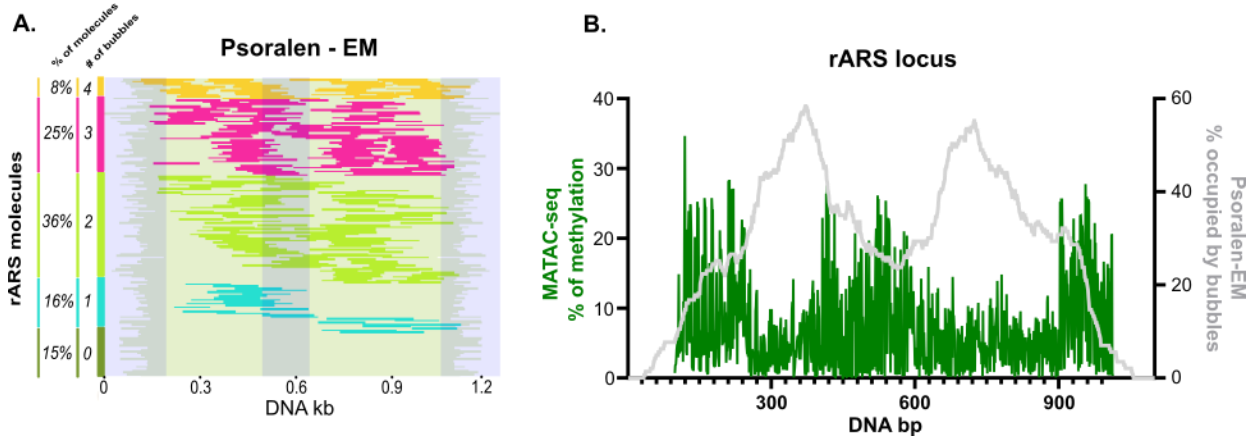


Figure 47. A. Five different groups of rARS (Y81) molecules according to the number of nucleosomes (0, 1, 2, 3 or 4) and its coverage amongst the whole population. Each molecule has been aligned to its symmetric center position. The shadowed areas indicate the middle and the two ends of the rARS molecules. **B.** Averaging nucleosome profile of 143 unoriented rARS molecules. The plot shows the probability of a nucleosome in a specific locus capturing the less protected rARS origin and the two most prominent neighboring nucleosomal sites which is comparable to the averaging methylation profile derived from MATAC-seq.

Despite the fact that EM approach does not orient the rARS molecules, we compared the averaging profile of methylation level produced by MATAC-seq and the averaging nucleosomal profile by psoralen crosslinking method focusing only on the two ends of rARS molecules and the centrally located ARS site as the known fix points within the molecules (**Figure 47B**). Importantly, the two techniques show highly concordant profiles. The highly methylated and accessible regions (ARS locus, beginning and end) are overlapping with low frequency of nucleosome occupancy based on psoralen crosslinking and the intermediate unmethylated inaccessible regions are overlapping with higher frequency of nucleosomal bubbles. Thus, we conclude that MATAC-Seq is robust and provides high-resolution chromatin accessibility maps of specific chromatin domains.

6.6.4.2. MATAC-seq and ChIP-Exo profiles

To further validate the significance of the single-locus origin datasets, we set out to compare the wildtype average methylation patterns generated by MATAC-Seq across all reads with bulk accessibility profiles of recently published ChIP-Exo datasets of canonical histone H3, Orc4, Orc5 and Mcm5 at these loci (Rossi et al., 2021). The methylation profiles of the four replication origins obtained by MATAC-seq revealed that the ARS regions in the center of the molecules were hyper-accessible and strongly overlapped with the ORC and MCM ChIP-Exo peaks (**Figure 48**). Interestingly, the chromatin accessibility of the two LI ARS regions was less accessible in comparison to the two EE origins. Additionally, potential nucleosomal regions surrounding the ARS with decreased methylation coverage match with H3-enriched regions based on the ChIP-Exo profiles. The purified chromatin domain of ARS315 also contained the promoter regions of the *TVS1* gene, which overlapped with a second highly accessible MATAC-seq site (**Figure 48B**), corresponding to the nucleosome-free region of this gene promoter.

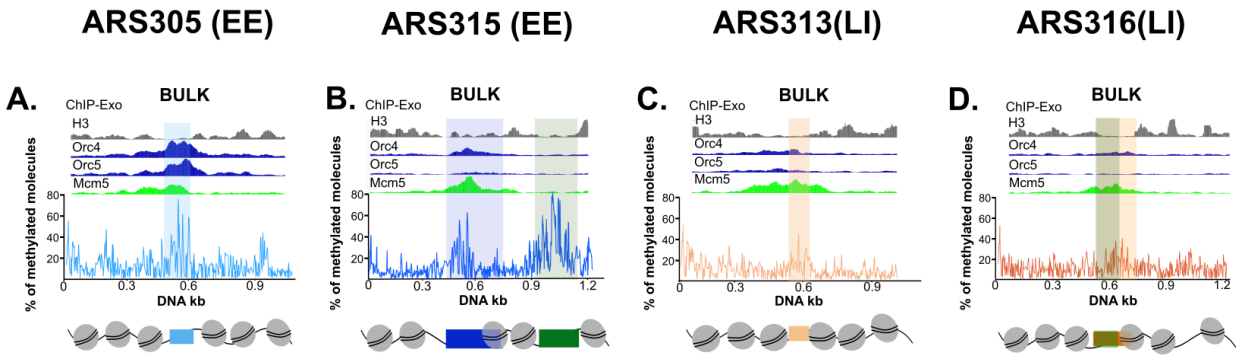


Figure 48. Average methylation profile around the replication origins (ARS305, ARS315, ARS313, ARS316) derived from MATAC-Seq recapitulates the known nucleosomal pattern derived from bulk ChIP-Exo analysis.

Together these data confirm the ability of MATAC-seq approach to reliably identify open chromatin regions.

6.7 Investigation of chromatin accessibility changes between wildtype and CRE mutant strains at targeted replication origins

In order to analyze the WT and CRE mutant datasets for each origin, all reads derived from forward and reverse strands were ranked according to their total methylation level. To this end, all basepairs containing CpG or adenine methylation were counted along the single molecule reads and the molecules with a higher DNA methylation count and thus higher accessibility were positioned on the top, whereas reads with less methylation and thus more chromatin protection were listed on the bottom. This visualization emphasized the high level of heterogeneity existing in one specific chromatin region among a cell population and reveals the co-existence of different chromatin states, which were previously masked by bulk methods or even genome-wide single molecule approaches (Shipony et al., 2020; Stergachis et al., 2020b).

In the resulting four origin maps derived from the control strain (**Figure 49**), strongly positioned nucleosomal footprints are displayed at multiple locations and origins, for example including the +/-3NS and +/-2NS in EE ARS305 (**Figure 49A**) and the +1NS/ +2NS in EE ARS315 (**Figure 49B**). Based on the size of these inaccessible regions (~150-160 bp), it can be inferred that the protection is caused by well-positioned nucleosomes. In addition, the intervening linker DNA showed strong accessibility along the majority of reads (see black arrows), proposing an organized structure with regularly phased nucleosomes. In both EE origins the ARS sites show accessibility in almost all of the single molecules, which is in agreement with the ChIP-Exo data (**Figure 48A-B**). Additionally, the nucleosome-free region of the TVS1 promoter (**Figure 49B**, green rectangle) site in ARS315 origin seems also adopted a uniformly accessible state along the cell population. In contrast to EE origins, the methylation pattern and therefore likely the nucleosomal landscape of the LI ARS313 and LI ARS316 is more diffuse and only a few regions show well positioned nucleosomes clearly separated by linkers, for example +3NS in ARS316 and +3NS in ARS313 (**Figure 49C-D**, orange arrows). Moreover, the accessibility of the ARS sites is not as pronounced as compared to EE origins, which is also in agreement with the ChIP-Exo data (**Figure 48C-D**).

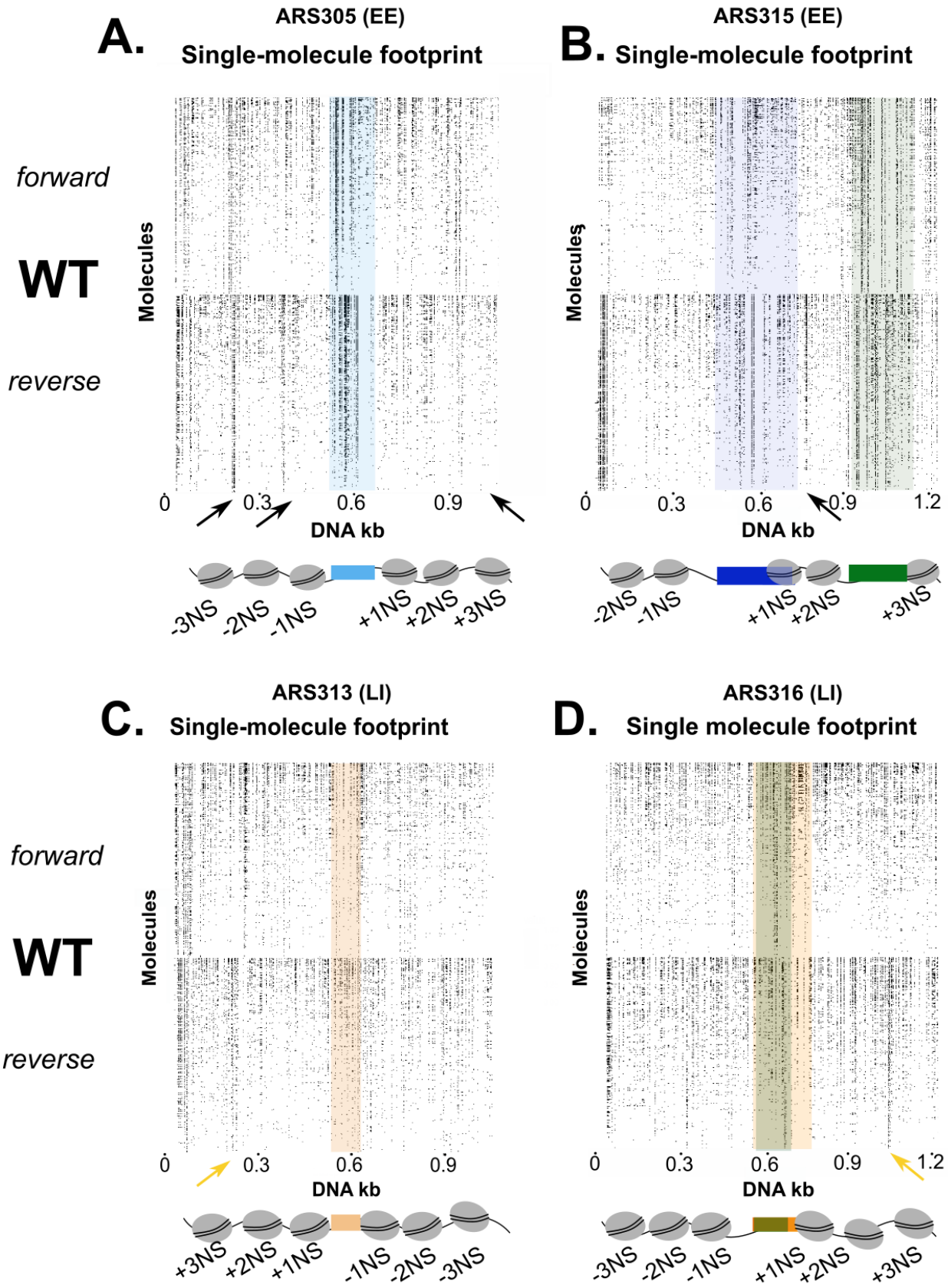
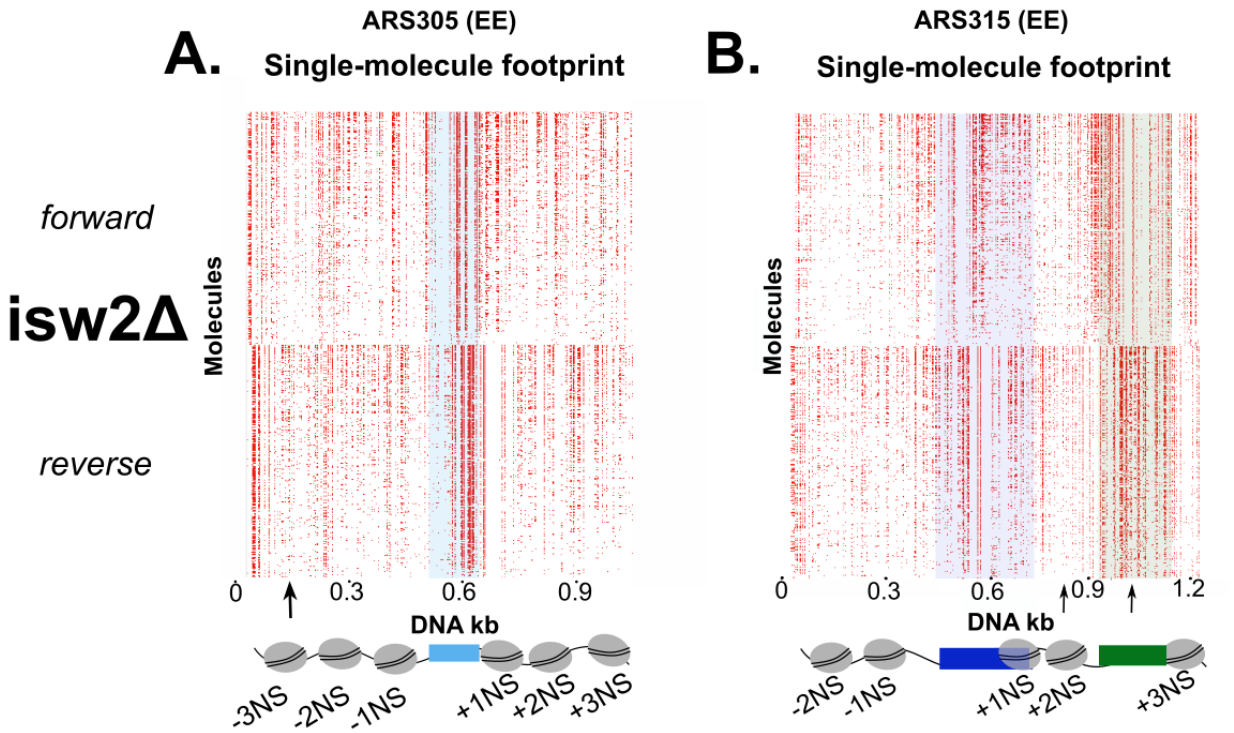

A.
ARS305 (EE)
Single-molecule footprint
forward
WT
reverse
Molecules
0
0.3
0.6
0.9
DNA kb
-3NS
-2NS
-1NS
+1NS
+2NS
+3NS
B.
ARS315 (EE)
Single-molecule footprint
Molecules
0
0.3
0.6
0.9
1.2
DNA kb
-2NS
-1NS
+1NS
+2NS
+3NS
C.
ARS313 (LI)
Single-molecule footprint
forward
WT
reverse
Molecules
0
0.3
0.6
0.9
DNA kb
+3NS
+2NS
+1NS
-1NS
-2NS
-3NS
D.
ARS316 (LI)
Single molecule footprint
Molecules
0
0.3
0.6
0.9
1.2
DNA kb
-3NS
-2NS
-1NS
+1NS
+2NS
+3NS

Figure 49. A-D. MATAC-seq chromatin accessibility maps around the native ARS loci of wildtype (WT) strains ((Y65 (ARS305), Y69 (ARS316). Y91 (ARS315), Y94 (ARS313)) strains. The methylated DNA bases of each molecule are depicted as dots and the reads have been organized according to their methylation level (higher on the top and lower on the bottom) in both forward and reverse DNA strands. Identical number of reads were used for forward and reverse strands in MATAC-seq plots.

Interestingly, the single-molecule maps of both CRE mutants display overall higher methylation level in all four origins (**Figure 50-51**). The nucleosome-free ARS and the promoter regions become even more accessible and several previously well-protected nucleosomal sites show increased methylation signal between the linker DNA (**Figure 50A-B**, black arrows). This effect was most pronounced in the ISW2 mutant (e.g. - 3NS in EE ARS305 and +2NS in ARS315), indicating nucleosome repositioning or eviction after CRE deletion (**Figure 50A-B**).



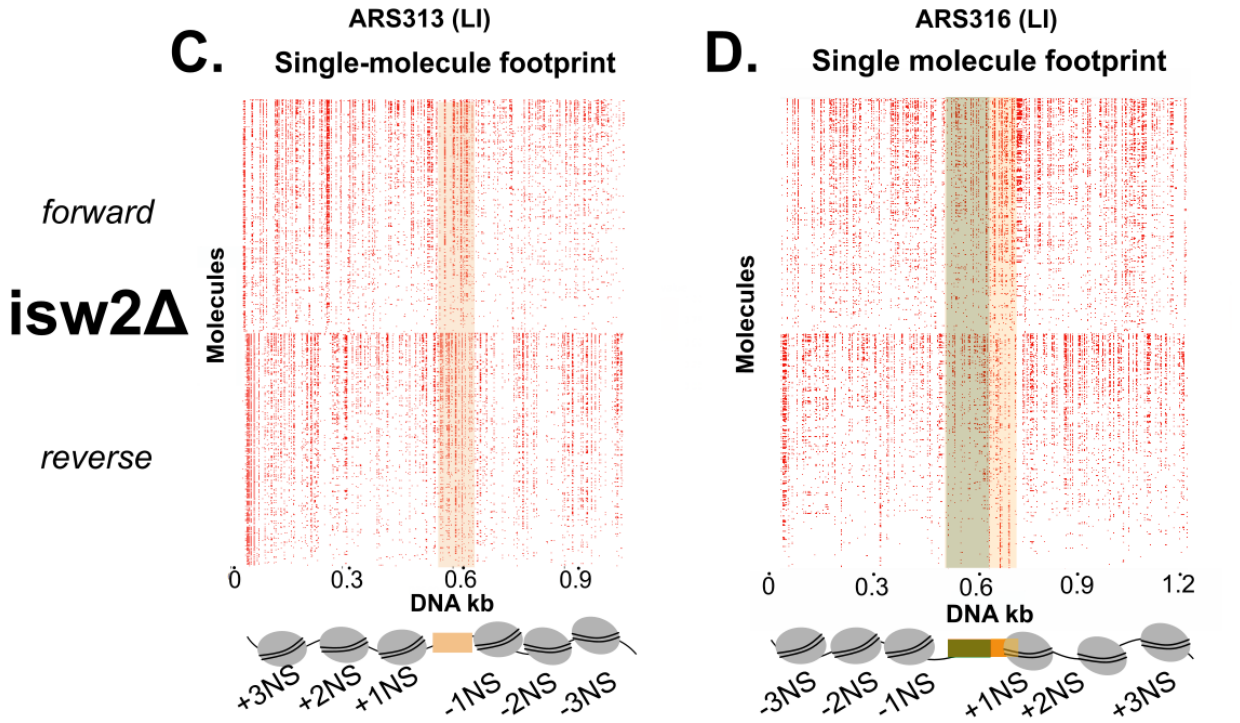
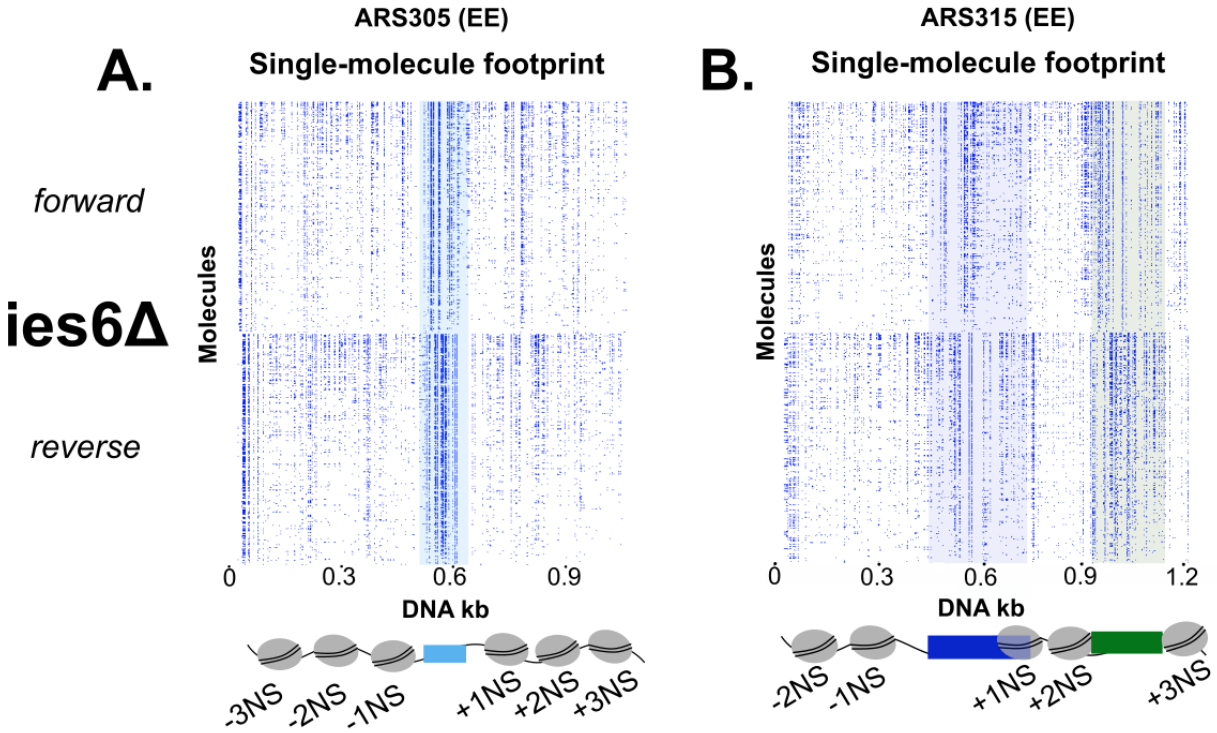


Figure 50. A-D. MATAC-Seq chromatin accessibility maps of the ARS regions in the isw2Δ strains. Identical number of reads were used for forward and reverse strands in MATAC-seq plots ((Y105 (ARS305), Y106 (ARS316), Y102 (ARS315), Y107 (ARS313)) strains.



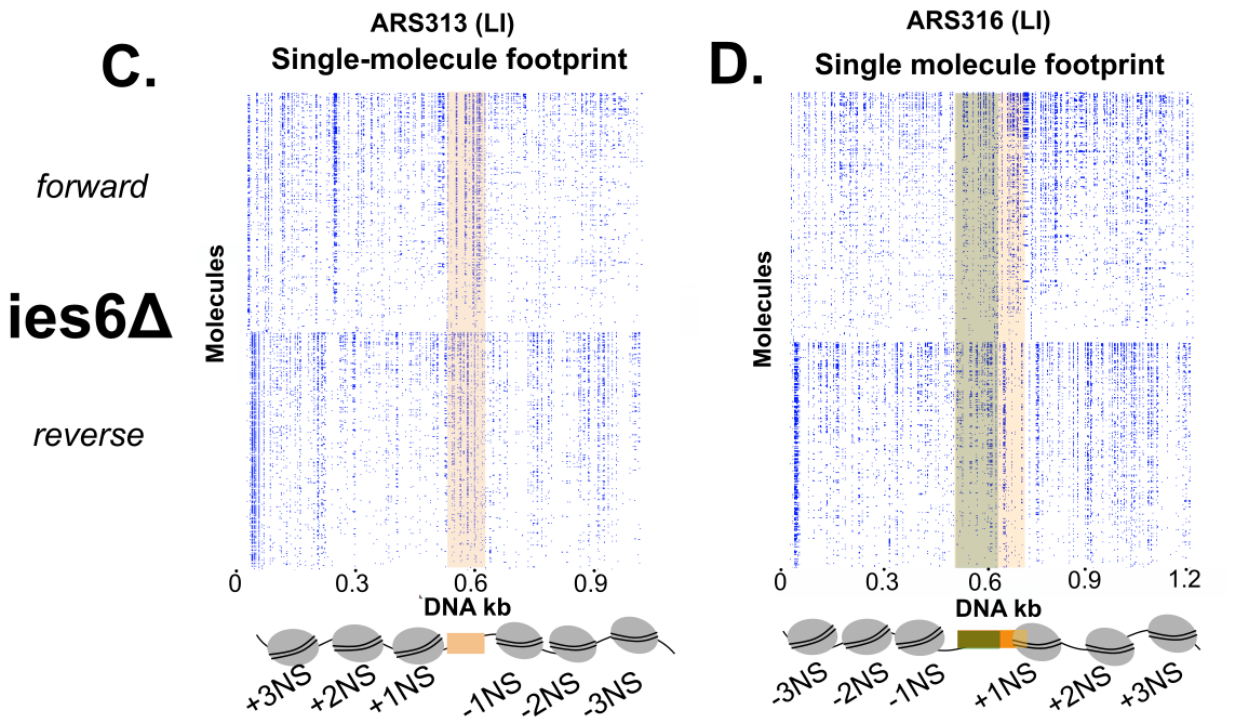
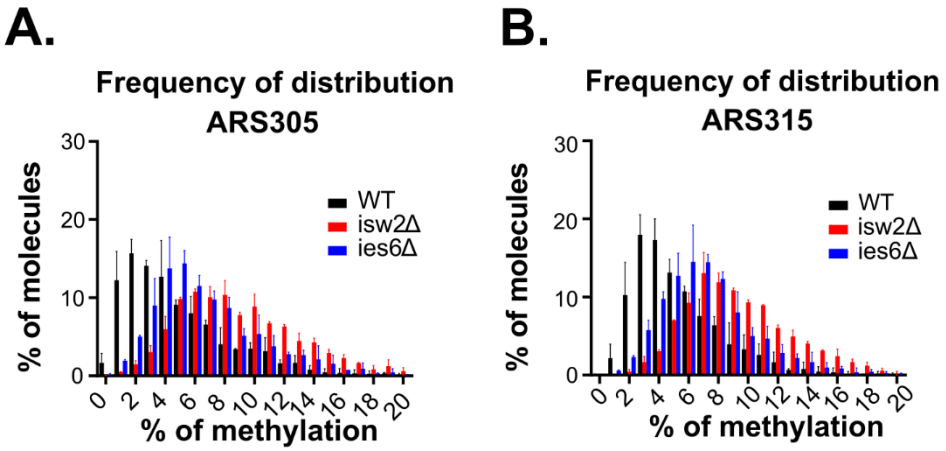


Figure 51. A-D. MATAC-Seq chromatin accessibility maps of the ARS regions in the *ies6*Δ strains ((Y127 (ARS305), Y129 (ARS316), Y130 (ARS315), Y128 (ARS313)). Identical number of reads were used for forward and reverse strands in MATAC-seq plots.

The stronger impact of *ISW2* deletion on nucleosomal occupancy compared to *IES6* deletion is also indicated by comparing the frequency of methylation distribution along the molecules in all four origins (**Figure 52**). The histograms have been generated by measuring the percentage of methylation in each molecule and both CRE mutants show a clear shift to higher methylation levels compared to the distribution observed in WT strains (**Figure 52**).



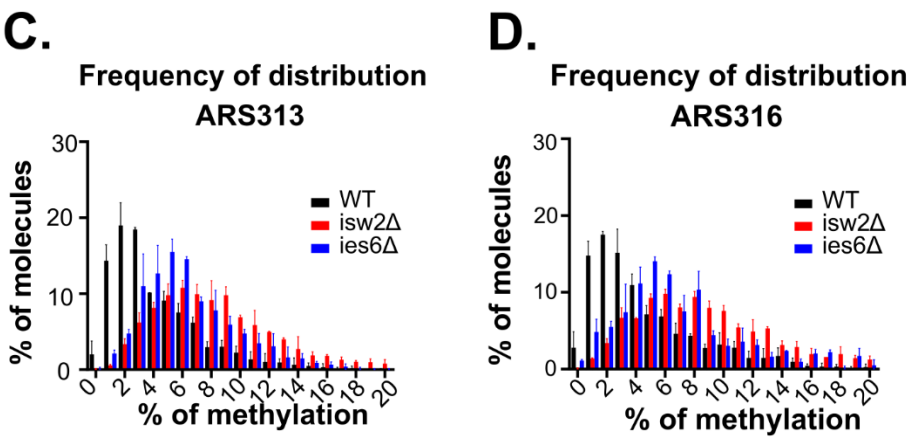


Figure 52. A-D. Frequency of methylation distribution. Comparative analysis of frequency of methylation distribution in all replication origins between WT and CRE mutants shows an overall increase of unprotected regions in CRE mutants. Average and standard deviations are from n= 2 biological replicates.

Overall, we conclude that the single-molecule footprint maps capture not only broad open chromatin regions such as ARS and gene promoter sites, but also the short and accessible linker DNA separating well-positioned nucleosomes are clearly detectable. Additionally, MATAC-seq reveals the increased methylation level in CRE mutants either with global frequency distribution measurements or local changes at individual nucleosomal positions where an increased methylation pattern can be observed along the single-molecule heatmaps.

6.8 MATAC-seq reveals an optimal length of NFR that is correlated with efficient replication origin firing.

Next, I asked whether the observed increased accessibility in the CRE mutants is uniform across the complete ~1 kb region or whether specific subregions of the domain were affected more than others. For this reason, we divided each chromatin locus into shorter intervals that - according to the H3 ChIP-Exo profiles – were likely occupied by nucleosomes (NS), part of linker regions between nucleosomes (L) or part of the accessible ARS or gene promoter sites. Depending on the genomic feature that is analyzed, different lengths for each region had to be chosen and we set the size of an NS region to 80 bp around the center of the ChIP-Exo peaks, the length of linker regions varies from 10 to 40 bp, the ARS sites are 100 bp and the *TVS1* promoter region on ARS315 domain is set to 175 bp (**Figure 53A-D**).

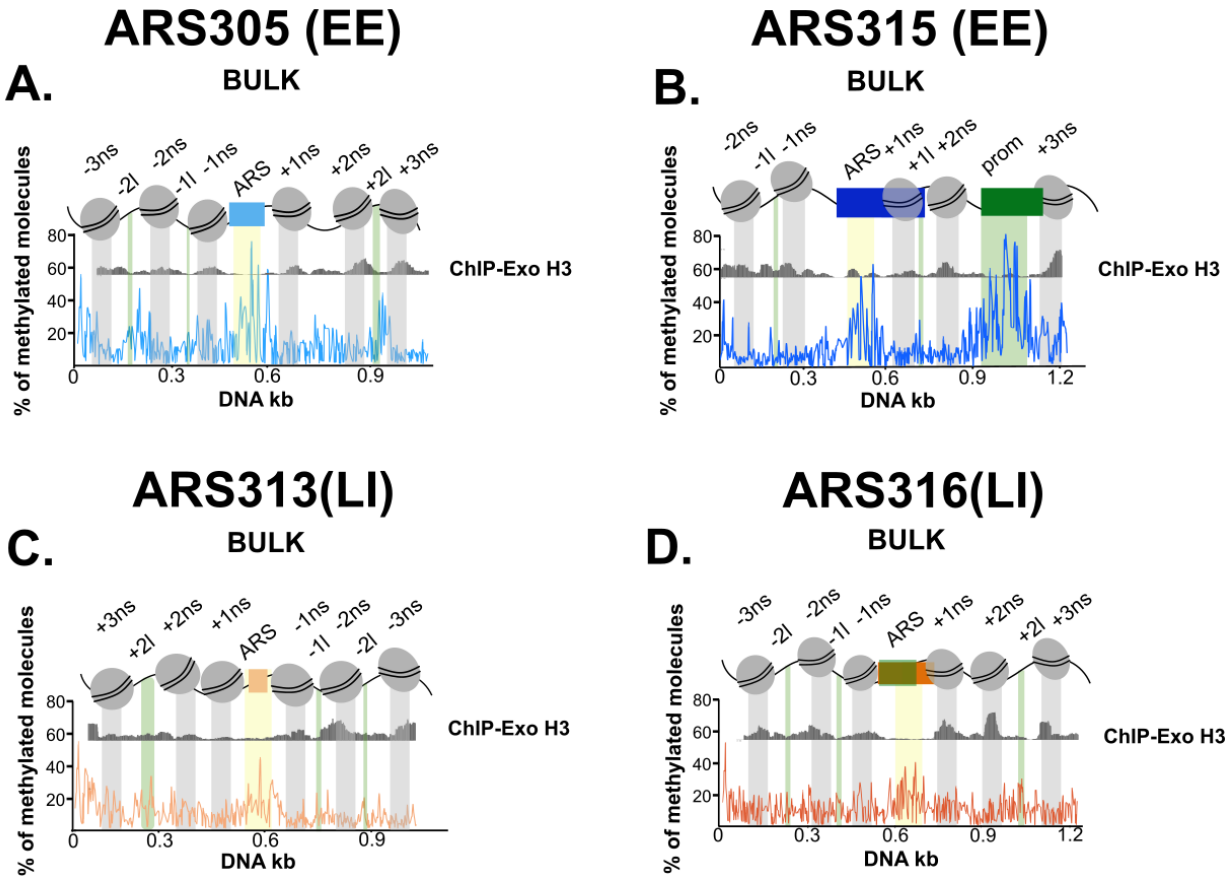


Figure 53. A-D. Schematic representation indicating the specific regions at which each chromatin locus has been further analyzed. Both MATAC-seq and ChIP-Exo profiles have been used in order to define the nucleosomal regions, ARS and linkers. Each ARS site is 100 bp long (yellow) and each nucleosome 80 bp long (grey). The linker length varies from 10b bp to 40 bp depending on the local chromatin accessibility (green). The width of each green bar is scaled according to the linker length. On the top of the graphs the name of each feature is written.

For the final average methylation plots, firstly we calculated the average methylation level of the spike-in *in vitro* nucleosomal array, as internal control of efficiency of DNA methyltranferase activity (**Figure 54A**). Then we calculated the average methylation per base-pair of each chromatin loci of all datasets and divided this by the average value of the respective nucleosomal array of each dataset. Given that the two biological replicates of each condition display similar methylation levels (**Figure 56A-D**), we pooled them together calculating the average of previously normalized methylation levels per base pair, to increase our statistical power.

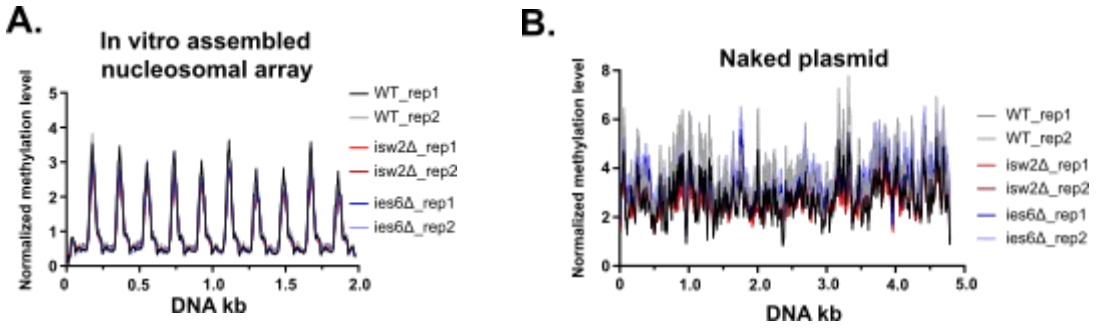


Figure 54. A-B The plots show the methylation pattern on the in vitro nucleosomal array and naked plasmid (K112) (controls) of both biological replicates after normalization to maximum methylation.

However, even though that the naked plasmid is free of nucleosomes and high identical level of methylation would have been expected, a large variety in DNA methylation is observed (**Figure 54B**). These differences could derive from basecalling variability. It is possible that the efficiency of methylation calling is affected by the surrounding DNA content sequence. To avoid comparisons of different DNA sequences, which could result in misleading conclusions, in the later steps we compared the chromatin landscape of each origin in the wildtype condition with the corresponding one of the CRE mutants so that the DNA sequence of the investigated region remains unchanged. No comparison was performed between two different origins that might display such DNA sequence biases.

Additionally, given that the spike-in naked plasmid also contained the ARS305 sequence, this allowed us to directly compare the methylation efficiency of the same DNA sequence in the context of naked DNA versus native chromatin. The chromatinized molecules showed clear protection from methylation except for the central ARS position that displayed similar accessibilities between the naked plasmid and native chromatin context (**Figure 55**).

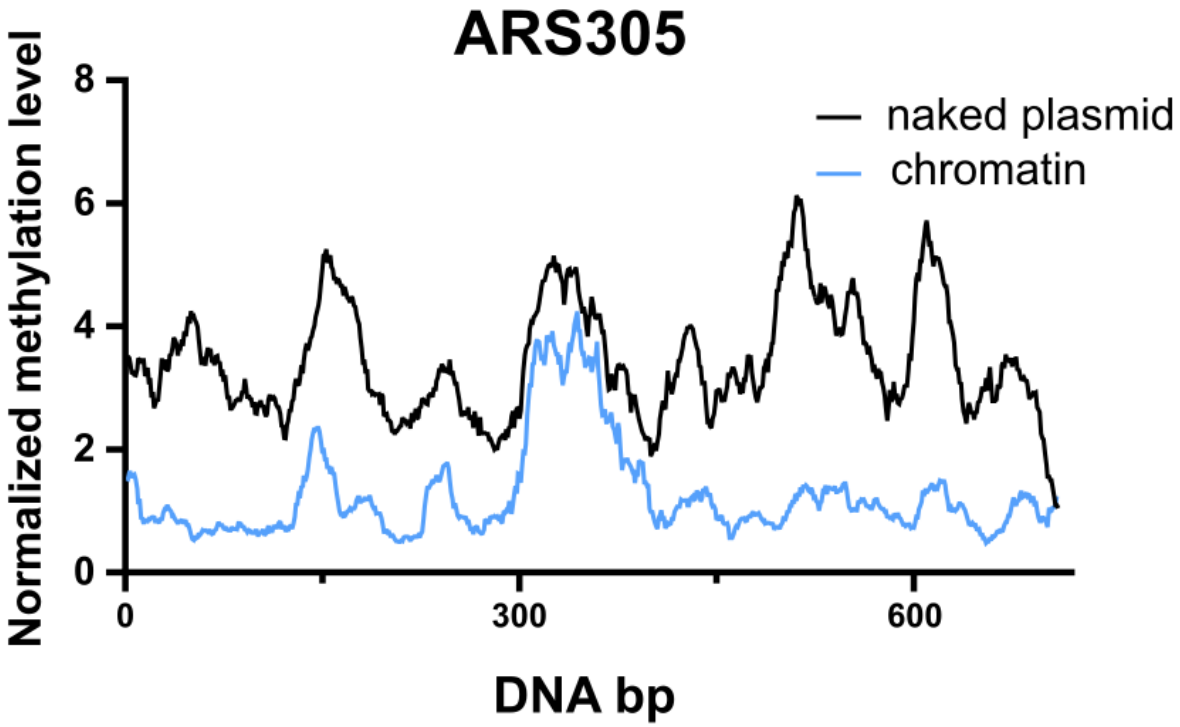
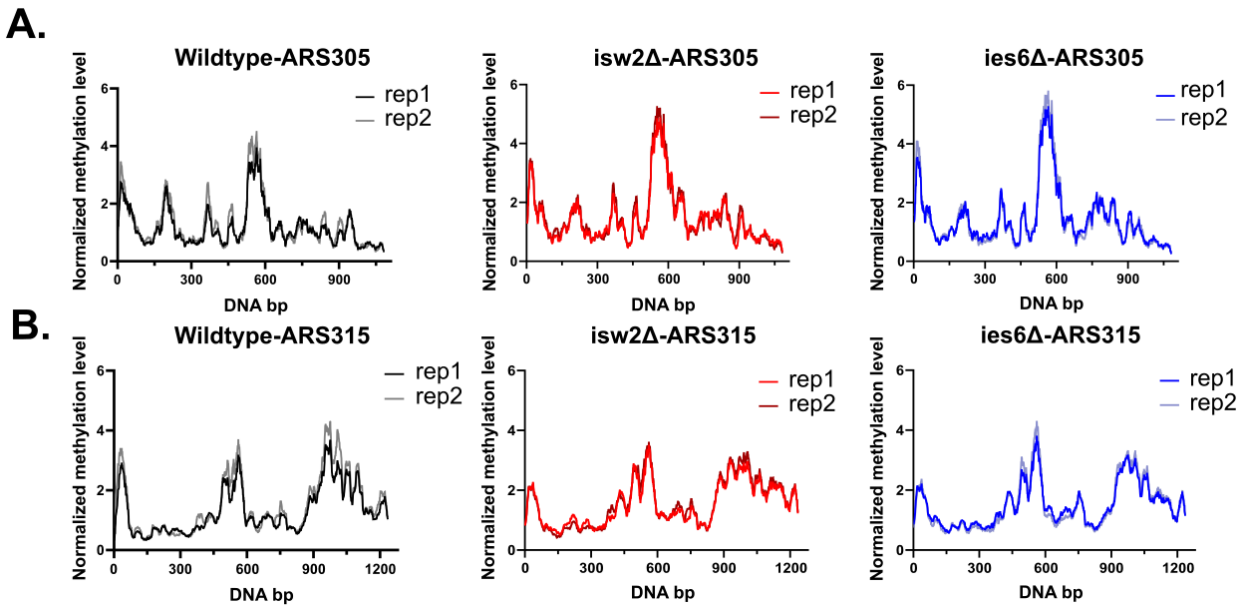


Figure 55. Comparative analysis of the same DNA sequence on ARS305 between naked plasmid (K112) and chromatin (Y65). The unprotected ARS domain of chromatin shows similar level of accessibility with the naked DNA. Their methylation level differs on the surrounding nucleosomal regions.



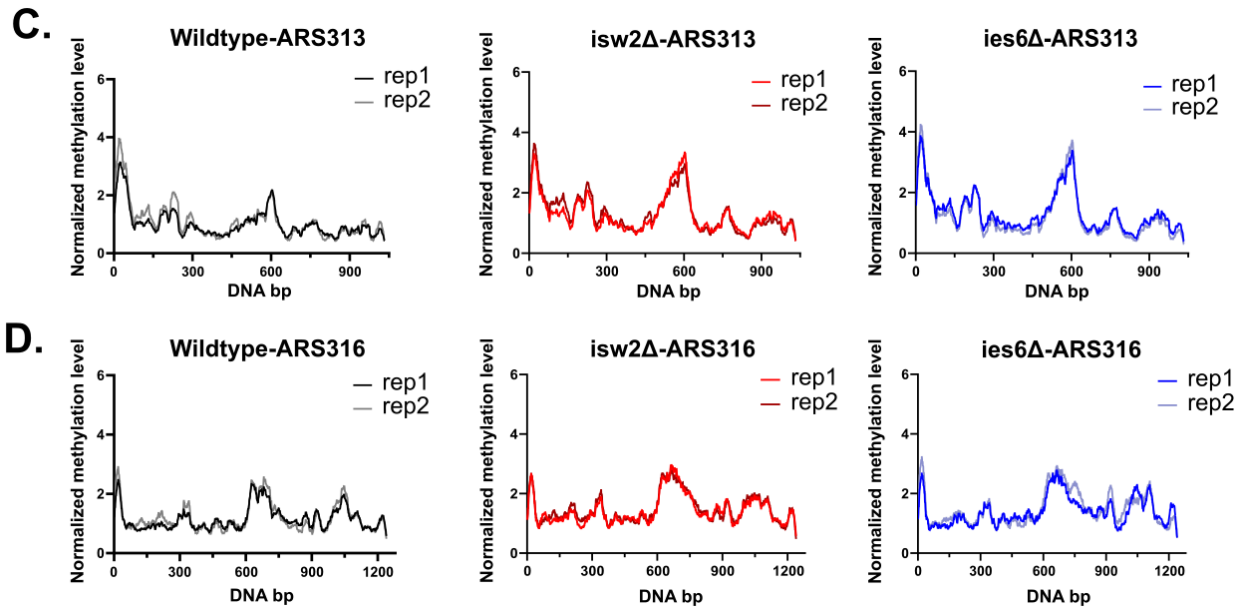


Figure 56. A-D. Methylation pattern on native chromatin domains at ARS loci for both biological replicates. For all the plots in **Figure 49-51**, the methylation level has been normalized and smoothened using a 30 bp window.

Interestingly, CRE mutant strains show consistently increased methylation pattern along the four chromatin loci, indicating that the two mutant strains may have overlapping functions at these genomic regions (**Figure 57A-D**). Comparison with the wildtype methylation footprints reveal striking differences at the ARS sites. where the actual length of the nucleosome-free region has been broadened in all four origins. In order to quantify this, I chose a specific threshold of DNA methylation levels for each origin and then analyzed the length of the ARS NFR in WT, *isw2*Δ and *ies6*Δ samples. Particularly, in EE ARS305 origin the NFR of ARS has been increased from 100 bp to 160 bp for *isw2*Δ and to 152 bp for *ies6*Δ. In both cases, the ARS broadening mainly affected the (+) side of the ARS305 local region leading to repositioning of +1NS further away from the ARS (**Figure 57A & Table 2**). Additionally, increased accessibility in +1L and +2L linker sites is observed indicating changes in nucleosome occupancy of +1NS, +2NS and +3NS. Similarly, the NFR length of EE ARS315 has been increased from 115 bp to 137 bp in both CRE mutants. This, in combination with the broadening of promoter site minimize the available space for +1NS and +2NS resulting in their repositioning or eviction compared to the wildtype strain (**Figure 57B & Table 2**). ISW2 and INO80 deletion also increased the accessibility on the (+) of the previously well protected LI ARS313 and ARS316 replication origins. Interestingly, the NFR of ARS313 has

been broadened from 30 bp to 104 bp in *IES6* deletion and to 125 bp in the *ISW2* deletion (**Figure 57C & Table 2**). In ARS316 the size of NFR changed from 170 bp to 270 bp in *isw2*Δ and 272 bp in *ies6*Δ, resulting to an overall increased methylation level along the (+) side, suggesting nucleosomal sliding or eviction (**Figure 57D & Table 2)**

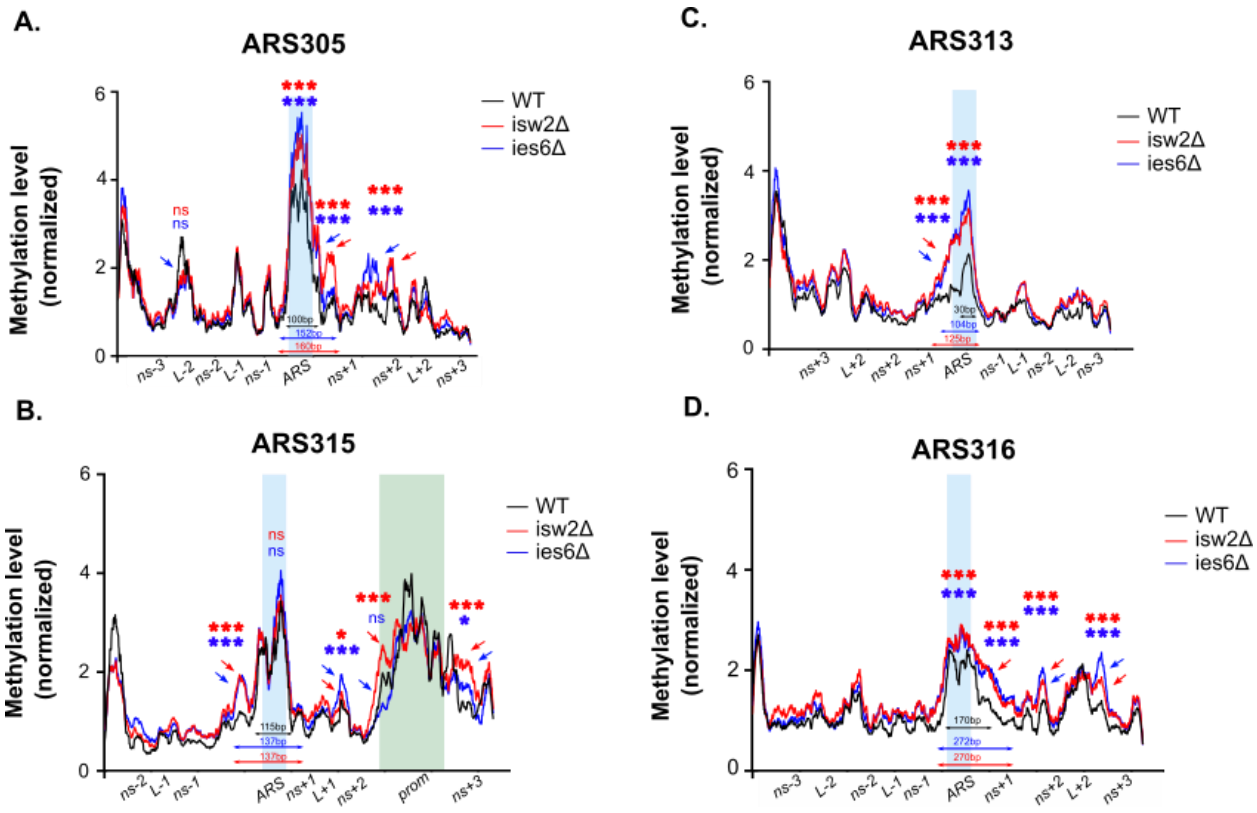


Figure 57. A-D. Comparative analysis of chromatin accessibility on replication origins between WT and CRE mutant strains shows statistically significant differences of specific features. The methylation level of each condition has been normalized to the maximum methylation value and smoothed using a 30 bp sliding window. The size of each the NFR of the ARS in wildtype and CRE mutants is indicated at the bottom using specific thresholds of mean methylation for each ARS (ARS305, ARS315 >1.2, ARS313 >1.4 ARS316 > 1.13). Statistical analysis performed between genomic bins representing nucleosomal positions (80 bp long), linker regions (10-40 bp) and the nucleosome-free ARS (100 bp) and Prom in ARS315 (175 bp).

Table 2. Overview of the changes on NFR of each ARS locus in combination with the changes on DNA replication efficiency.

	WT	isw2Δ	ies6Δ
ARS305	100 (E)	160 (IN)	152 (IN)
ARS313	30 (IN)	125 (IN)	104 (E)
ARS315	115 (E)	137 (IN)	137 (IN)
ARS316	170 (IN)	272 (IN)	272 (IN)

Overall, these shifts in NFR size in all four origins correlate with their observed altered phenotype after the deletion of Isw2 and Ies6 subunits. Firstly, the two originally efficient origins ARS305 and ARS315 decrease their replication efficiency in the CRE mutant condition. Additionally, broadening of the NFR at both ARS loci to larger sizes compared to the wildtype (100 bp -ARS305) and (115 bp -ARS315) was observed. Secondly, ARS313 changed its replication profile from inefficient to more efficient upon ies6 deletion and, in parallel, the previously very short NFR of ARS313 (30 bp) increases to 104 bp reaching the NFR level of the two efficient origins ARS305 and ARS315. However, the ARS313-*isw2*Δ mutant does not show any phenotypic change in replication timing and remains inefficient. Interestingly, the ARS313 NFR in the isw2 mutant presents a length of 125bp and is thus exceeding the size of ARS313-*ies6*Δ as well as the size of the two wildtype efficient origins ARS305 and ARS315. Finally, the inefficient ARS316 shows an inefficient replication profile in both CRE mutants increasing in parallel the already long NFR size from 170 bp to 270 bp and 272 bp.

Together, these results suggest that NFR needs an optimal narrow size that ranges according to this analysis between 100 – 115 bp in order to support efficient DNA replication, while the shorter or longer NFR length of ARS loci seems to be correlated to decreased efficiency.

6.9 Hierarchical clustering

Apart from measuring the average methylation per base pair, we also calculated the mean methylation level of each defined feature in each molecule and plotted the distribution per molecule in violin plots (**Figure 58**). Similar to the averaging methylation graphs, the ARS sites of the origins show striking differences between WT and CRE mutants, since the majority of the molecules display higher methylation levels at this feature. Additional to ARS, there are nucleosomal as well as linker regions with higher variability of chromatin states compared to the WT (**Figure 58A.** the -2L of *ies6*Δ or +1NS/+2NS of *isw2*Δ in ARS305, **B.** the +3NS of isw2Δ in ARS315, **C.** the +3NS and +2L of both CRE in ARS313 and **D.** -2NS and +2L of *isw2*Δ in ARS316). This observation indicates high heterogeneity in CRE mutants locally along each molecule, which might additionally be correlated to the phenotypic changes in replication efficiency.

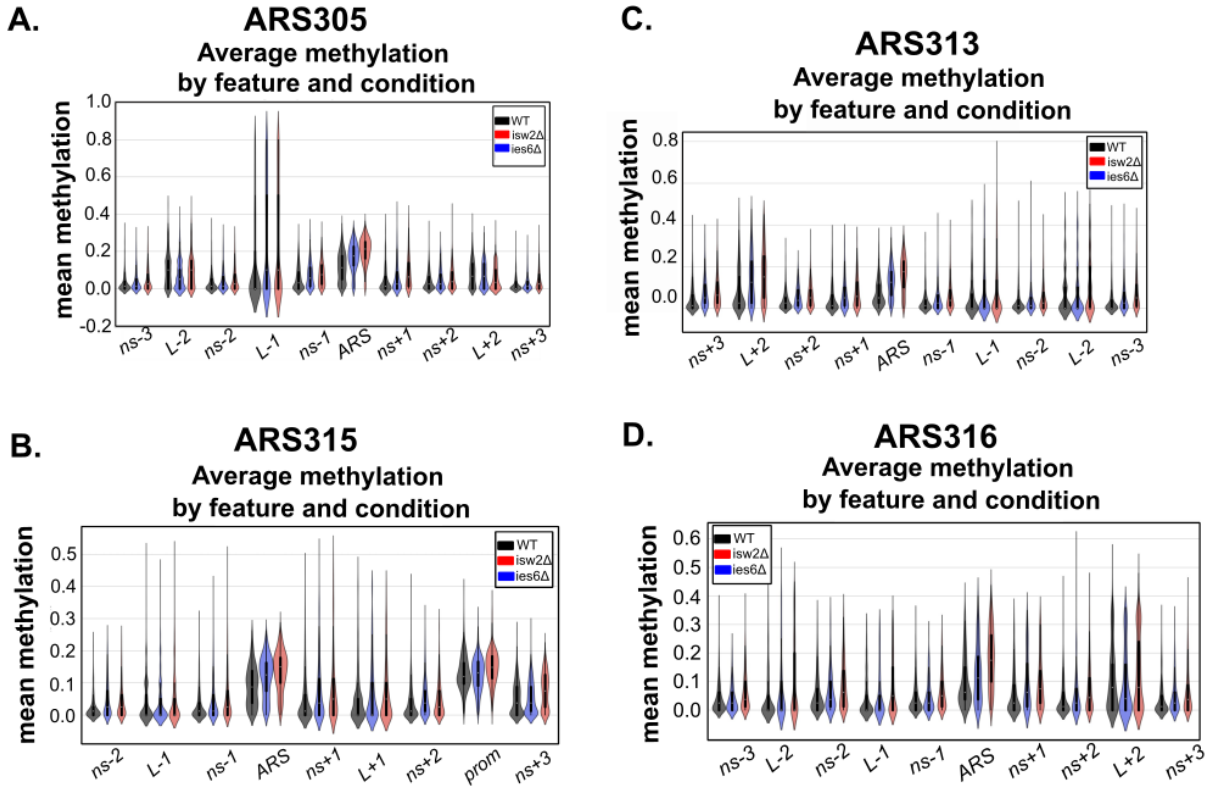


Figure 58. A-D. Comparative analysis of the average mean methylation of each specific features along the molecules of different replication origins between WT and CRE mutants. The size of ARS and the nucleosomal regions is 80 bp. The length of linker regions varies from 10 to 40 bp and the promoter region on ARS315 domain is 175 bp long.

To leverage the power of our single-molecule datasets and identify groups of molecules with common chromatin accessibility states, we performed hierarchical clustering. For clustering, we scaled normalized average feature methylation to zero mean and a standard deviation of one, and calculated the distance matrix as $\sqrt{(1-\rho)/2}$, where ρ is the Spearman's correlation coefficient between molecules. Based on these values, hierarchical clustering analysis of each origin produced a tree that was pruned to a height of 5 distinct clusters, so that each cluster presented a unique chromatin accessibility state. To ensure that our analysis is biologically meaningful and not derived from technical noise stemming from the pooled biological replicates, the contribution of each replicate to each of the five clusters was analyzed (**Figure 59A**). In this analysis, both replicates show similar contribution of reads to each cluster. Additionally, we confirmed that both DNA strand are similarly included in each cluster and there is no forward or reverse DNA strand overrepresented in one cluster, indicating that the clustering algorithm did not segregate or skew the results by DNA sequence and thus potentially different numbers of methylation sites present on the two DNA strands (**Figure 59B**).

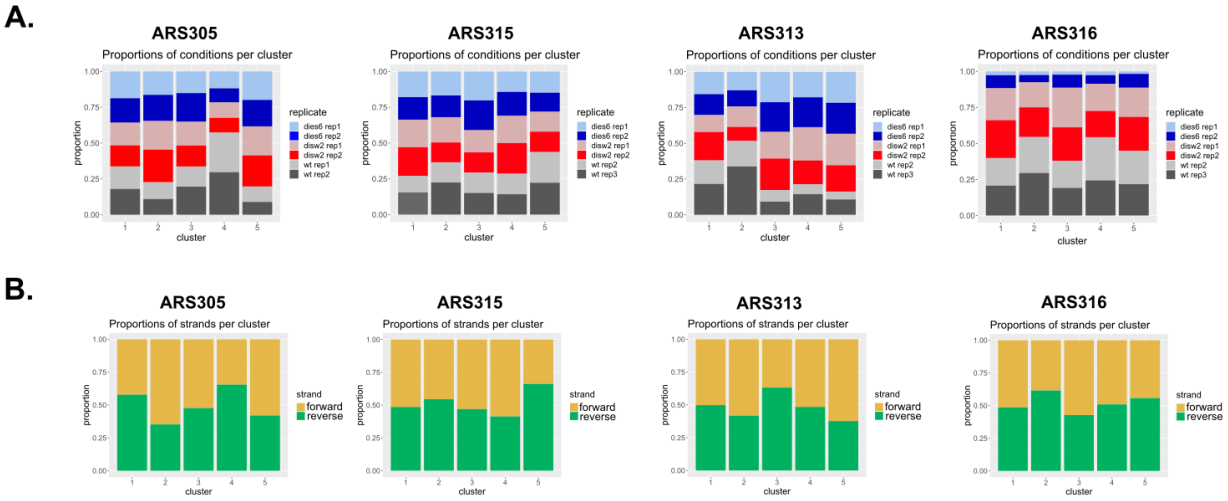


Figure 59. A. Bar graph plots show the read distribution of wildtype and CRE mutants per cluster of each replicate. **B.** Bar graph plots show the proportion of forward and reverse DNA strands per cluster for each replication origin.

For data visualization, a cartoon figure was created that visually reflects the chromatin state of each cluster by assigning each bin to one of three methylation states (low = strongly positioned nucleosome, intermediate = relatively occupied and high = free) (**Figure 60A-D**) and then created bar-plots containing the normalized proportion of reads of each condition (WT, *isw2*Δ and *ies6*Δ) in each cluster (**Figure 60E-H**).

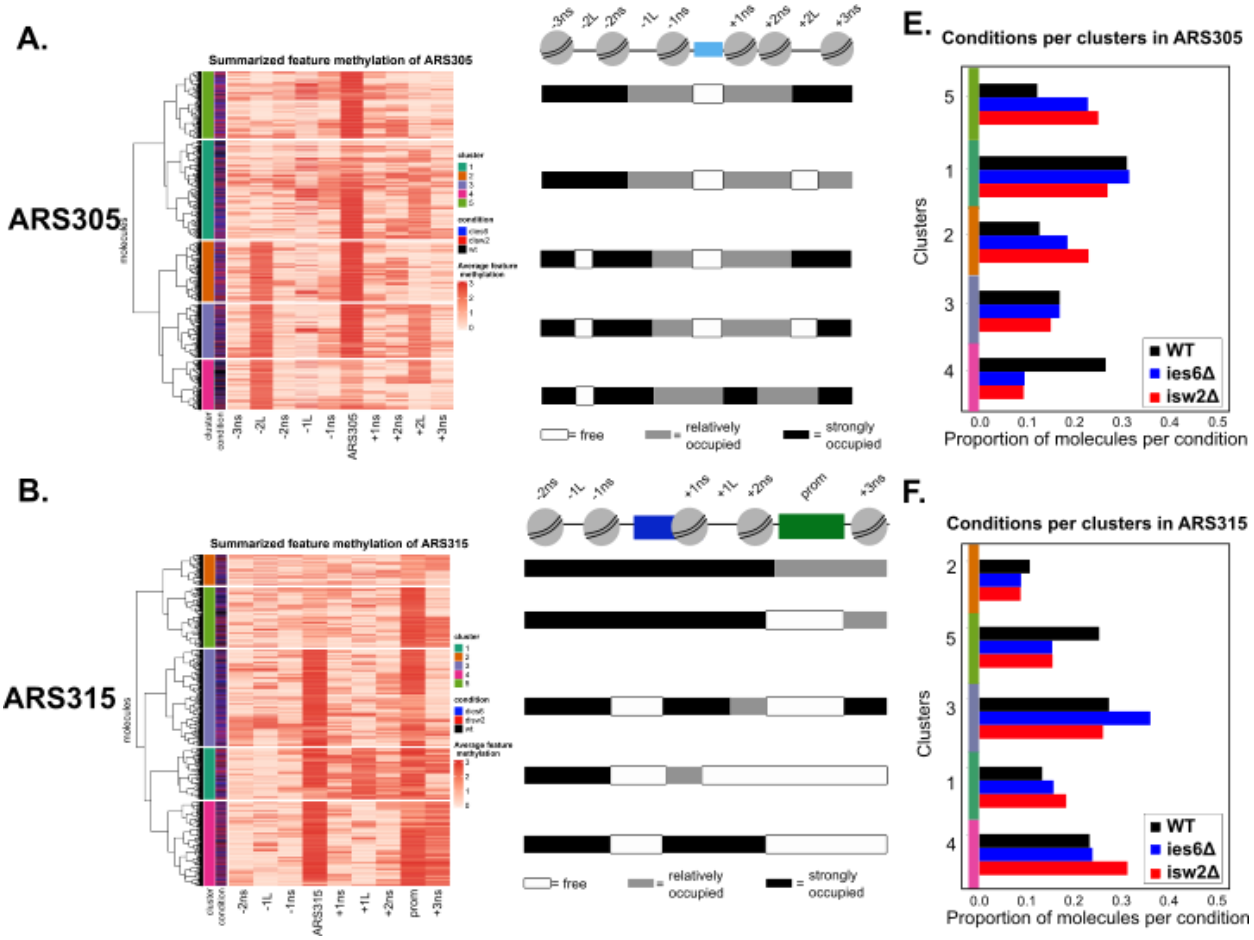
MATAC-Seq revealed a dynamic set of clusters whose occupancies between wildtype and CRE mutants strongly varied at the investigated EE and LI origins. In ARS305, cluster 1, which is characterized by open ARS site and well positioned +2 linker, contains the majority of the molecules in WT and both CRE mutants (**Figure 60A, E**). Unexpectedly, cluster 4 with intermediate methylation levels of ARS and +2L regions and free -2L contains also a large proportion of the WT molecules, highlighting the large heterogeneity and revealing so far hidden aspects of this EE origin. This cluster is underrepresented in CRE mutants, and the majority of their molecules, apart from cluster 1, is distributed to clusters 2 and 5. These two clusters share strongly protected +2L, potentially derived from nucleosome repositioning and resulting in “fuzzy” nucleosomal landscape on the + side. This observation might be correlated to the decreased replication efficiency upon CRE deletion (**Figure 60A, E**).

In ARS315, the nucleosome-free ARS and Prom sites surrounded by well positioned nucleosomes (cluster 3) is the most abundant chromatin state in WT cells (**Figure 60B, F**). Clusters 4 and 5 are the second more frequently occurring WT sub-groups. Interestingly and similar to ARS305, cluster 5 is characterized by protected ARS region challenging previous models connecting the EE origins strictly to open ARS regions. In CRE mutants, the majority of the molecules is distributed to clusters 3, 4 and 1 instead to 5. These clusters exhibit low nucleosome occupancy around the ARS and Prom sites, denoting again nucleosomal re-arrangement in the + side of the ARS315 chromatin locus (**Figure 60B, F**). The correlation of ARS315 with lower replication efficiency and disrupted +1NS, +2NS and +3NS positions is in agreement to the changes observed for ARS305 origin, stressing the importance of well positioned nucleosomal distribution on + side.

ARS313 is an LI origin with advanced replication efficiency upon ies6 deletion. In this origin, clusters 1 and 2 represent the chromatin state of the majority of molecules in the WT cells. Both clusters present a strongly protected chromatin state in contrast to the other three clusters. Particularly cluster 3 and 4 subgroups, which are abundant in CRE mutants, display an open ARS locus flanked by well positioned nucleosomes (**Figure 60C, G**). Cluster 3 contains additionally an open +2L region, a similar pattern to cluster 1 in EE ARS305 origin (**Figure 60A, E**), indicating again the potential importance of the + side of an origin with well positioned nucleosomes. This result would also fit with the data from the second EE origin ARS315 if one considers the

nucleosome-free Prom-site as a corresponding large linker region between the +2 and +3 nucleosomes.

Finally, the second LI origin ARS316 does not show any phenotypic change in replication efficiency between WT and CRE mutants. The majority of MATAC-Seq reads in wildtype cells corresponded to cluster 3 containing an open ARS flanked by relatively well but not strongly occupied regions (**Figure 60D, G**). This cluster was even further enriched in the CRE mutant samples at the cost of clusters 2 and 4. As the common feature of clusters 2 and 4 was again a large open +2L region and the mutants showed this state less frequently, this result provided further support that the combination of an open ARS with an optimal NFR size together with an open +2L region flanked by well positioned nucleosomes is a chromatin state that supports or at least highly correlates with early and efficient replication of these four origins.



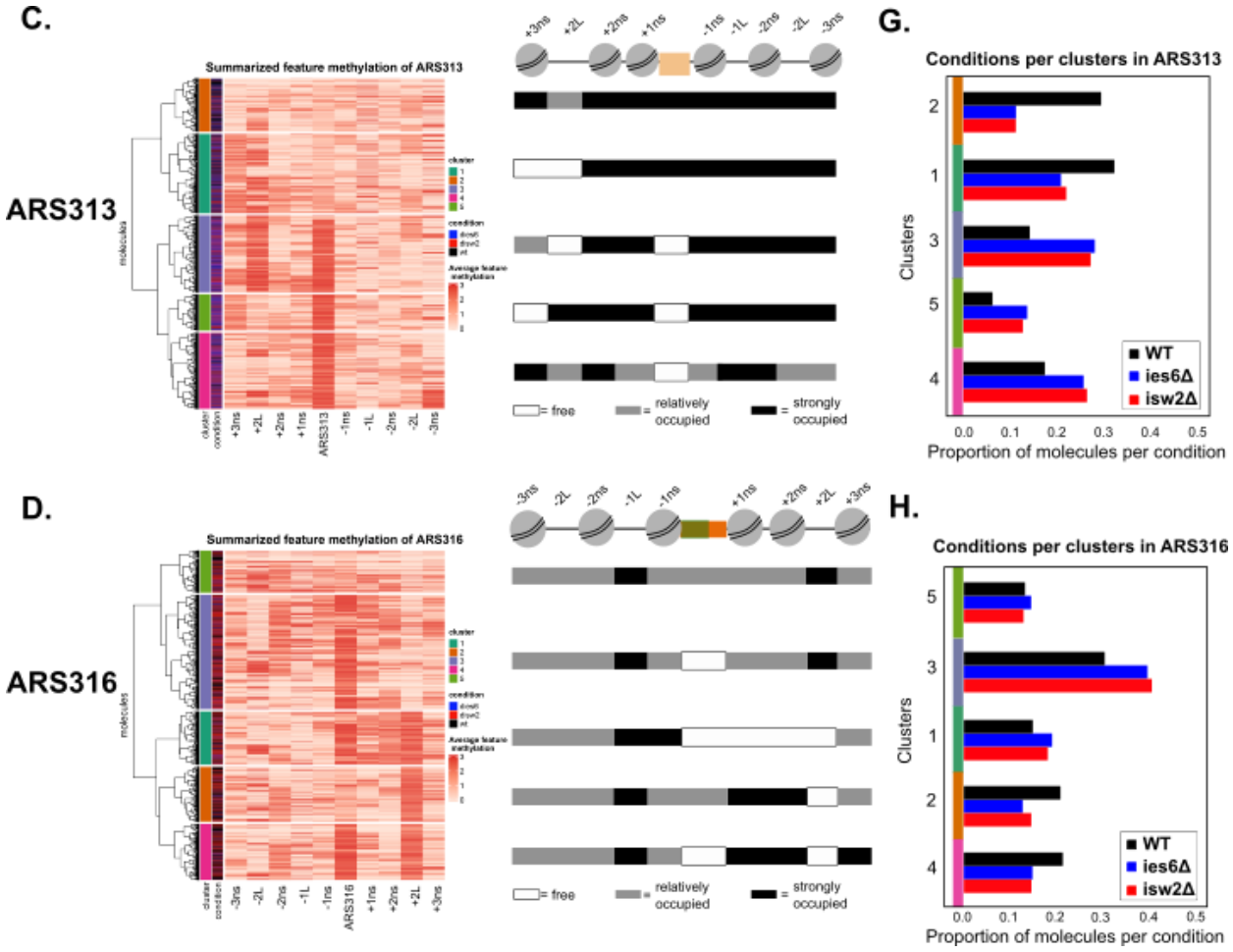


Figure 60. A-D. Heatmaps showing hierarchical cluster analysis of the summarized methylation levels per feature along the single ARS molecules. For each origin, the dendrogram was cut after 5 clusters. The cartoons on the right represent a visual representation of the associated chromatin accessibility state. Regions with high methylation level are considered as NFR and depicted as white. Regions with medium methylation level are considered as relatively occupied nucleosomal sites and depicted as grey and the low methylated features are considered as fully nucleosomal and represented as black. **E-H.** The plots show the proportion of the reads per condition and cluster. The total number of reads has been normalized to one.

7. Discussion

DNA replication is a fundamental biological process. Disruption of normal function of DNA replication could lead to phenotypic changes, diseases and cancer (Lomonosov et al., 2003; Poveda et al., 2010; Copeland, 2012). DNA replication is initiated at replication origins. In mammalian cells, the exact location of each replication origin is unknown and the genome is separated into early replicating (euchromatin) and late replicating regions (heterochromatin) (Miura et al., 2020). In contrast to mammalian cells, the yeast genome is smaller and the DNA replication origins are well-defined by specific DNA sequence elements named ARS (Bell and Dutta, 2002). Despite the fact that DNA replication origins in yeast contain a conserved DNA sequence and have been studied for several decades, it is still unknown why a given replication origin shows a distinct replication profile regarding its timing and efficiency (Yabuki et al., 2002). Additionally, it has been shown that there are no two cells with identical replication pattern at a specific time point (Czajkowsky et al., 2008b), suggesting high level of origin firing stochasticity and emphasizing the need of single molecule analysis. In this work, I hypothesized that one major factor determining the replication profiles could be the local nucleosome positioning. To exclude the genome-wide background noise deriving from population-based methods, we developed MATAC-seq, an affinity purification system on selected origins and performed single-molecule analysis of their chromatin accessibility pattern. The main findings of the project reveal strong correlation of an optimal size of the NFR at the ARS locus and an open region upstream of the +2NS with efficient replication origin initiation.

7.1 Establishment of MATAC-seq facilitates the purification of functional chromatin rings

To date, the chromatin landscape of EE and LI origins have mostly been investigated by bulk assays failing to reveal small details in the structure among the population and co-existence of variable chromatin states of a specific locus (Schones et al., 2008b; Bianco et al., 2015b; Ishii et al., 2015b; Chanou and Hamperl, 2021b). During my PhD project, I developed MATAC-Seq as a single-molecule method for purification of a targeted chromatin locus and investigation of its chromatin landscape at a single molecule level using Oxford Nanopore Technologies. In this way MATAC-seq generates chromatin occupancy heatmaps of a selected locus with unprecedented resolution and coverage. So far, large-scale genome-wide single-molecule techniques have limitations on their coverage and the detailed analysis of alternative chromatin states and their heterogeneities have been restricted to multi-copy gene loci (Abdulhay et al., 2020; Stergachis et

al., 2020b; Shipony et al., 2021). MATAC-seq has the potential to overcome these obstacles providing a solid basis for detailed characterization of alternative and rare chromatin states at a single-copy locus of interest.

One of the greatest advantages of MATAC-seq is that it can be applied to any yeast genomic locus with similar efficiency (**Figure 33**) without resulting in changes of nucleosome occupancy (**Figure 29**). For this project, MATAC-seq was established to compare the chromatin states of four different DNA replication origins, but the method could be applied to other genomic regions such as specific genes, promoters, centromeric or telomeric regions. Previously, single-molecule nucleosome occupancy landscapes of single-copy genes were obtained by psoralen-crosslinking EM analysis (Brown et al., 2013). MATAC-Seq provides a complementary tool and avoids the need for limited and specialized electron microscopy equipment for such technically more challenging single-molecule analysis (Brown and Boeger, 2014b). The recombination induction of the locus of interest by R-recombinase leads to excision of a functional chromatin ring (**Figure 28**). However, the DNA replication efficiency decreased by 19 % compared to the genomic control remaining integrated into the endogenous locus (**Figure 30**). One explanation could be that after recombination, the excised chromatin ring may become more diffusive in the nucleus and as a consequence, not all chromatin rings may receive the signal for replication initiation. For example, less limited initiation factors such as Cdc45 might have been loaded on ARS315 origin once it is recombined and removed from the initial chromatin environment. Apart from Cdc45, additional limiting factors which could possibly play critical roles for replication initiation are Sld2, Sld3, Dbf4, Dpb11 and Sld7. According to the model proposed by Mantiero et al 2011, the early and/or efficient origins are more competitive for recruiting these limiting factors compared to late and/or inefficient origins explaining their preferred activation in S-phase (Mantiero et al., 2011). However, the ARS315 origin might lose the ability of competed loading out of the chromatin landscape and hence replication is initiated less efficiently. his could be related to the subnuclear localization of each chromatin ring and 3D genome architecture. It has been shown that in yeast an early and efficient origin could be transformed into a late origin when placed into a transcriptionally silenced region close to the nuclear periphery (Ferguson and Fangman, 1992; Stevenson and Gottschling, 1999). Conversely, excision of a late ARS out of its chromosomal context and placement into a plasmid result in early replication (Ferguson and Fangman, 1992; Friedman et al., 1996). For this reason, apart from the local chromatin landscape, surrounding DNA elements of a larger two-dimensional or three-dimensional context might also contribute to the molecular mechanism controlling the replication timing.

7.2 Validation of MATAC-seq as a tool for single-molecule chromatin accessibility analysis

Given that MATAC-seq is a new approach, its validation though comparison with complimentary and well-established methods was necessary. For this purpose, psoralen crosslinking and EM analysis was used to detect nucleosome positioning on a single molecule level. This is a well-established method and is characterized as gold-standard approach for chromatin analysis in a single molecule level (Dammann et al., 1993; Brown and Boeger, 2014b; Hamperl et al., 2014a). Despite its resolution, the successful application of this method was a major challenge, since it requires not only specialized equipment but also high input amount of DNA because of the multiple washing and staining steps of the protocol. For this reason, the comparison between the MATAC-seq and psoralen crosslinking could only be performed on the multi-copy ribosomal ARS locus, since the yield from our 4 selected single-copy origins was not sufficient. Apart from that, psoralen crosslinking did not allow us to assure correct orientation along the linearized DNA molecules (**Figure 47**), even though previous studies showed that the LEXA binding sites at one end of the molecules could lead after denaturation to the formation of a Y-structure indicating the orientation (Hamperl et al., 2014b). Unfortunately, in this study the described Y-structure was only observed in a very small percentage of molecules, which could derive from stronger psoralen crosslinking or suboptimal denaturation conditions. For this reason, we could only compare the rARS profile of nucleosomal bubbles in a random orientation. Despite this drawback, MATAC-seq and Psoralen crosslinking rARS profiles mirrored each other well, suggesting good reproducibility of these two complementary methods. Nevertheless, the accuracy and precision of both methods could still be boosted by correctly-oriented psoralen-EM rARS nucleosomal patterns. In general, these technical challenges strengthen the usefulness and feasibility of MATAC-Seq to monitor chromatin accessibility of our targeted single-copy domains. Apart from the multi-copy rARS locus, we also compare MATAC-seq averaging profiles of the four single-copy replication origins with the bulk ChIP-Exo profiles (Rossi et al., 2021). In both cases we confirm that MATAC-seq detects successfully the NFR regions of the replication origins as well as the neighboring gene promoter in ARS315 (**Figure 48, 49**).

7.3 An optimal NFR length of ARS is correlated with efficient DNA replication

Given that MATAC-Seq provides the first investigation of EE and LI origins after their *ex vivo* isolation from native chromosomes, several important conclusions could be drawn that are

consistent but also challenge previous findings. Interestingly, MATAC-seq shows notable level of heterogeneity at all selected origins. There are clearly molecules at the EE origins ARS305 and ARS315 which exhibit low signal of methylation and therefore could be considered as fully nucleosomal but also subpopulations with overall high accessibility and therefore lacking nucleosomes across the whole domain. This high level of heterogeneity was unexpected given the well-established model that ARS regions are NFR and surrounded by well-positioned nucleosomes (Eaton et al., 2010). Similarly, even though that the overall MATAC-seq profile of LI origins presents less structured chromatin landscape compared to EE origins, which is agreement with older studies (Berbenetz et al., 2010b), there are also individual molecules showing molecules with open ARS locus, well-positioned nucleosomes and defined linker regions. How such high levels of heterogeneity functionally impact the stochastic activation of replication origins is an important question. To this end, we applied MATAC-Seq to CRE mutant strains of ISW2 and INO80, implicated in nucleosome sliding and eviction (Shen et al., 2003; Udugama et al., 2011; Clapier et al., 2017). Both EE origins decrease their replication efficiency after isw2 and ies6 deletion and in parallel increase the accessibility of the ARS locus (**Figure 57**). This broadening of NFR of ARS leads to disruption of the +/- 1 nucleosomal site, implying the importance of a well-positioned +/-1 nucleosome, which has been previously suggested (Lipford and Bell, 2001b). On the contrary, the LI ARS316 origin does not advance DNA replication efficiency but remains inefficient in CRE mutants despite the increased accessibility. In wildtype cells, the size of NFR region of ARS316 is 170 bp and is extended to 270 bp in both mutants (**Figure 57**). This indicates that high accessibility of ARS region alone is necessary but not sufficient for efficient replication, as it has been shown earlier (Eaton et al., 2010). Interestingly, the LI ARS313 origins in the wildtype has an ARS NFR size of 30 bp, but once ARS313 show efficient replication initiation in ies6 mutants, its NFR is increased to 104 bp. In other words, the length of the accessible NFR of ARS313 locus reaches similar levels to EE origin in the wildtype, suggesting that an optimal NFR size may assist in efficient DNA replication. One possible explanation for this optimal size is the strict and previously proven dependency of ORC binding with the +/-1 nucleosome, since it has been shown that the ORC can alter nucleosome position (Li et al., 2022) and also bind nucleosome through the bromo-adjacent homology (BAH) domain (De Ioannes et al., 2019). When the ARS is too short, as in wildtype ARS313 (30 bp), there might not be enough space for the efficient loading of the MCM2-7 complex (**Figure 57**), which could possibly lead to reduced number of MCM DH and hence less efficient DNA replication initiation. On the contrary, when the NFR of ARS is too long, the ACS is located far away from the first nucleosome, which works as

boundary, and ORC may lose its ability for efficient MCM loading and subsequent DNA replication initiation.

7.4 Open ARS region in combination with well-defined accessible +2 linker support efficient DNA replication

Hierarchical clustering uncovered a large range of heterogeneity between the alternative chromatin states of protected or unprotected regions for each molecule and origin. The molecules from all selected origins were classified and divided into 5 clusters according to the mean methylation level of the distinct features. This choice was derived by the fact that the resulting 5 clusters did not show major biases in their contribution from biological replicates and/or the ratio of forward/reverse strands (**Figure 59**) and therefore likely represent biological and not technical differences in chromatin accessibility. Analyzing and comparing the clusters of the four origins, we found out that the presence of open ARS in combination with a well-defined, accessible region following the +2NS and surrounded by well-positioned nucleosomes is strongly correlated with efficient DNA replication initiation (**Figure 60**). Prerequisite for the co-existence of this chromatin landscape and the replication profile is the presence of the optimal NFR of ARS locus. It is also very interesting that this finding slightly implicates an important role of the (+) side of our chromatin loci. As noted earlier, as (+) was named the side closer to ACS motif and thus ORC binding site. However, the mechanistic link of how such an asymmetric chromatin state could favor the bidirectional and thus symmetric process of replication initiation is an open question. One possibility would be that the open site after the +2NS would provide the suitable ground, regarding distance and orientation from ACS, for additional binding of other essential replication initiator factors e.g., MCM2-7 complex, Cdc45 or Sld2 (Mantiero et al., 2011). Especially for MCM2-7 complex it has been shown that early and efficient origins are characterized by higher number of loaded MCM2-7 complexes (Das et al., 2015b; Dukaj and Rhind, 2021b). Even though that the hierarchical clustering analysis uncovered key structure elements, it is clear that some of the clusters, for example: a) in ARS305 cluster 4 +2L, b) in ARS315 cluster 5 +3NS, c) in ARS313 cluster 4 -3NS and d) in ARS316 cluster 3 -2L, contain additional smaller subpopulations. At this point it is not clear whether the subpopulations are biologically meaningful and await further stratification of the data of reverse and forward DNA strand to reveal their potential biological function.

7.5 Limitations

Clearly MATAC-seq gave us the opportunity to correlate and further dissect the functional relationship between the chromatin accessibility landscape and the replication profiles of these four ARS loci. We also investigated the high level of heterogeneity that exist in one single locus and challenge previous studies supporting the uniformly high accessible ARS as major factor of DNA replication (Eaton et al., 2010). However, the establishment and profound characterization of the recombination-proficient yeast strains limited the number of selected replication origins as part of this work. Therefore, it is hard to draw global conclusions concerning all replication origins and possibly increase the impact of this study. Additionally, given that so far MATAC-seq has been applied only on a very defined and rather small region of 1kb around an origin, it was not possible to analyze the interplay of chromatin states between neighboring origins. For example, it would be very interesting to investigate whether the increased accessibility of a replication origin and the change in replication efficiency is accompanied with increased or decreased accessibility of its neighboring origin and also possible changes on their heterogeneity level. To address this question, strains containing two or more origins in the recombined domain should be constructed and used for MATAC-seq analysis. Apart from that, this project faced limitations regarding the available algorithms used for the analysis of the nanopore datasets. Even though the chromatin rings were treated with three different DNA methyltransferases responsible for CpG, GpC and m6A methylation, only two of them could be eventually analyzed using Megalodon, since there is no algorithm detecting all these modifications in parallel (Liu et al., 2021). Fortunately, the yeast genome is AT-rich, so the recovered resolution was enough for a robust analysis based on these two DNA methylation events. Another possible limitation of this study is the reduced number of recovered reads of the ARS316 locus in the ies6Δ CRE mutant. Given that this observation was reproducible across both biological replicates, a biological reason appears likely. One possible explanation is that the chromatin changes generated by the deletion of the ies6 subunit resulted in nucleosomal sliding and re-positioned a nucleosome onto the previously unprotected RS/*LEXA* site. As these genetic elements are important for recombination and efficient affinity purification, this alteration could possibly explain the low coverage of this locus by MATAC-seq in the ies6Δ CRE mutant.

7.6 Outlook

The specific enrichment of four different replication origins and their later investigation of chromatin structure indicates that this approach is applicable to any other single-copy locus in the yeast genome, for example on promoter regions in correlation to active or inactive transcription states (Brown et al., 2013). Apart from that, instead of limiting this method to a handful of selected regions, it appears feasible to extend MATAC-seq towards targeting and enriching larger chromosomal regions of several 10 - 100 kilobases combined with the advantage of third generation long-read sequencing technologies. Moreover, investigation and comparison of the proteomic pool of wildtype to the CRE mutants would provide additional evidence about important protein factors for DNA replication and provide complimentary information to this current work. Last but not least, the footprints in MATAC-seq could be also generated by any protein complex tightly bound to DNA not only nucleosomes. Hence, it could result in smaller footprints or even transcription factor binding footprints, reveal their loading sites and compare the findings with older studies (Kleinendorst et al., 2021) .

8. Material & Methods

8.1. Common techniques for yeast manipulation

A. Preparation of competent yeast cells

Yeast cells cultured in 50 ml medium and OD0.5 were collected after centrifugation at 2000 rpm, 5 min and RT. The pellet was washed 2x by sterile water and then they resuspended into 1 ml of SORB. Aliquots of 100ul were transferred to fresh and sterile reaction tubes and kept at -80 °C.

B. Transformation of competent yeast cells

The required number of reaction tubes/competent yeast cells kept at -80°C plus one background control were thawed on ice. Cells were pellet (2000 rpm, 2 minutes and 4 °C). After removing the supernatant the cells were resuspended into the transformation solution (260 µl PEG, 12µl LiAc, 10 µl Salmon sperm DNA already boiled at 95 °C, 1 µg – 5 µg desired DNA fragment) and heat-shocked at 42 °C for 20 min. The yeast cells were pellet again, resuspended into sterile water and transferred into selective agar plates.

C. Liquid culture

Single colony was picked from plates and cultured in 4-5 ml respective medium at 30 °C overnight. The day of the experiment cells were inoculated at OD 0.2 and grown until they reach the required OD. The culture volume should not exceed 1.3 of the corresponding flask, temperature is 30 °C and 200 rpm shaking.

D. Permanent yeast glycerol stocks

Yeast cells at stationary phases were mixed 1:1 ration with sterile 50% glycerol solution and kept at -80 °C

E. Establishment of MNase fusion strains

We established the Chec strain by using the plasmid K155 gifted by Joachim Griesenbeck (Merz et al., GAD 2008), which has a MNase-HA3-URA cassette cloned into a pBlueScript background. We fused MCM2 to MNase by conducting a PCR using overhang primers consisting of a 5' sequence complementary to the target gene (a 50 bp located right after the stop codon) and a 3' sequence complementary to the region after the previous 50 bp, allowing us to insert the cassette directly following the MCM2 gene. The resulting PCR product (overhang-Mnase-HA3-URA-overhang) was inserted next to MCM2 by homologous recombination. The positive strains were selected on uracil medium.

F. Establishment of yeast strains with integrated RS and LEXA binding sites in the ARS locus

The plasmids: K113 containing RS/LexA after the 3rd nucleosomal pair of ARS305, K116 containing RS/LexA after the 3rd nucleosomal pair of ARS316, K273 containing RS/LexA after the 3rd nucleosomal pair of ARS315 were amplified by primers 223/224, 251/252, 858/859 respectively. The resulting PCR products with 100 basepair homologous arms were inserted into the strains Y18, Y21 and Y89 respectively (see yeast, plasmid and primers tables below). The positive clones were selected on FOA medium (see Results Figure 26). Regarding the plasmid K170 different strategy was followed. The plasmid was digested after maxiprep by EcoRI and HindIII. The area containing RS/LexA after the 3rd nucleosomal pair of ARS313 was inserted into the Y44 strains by homologous recombination. FOA plates were used for positive clones selection. For all transformations 6ug of DNA were used.

8.2. Manipulation of yeast cells for tandem affinity purification method

A. Culture of yeast strains expressing the recombination cassette

Yeast strains capable of recombination were grown in 2 L YPR medium at 30°C with a rotation speed of 200 RPM, starting from an optical density (OD600) of 0.2. Once the OD600 reached 1.0, cells were arrested in the G1 phase by adding α-factor (50 ng/ml) and inducing recombination by introducing galactose to a final concentration of 2% (w/v) for 2 hours at 30°C. The cells were harvested through centrifugation (6000 rpm, 10 minutes, 4°C), the supernatant was discarded, and the cell pellets were resuspended in water and placed in a sealed 25 mL syringe. After another centrifugation step (3000 rpm, 12 minutes, 4°C), the supernatant was discarded and the

cells were frozen in liquid nitrogen, breaking them into small pieces, and stored at -80°C for future use.

B. Specific chromatin domain isolation with affinity purification

A TEFAL Prep`line coffee grinder was first cooled by grinding 30 grams of dry ice for 30 seconds. The resultant dry ice powder was then discarded. An appropriate amount of frozen cells, either 3-4 grams of multi-copy rARS circles or 5-6 grams of single-copy ARS circles, was mixed with approximately 60 grams of dry ice and processed in the coffee grinder. The grinder was run 10 times for 30 seconds each with short intervals in between. To prevent the ground cells from sticking to the grinder's wall, the grinder was occasionally tapped using a spatula. The ground yeast was then transferred into a plastic picker until the dry ice had fully evaporated. When chromatin methylation and nanopore sequencing were performed, each yeast strain's frozen pellet was ground separately and then combined into the same picker.

After the dry ice had evaporated, the powder was dissolved in 0.75 milliliters of cold MB buffer with Protease and Phosphatase inhibitors, in a ratio of 1:1 with the ground yeast cells. The resulting cell lysate was transferred into 2 milliliters of low-binding eppis and cleared of cell debris by centrifuging for 30 minutes at 16,000 x g and 4 °C. The chromatin ring-containing supernatant was then transferred into clean 2 milliliters of low-binding eppis and supplemented with magnetic beads coupled with IgG. Input IN DNA and protein samples (0.1 % and 0.05 %, respectively) were taken from the supernatant before the addition of magnetic beads. The supernatant was then incubated with the beads for 2 hours in a rotating wheel at 4°C. The beads were separated from the supernatant using a magnetic rack and washed 5 times with cold MB buffer with Protease inhibitors and once with cold MB buffer without Protease inhibitors. Flow-through samples for DNA and protein analysis (0.1 % and 0.05 %, respectively) were taken from the supernatant of the IgG affinity purification after each washing step.

The LexA-Chromatin ring complexes were then released by proteolytic cleavage of the LexA-TAP fusion protein using 200 U of 6xHis-tagged recombinant TEV protease for a single yeast strain or 1000 U for all yeast strains together, in a total volume of 300 microliters for a single yeast strain or 400 microliters for all yeast strains together. The final elution, containing the chromatin rings, was separated from the beads and DNA and protein samples were taken from the beads (B) and final elution (E) (0.1 % and 0.05 %, respectively).

C. Specific chromatin domain isolation with affinity purification – DNA analysis

The DNA samples from the purification process were mixed with H2O to reach a final volume of 100 µl. A control for the DNA extraction efficiency was added in the form of 1.13 ng of plasmid DNA (K71) to each sample. The DNA samples were then supplemented with 100 µl of IRN buffer (50 mM Tris–HCl pH 8, 20 mM EDTA, 500 mM NaCl) and 2µl RNAse A (10mg/ml) and incubated at 37 ºC for 1 hour. Next, 5 µl Proteinase K (10mg/ml) and 10 µl of SDS 20% were added and incubated for an additional hour at 56 ºC. Afterward, 200µl Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) was added, followed by 2x10 sec of vortexing. The mixture was centrifuged for 7 minutes at 21,000 x g to separate the organic and aqueous phases. The supernatant was transferred to a fresh 1.5 ml tube containing a mixture of ethanol (2.5 : 1) and 2 µl glycogen (10mg/ml) and stored at -20 °C overnight. After a 45-minute centrifugation step at 4 °C with 21,000 x g, the supernatant was discarded and 150 µl of 70% ethanol was added to the pellet. Another centrifugation step with 16,000 x g at 4 °C for 10 minutes resulted in a dried DNA pellet at room temperature for 10 minutes. Finally, the DNA samples were digested using the BsrGI restriction endonuclease (NEB) in a total volume of 35 µl and the DNA content was analyzed by qPCR.

8.3. Quantitative PCR

The measurement of a specific DNA fragment was done using qPCR to ensure high accuracy. The presence of DNA at the end of each PCR cycle was detected through the fluorescence of SYBR-Green, a dye that emits fluorescence when bound to DNA but not in solution. The intensity of the fluorescence signal determines the amount of DNA in the sample. Each qPCR reaction had a volume of 12 µl, consisting of 2 µl of DNA sample and 10 µl of master mix. The master mix included the forward and reverse primer and a ready SYBR GREEN solution. The qPCR was performed on a LightCycler480II system by Roche and the fluorescence was recorded at 510nm after excitation at 480 nm. Data analysis was done through the comparative quantitation module of the a LightCycler480II analysis software.

8.4.1 Western Blot Analysis

After SDS-PAGE, the proteins are bound with SDS and carry a negative charge, allowing for semi-dry transfer to a PVDF membrane through an electric current using a BIORAD semi-dry transfer apparatus. To facilitate the transfer, three layers of Whatman paper saturated with blotting buffer (25 mM Tris, 190 mM glycine, 20 % methanol, pH 8.3) are placed on the anode of the

transfer device. The membrane (Immobilon PSQ 0.2µm, Millipore) is prepped by first soaking it in methanol and then in the blotting buffer, before being placed onto the layers of Whatman paper. Care must be taken to remove any air bubbles, as they can obstruct the flow of the electric current. The membrane must be kept moist with blotting buffer throughout the transfer process. After disassembling the gel apparatus, the gel is placed onto the membrane, air bubbles are removed, and the gel is covered with three additional layers of soaked Whatman paper. The blotting process is conducted at 24V for 1.5 hours. Finally, the marker bands and lanes should be marked with a pen for reference.

8.4.2 Detection of proteins by chemiluminescence

The membrane was treated with a blocking solution (5 % milk powder in 1x PBST) to prevent non-specific binding of the antibody. The blocking was done either for 1 hour at room temperature or overnight at 4 °C while being agitated. The membrane was placed in a 50 ml falcon tube containing the primary antibody at the appropriate dilution in 1x PBST with 5 % milk powder (3 ml for large membranes) and rotated at room temperature for 1 hour. After three washes of five minutes each with 1x PBST in a tray, the membrane was placed in a 50 ml falcon tube containing the secondary antibody at the appropriate dilution in 1x PBST with 5 % milk powder (3 ml for large membranes) and rotated at room temperature for 30 minutes. The membrane was then washed three times for five minutes with 1x PBST and then 30 min incubation at RT with the secondary antibody was followed. After 3x washing with 1x PBST the membrane was put between two sheets of a thin plastic bag (Roth) and covered with a liquid film of reaction substrates (SuperSignal West Pico PLUS Chemiliminescent Substrate, ThermoScientific) and was developed by chemiluminescence.

8.5. Restriction Enzyme Accessibility assay and Micrococcus Nuclease assay

A. Preparation of nuclei

All steps were performed at 4 °C. The cells were first washed three times with 0.6 ml of buffer A containing 1× protease inhibitors. They were then collected by centrifugation (16,000 × g, 2 min) and the supernatant was discarded. The cells were suspended in 350 µl of buffer A with 1× protease inhibitors and mixed with ~500 µl of glass beads. It is important to ensure there is enough buffer solution to cover the beads by a thin layer. The cell disruption was done for 10 min at maximum speed on the Vibrax shaker at 4 °C. The cell lysates were collected by piercing the bottom and cap of the microtubes with a hot needle and transferring the crude cell lysates to new

1.5 ml microtubes. The nuclei were then isolated by centrifugation (16,000 × g for 2 min) and the supernatants were removed. The nuclei pellets were suspended in 0.6 ml of buffer A containing 1× protease inhibitors and the solution was centrifuged again (16,000 × g for 2 min).

B. For restriction enzyme Accessibility assay:

The supernatants were discarded and the nuclei pellets were resuspended with buffer A containing 1x protease inhibitors and 2 mM $MgCl_2$ in total volume of 400 ul. Each sample was split equally to reaction tubes according to the number of the enzymes. Enzymatic digestion reaction was performed for 1 h and then IRN buffer was added to terminate the reaction before proceed to DNA extraction.

C. For Mnase assay:

The nuclei were resuspended in 550 µl Mnase buffer (Buffer A without EDTA and 2 mM $CaCl_2$). After 3 min nuclei pre-incubation at 30 ºC and 750 rpm, the samples were supplemented by 3.5 U (Kunitz Units) of micrococcal nuclease (NEB) in order to initiate digestion. Once the digestion is complete 100 µl of IRN were added into the samples.

After mixing with 2 µl RNAse A (10mg/ml) and the samples were incubated at 37 °C for 1 h. Then 2.8 U of Proteinase K (from Sigma) were added, followed by incubation at 56 °C for 1 h and further cleaned through a phenol/chloroform extraction. The DNA is then precipitated using ethanol with 100 µg of glycogen/ml as a stabilizer overnight at -20 °C. The DNA is redissolved in 20 µl and subjected to Southern blot analysis

8.6 Southern Blot Analysis

The Southern blotting method was used to transfer DNA from an agarose gel to a positively charged nylon membrane (PositiveTM Membrane, MP Biomedicals). The double-stranded DNA was denatured by incubating the gel twice for 15 minutes in a solution of 0.5 M NaOH and 1.5 M NaCl, followed by two 15-minute incubations in transfer buffer (1 M NH_4OAc). The DNA was transferred onto the membrane through capillary flow, using a blotting pile consisting of Whatman papers, the gel, the membrane, and a weight to apply pressure. The pile was covered and left to blot overnight or for at least 6 hours. Finally, the DNA was crosslinked to the membrane by exposing it to ultraviolet light (0.3 J/cm^2), covalently binding the thymine bases to the amino groups of the membrane. The membrane can be dried and stored at room temperature.

8.6.1 Hybridisation

Up to four membranes can be stacked in a single hybridization tube, separated by meshes. The membranes were first wetted with 2x SSC and pre-hybridized for 1 hour at 65 °C using a hybridization buffer consisting of 0.5 M sodium phosphate buffer (pH 7.2) and 7 % SDS. After discarding the pre-hybridization buffer, 15ml of new, pre-warmed hybridization buffer was added to the tube. The probe was mixed with salmon sperm DNA, boiled for five minutes, and added to the tube. The hybridization was carried out overnight at 65 °C while the tube was rotated gently in a hybridization oven. The blots should be kept covered with liquid and should not roll or stick to the walls of the tube.

After hybridization, the probe in the hybridization buffer can be stored at -20 °C for reuse. The blots were then rinsed once with 30 ml 3x SSC and 0.1 % SDS. They were washed at 65 °C while the tube was rotated in the hybridization oven, using three washing buffers with decreasing salt concentration and increasing SDS concentration: Wash 1 (0.3x SSC, 0.1% SDS), Wash 2 (0.1x SSC, 0.1 % SDS), and Wash 3 (0.1x SSC, 1.5% SDS). Each wash step was repeated twice for 15 minutes. Finally, the blots were dried and stored at room temperature.

8.6.2 Detection of radioactive probes

The BAS-III imaging plate (IP) was erased using the Eraser (Raytest). The blot was placed in a BAS cassette 2040, and the IP was removed from the eraser in the dark and placed onto the blot. The length of exposure was determined based on the radioactive signal. The IPs were scanned with 100 µM resolution using a phosphor imager (FLA3000 by Fujifilm)

8.7 Replication timing

Yeast cells were grown in 2 L YPD medium at 30°C and 200 rpm, starting with an OD600 of 0.2. At an OD600 of 0.6, the cells were arrested in G1-phase using α-factor (50 ng/ml) for 2 hours. Samples of 1 ml were taken every 30 minutes (0, 30, 60, 90, 120 min) for flow cytometry analysis. After the 2-hour G1-phase arrest, the cells were released into S-phase by adding 125 U of Pronase (Sigma-Aldrich, 53702-25KU) and 20 mM potassium phosphate buffer. Samples for flow cytometry (1 ml) and quantitative PCR (4.5 ml) were taken every 8 minutes for 48 minutes.

8.8 Flow cytometry

Cells were spun at 16000 x g for 2 minutes. The supernatant was discarded and 1 ml of cold 70 % ethanol was added drop-by-drop with gentle mixing. The cell suspensions were stored at 4 °C until further use. 500-600 µl of the samples were transferred to a clean 1.5 ml reaction tube, spun, and the resulting pellet was resuspended in 300 µl of 50 mM Na-citrate and 0.1 mg/ml RNAse A. After a 2-hour incubation at 50°C, 3 µl of Proteinase K (10 mg/ml) was added (at a 1:100 ratio), followed by another 2-hour incubation at 50°C. 30 µl of each sample was mixed with 170 µl of 50 mM Na-citrate and 0.5 µM Sytox Green (S7020, ThermoFisher). Before FACS analysis, the samples were briefly sonicated (using a Bioruptor with 30-second intervals of 5 minutes ON and OFF) to separate cell clumps.

8.9 Chromatin methylation

The chromatin methylation treatment protocol was adopted and improved from (Shipony et al.,2021). The final chromatin purification (~200 ng total DNA in 400 µl) was combined with 100 ng of a nucleosomal array assembled in vitro and 1 µg of a naked plasmid DNA (K112) as control for methylation efficiency and nucleosome detection. The chromatin and naked DNA were first treated with 200 U of EcoGII and M.CviPi (NEB), 0.6 mM SAM (S-adenosylmethionine) (NEB), and 300 mM sucrose, then incubated at 30°C for 7.5 min. Another 100 U of each enzyme and 0.05 mM SAM was added and incubated for another 7.5 min at 30°C. Then, 120 U of M.SssI (NEB), 10 mM $MgCl_2$, and 0.05 mM SAM were added and incubated for another 7.5 min at 30°C. 0.05 mM SAM was added and the incubation continued for 7.5 min at 30°C. The methylation reaction was stopped with a 1:1 volume of Stop Buffer (20 mM Tris-HCl pH 8.5, 600 mM NaCl, 1 % SDS, 10 mM EDTA) followed by DNA extraction. All reactions were conducted in low-binding Eppendorf tubes.

8.10 Library preparation

The recovered DNA was converted into libraries for MATAC-seq using the Ligation Sequencing Kit 1D (Oxford Nanopore Technologies, SQK-LSK109) as per the manufacturer's instructions. The sequencing was performed on R9.4 MinION flowcells from Oxford Nanopore Technologies for up to 24 hours.

8.11.1 Nanopore basecalling

The software "Megalodon" was used to detect CpG and m6A methylation by base calling. This software package included versions 2.4.0 of "megalodon," 5.0.11 of "guppy," 2.17 of "minimap," and "res_dna_r941_min_modbases-all-context_v001.cfg." To determine threshold values for binarizing methylation states, methylation scores of 10 million sites were analyzed on unmethylated and fully methylated plasmid samples and a score cutoff was identified to achieve high specificity, separately for m6A and CpG methylation.

8.11.2 Analysis of nanopore sequencing data

The binary count data was normalized based on the average methylation levels of the controls from the spiked-in nucleosome array. The replicates from the same condition were combined, the mean methylation levels per position were calculated, and a Wilcoxon-rank-sum test was used to determine the p-values. The raw binary data is shown in heatmaps. The average methylation per feature was summarized by calculating the mean normalized methylation for each molecule and feature. To cluster the data, the normalized average feature methylation was scaled to have a mean of zero and a standard deviation of one, and the distance matrix was calculated using the Spearman's correlation coefficient between the molecules. Hierarchical clustering was performed using the ward.D2 method in R's hclust, and the dendrogram was cut after 5 clusters. The heatmaps display the average feature methylation per molecule, the cluster assignment, and the condition. The bar plots show the proportion of reads per condition and cluster, normalized to a total of 1 for each condition.

8.12.1 Psoralen crosslinking

The purified chromatin rings in 400 µL eluate were added to two wells of a 24-well culture plate. Each sample was mixed with 10 µL Trimethylpsoralen (TMP) (0.2 mg/ml in ethanol) and incubated in the dark on ice for 5 minutes. The samples were then exposed to ultraviolet light for 5 minutes, 5 cm away from a 366-nm ultraviolet bulb in a Stratalinker. After each 5-minute exposure, 10 µL of TMP was added and the sample was subjected to additional 6-minute, 7-minute, and 8-minute incubations. After the final exposure, the samples were transferred to microtubes and combined with 200 µL IRN. RNase A was added to a final concentration of 0.33 mg/ml and incubated at 37 °C for 2 hours. The samples were then treated with 0.33 mg/ml of Proteinase K and 0.5% SDS and incubated for 4 hours at 55 °C. The DNA was extracted with phenol/chloroform, precipitated overnight at 4 °C, and digested with 3 µL NcoI restriction enzyme (NEB) in a 23 µL total reaction volume.

8.12.2 Denaturing spreading and analysis by electron microscopy (EM)

The denaturing spreading reaction involved a mixture of 2.0 µl formamide, 0.4 µl glyoxal, and 2 µl of chromatin rARS locus, which was incubated for 10 minutes at 42°C using a thermo-mixer, and then transferred to an ice water bath. Next, the reaction was spread on carbon-coated 400-mesh magnetic nickel grids by using the BAC (benzyldimethylalkylammonium chloride) method over a water surface. The DNA was then made electron dense by platinum-carbon rotary shadowing with a high vacuum evaporator (MED 020; Bal-Tec). The grids were scanned with a transmission electron microscope (Tecnai G2 Spirit; FEI; LaB6 filament; high tension ≤ 120 kV) and images were captured using a side-mount CCD camera (Orius 1000; Gatan, Inc.). The images were processed with DigitalMicrograph Version 1.83.842 (Gatan, Inc.) and analyzed using ImageJ64.

8.13 Materials

Media		
YP	Yeast extract Peptone Agar (plates) Autoclave	10g/l 20g/l 20g/l
YPD medium	YP + Glucose	20g/l
YPR medium	YP + Raffinose	20g/l
YPD- kanamycin	YPD + kanamycin	200mg/ml

Consumables

Name	Company	Katalog number
Yeast Extract	BD (VWR)	212720
Bacto Agar	BD (VWR)	214010
Bacto Peptone	BD (VWR)	211820
D-(+)-Galactose	Sigma Aldrich	G0625-1KG / 5KG
Polyethylenglykol 4000 (Rotipuran)	Carl Roth	156.3

A-D-Raffinose	SERVA	34140.03
Triton X-100	Sigma Aldrich	X100-100ML
Agarose (SERVA Wide Range)	Serva	11406.03
LITHIUM ACETATE DIHYDRATE	Santa Cruz	Sc-257671
Yeast Nitrogen Base without amino acids	Sigma Aldrich	Y0626-1KG
Yeast Nitrogen Base with amino acids	Sigma Aldrich	Y1250-250G
Yeast Synthetic Drop-out medium Supplements without URACIL	Sigma Aldrich	Y1501-20G
Yeast Synthetic Drop-out medium Supplements without LEUCINE	Sigma Aldrich	Y1376-20G
Yeast Synthetic Drop-out medium Supplements without HISTIDINE	Sigma Aldrich	Y1751-20G
Kanamycin	Sigma-Aldrich	K400-25G
Tween20	Kraft	18014332
DTT (Dithiothreitol)	Thermo Scientific	R0861
β- Mercaptoethanol	Sigma-Aldrich	07604-100ml
Sodium azide	Santa Cruz Biotechnology	sc-208393
Psoralen	Sigma Aldrich	
Protease inhibitor cocktail 100x (Halt, EDTA-free)	Thermo Scientific	78445
PEG-8000	Biosciences	RC-077
Ammonium Sulfate	Santa Cruz	Sc-29085
Protease and Phosphatase Inhibitor Cocktail 100x	Thermo Fisher Scientific	78446
Acetone	Carl Roth	5025.1
Lithiumchlorid	CarlRoth	3739.1
SYBR green I nuclei acid dye & SYBR green II RNA gel stain	ThermoFisher Scientific	S32717
Sodium hydroxide	EMSURE	1.06498.1000
Sodium chloride	Sigma	S9888-1KG
Ammonium acetate	Sigma	A7262-5KG
Magnesium acetate tetrahydrate	Sigma	M5661-250G
SuperSignal West Pico PLUS Chemiliminescent Substrate	ThermoScientific	34580

Plasmid Table

K004 (pM49.2)		Plasmid pM49.2 is a derivative of pABX22, and has been modified by addition of a LexA-binding cluster juxtaposed to an RS element.
K005	Yeast expression vector for constitutive expression of LexA-TAP under control of TEF2 promoter and inducible expression of R Recombinase under control of GAL1-10 promoter; LEU2 selection marker framed with RS sites (Hamperl et al., 2014)	
K009	*E. coli*/yeast shuttle vector used for genomic integration of CYC1 LexATAP GAL1-10 RecR expression cassette by recombination in URA3 locus (Hamperl et al., 2014)	
K18		pBlueSkript SK (+)
K94	Plasmid containing wildtype ARS305 locus in pBlueScript backbone (Weiß et al., 2023)	HindIII/PstI cut amplicon (primer 0016/0017) from yeast gDNA into K18
K95	Plasmid containing wildtype ARS316 locus in pBlueScript backbone	EcoRI/PstI cut amplicon (primer 0020/0021) from yeast gDNA into K18
K102	Vector with ARS305 +-2 nucleosome sequence next to lexA/RS sites (Weiß et al., 2023)	Gibson assembly of two fragments (primer 0119/0120 with K94 as template + HpaI/XhoI cut backbone from K0004)
K103	Vector with ARS305 +-3 nucleosome sequence next to lexA/RS sites (Weiß et al., 2023)	Gibson assembly of two fragments (primer 0121/0122 with K94 as template + HpaI/XhoI cut backbone from K0004)

K112	Vector for yeast transformation in order to modifiy ARS305 locus and insert RS sites next to NS+-2 (Weiß et al., 2023)	Gibson assembly of two fragments (primer 0046/0047 with K94 as template + primer 0048/0049 with K102 as template)
K113	Vector for yeast transformation in order to modifiy ARS305 locus and insert RS sites next to NS+-3 (Weiß et al., 2023)	Gibson assembly of two fragments (primer 0050/0051 with K94 as template + primer 0052/0053 with K103 as template)
K116	Vector for yeast transformation in order to modifiy ARS316 locus and insert RS sites next to NS+-3 (Weiß et al., 2023)	Gibson assembly of two fragments (primer 0074/0075 with K95 as template + primer 0076/0077 with K106 as template)
K139	*E. coli*/yeast shuttle vector used for genomic integration of pCYC1 LexA-TAP GAL1-10 RecR expression cassette by recombination in 500bp homology region from K121 of yeast chromosome I, LEU2 selection marker framed with RS sites (two mutations in the lexA gene that stop binding to lexA binding site: V11A, N171D) (Weiß et al., 2023)	AscI/NheI cut amplicon (primer 0270/0271 template K121) inserted with AscI/NheI cut amplicon (primer 0272/0273 template K009). Resulting plasmid was cut with NheI/PacI and inserted with NheI/PacI cut amplicon (primer 0274/0275 template K009).
K155	Vector containing MNase-HA3-URA cassette cloned into a pBlueScript background (gifted by Prof. Dr. Joachim Griesenbeck)	
K173	Plasmid containing wildtype ARS313 locus in pBlueSkript backbone	EcoRI/HindIII cut amplicon primers 319and 320 from gDNA into K18
pM49.2_ARS313+-1	Vector with ARS313+-1 nucleosome sequence next to lexA/RS sites	Gibson assembly of two fragments primers 321and 322 with K173 as template

		and HpaI/XhoI cut backbone from K004
pM49.2_ARS313+-2	Vector with ARS313+-2 nucleosome sequence next to lexA/RS sites	Gibson assembly of two fragments primers 331and 332 with K173 as template and HpaI/XhoI cut backbone from K004
pM49.2_ARS313+-3	Vector with ARS313+-3 nucleosome sequence next to lexA/RS sites	Gibson assembly of two fragments primers 333and 334 with K173 as template and HpaI/XhoI cut backbone from K004
K168	Vector for yeast transformation in order to modify ARS313 locus and insert the RS sites next to NS+-1	Gibson assembly of two fragments primers 346and 347 with pM49.2ARS315+-1 as template and primers 348 and 349 with K173 as template
K167	*E. coli*/yeast shuttle vector used for genomic integration of pCYC1 LexA-TAP GAL1-10 RecR expression cassette by recombination in 500bp homology region from K121 of yeast chromosome I, LEU2 selection marker framed with RS sites (Weiß et al., 2023)	Insert from K009 (NsiI/BlpI) cloned into K139
K169	Vector for yeast transformation in order to modify ARS313 locus and insert the RS sites next to NS+-2	Gibson assembly of two fragments primers 350and 351 with pM49.2ARS315+-2 as template and primers 352 and 353 with K173 as template
K170	Vector for yeast transformation in order to modify ARS313 locus and insert the RS sites next to NS+-3	Gibson assembly of two fragments primers 366and 367 with pM49.2ARS315+-3 as template and primers 368 and 369 with K173 as template
K174	Plasmid containing wildtype ARS315 locus in pBlueSkript backbone	EcoRI/HindIII cut amplicon primers 323and 324 from gDNA into K18
pM49.2_ARS315+-2	Vector with ARS315+-2 nucleosome sequence next to lexA/RS sites	Gibson assembly of two fragments primers 775and 776 with K174 as template and primers 771 and 772 with K004 as template

pM49.2_ARS315+-3	Vector with ARS315+-3 nucleosome sequence next to lexA/RS sites	Gibson assembly of two fragments primers 777and 778 with K174 as template and primers 771 and 772 with K004 as template
K238	*E.coli*/yeast shuttle vector used for genomic integration of pTEF2 LexA-TAP GAL1-10 RecR expression cassette by recombination in 500bp homology region from K121 of yeast chromosome I, LEU2 selection marker framed with RS sites (Weiß et al., 2023)	NcoI/AflII cut amplicon (primer 0769/0770 template K005) into K167
K293	Vector for yeast transformation in order to modify ARS315 locus and insert the RS sites next to NS+-2	Gibson assembly of two fragments primers 815and 816 with pM49.2ARS315+-2 as template and primers 817 and 818 with K174 as template
K273	Vector for yeast transformation in order to modify ARS315 locus and insert the RS sites next to NS+-3	Gibson assembly of two fragments: primers 819 and 820 with pM49.2ARS315+-3 as template and primers 821 and 822 with K174 as template
K207		Bar ko plasmid
K303	Vector for yeast transformation in order to delete the les6 gene and replace it with histidine	Gibson assembly of four fragments: Primers 1049 and 1050, 1053 and 1054 with gDNA as template, Primers 1051 and 1052 with K207 plasmid as template, K18 digested by

Yeast strains

Y0001 (Y01408)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4	Wildtype strain	EUROSCARF
Y0008 (ySH089)	MATa; ade2-1; ura3-1; trp1-1; leu2-3,112; his3-11; can1-100, URA3::LEU2 pTEF2 LEXA-TAP pGAL RecR		Gifted by Joachim Grieschemberg
Y0010 (yMW2)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; ARS305::URA3	ARS305 exchanged for *URA3*	Transformation of amplicon derived from K001 (primer 0040/0041) into Y0001. Selection on Ura
Y0011 (yMW3)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; ARS316::URA3	ARS316 exchanged for *URA3*	Transformation of amplicon derived from K001 (primer 0080/0081) into Y0001. Selection on Ura
Y0018 (yMW6)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS305_NS+3_RS	RS sites and lexA binding sites at ARS305 after +-3 nucleosomes	Transformation of amplicon derived from K113 (primer 0223/0224) into Y0010. Selection on FOA
Y0021 (yMW9)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS316_NS+3_RS	RS sites and lexA binding sites at ARS316 after +-3 nucleosomes	Transformation of amplicon derived from K116 (primer 0251/0252) into Y0011. Selection on FOA
Y0042 (yAC02)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; ARS313::URA3	ARS313 exchanged for *URA3*	Transformation of amplicon derived from K001 (primer 0338/0457) into Y0001. Selection on Ura

Y0043 (yAC03)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; ARS315::URA3	ARS315 exchanged for *URA3*	Transformation of amplicon derived from K001 (primer 0340/0458) into Y0001. Selection on Ura
Y0044 (yAC04)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS313_NS+3_RS	RS sites and lexA binding sites at ARS313 after +-3 nucleosomes	Transformation of EcoRI/HindIII digested plasmid K170 into Y0042. Selection on FOA
Y0065 (yTS3)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS305_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR	RS sites and lexA binding sites at ARS305 after +-3 nucleosomes. Expression cassette for R Recombinase and lexA (*TEF2* promoter)	Transformation of SbfI digested plasmid K238 into Y0018. Selection on Leu
Y0066 (yTS4)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR	WT strain with expression cassette for R Recombinase and lexA (*TEF2* promoter)	Transformation of SbfI digested plasmid K238 into Y0001. Selection on Leu
Y0069 (yMW38)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS316_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR	RS sites and lexA binding sites at ARS316 after +-3 nucleosomes. Expression cassette for R Recombinase and lexA (*TEF2* promoter)	Transformation of SbfI digested plasmid K238 into Y0021. Selection on Leu
Y0084 (yMW53)	MATa; ade2-1; ura3-1; trp1-1; leu2-3,112; his3-11; can1-100; bar1::his3	Bar1 knockout	Transformation of amplicon derived from K207 (primers 932/933) into Y8. Selection on His
Y0089 (yAC14)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS315_NS+3_RS	RS sites and lexA binding sites at ARS315 after +-3 nucleosomes	Transformation of amplicon derived from K273 (primer 0858/0859) into Y0043. Selection on FOA
Y0091 (yAC16)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0;	RS sites and lexA binding sites at	Transformation of SbfI digested

	bar1::kanMX4;RS_LEXA_NS-3_ARS315_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR	ARS315 after +-3 nucleosomes. Expression cassette for R Recombinase and lexA (*TEF2* promoter)	plasmid K238 into Y0089. Selection on Leu
Y0094 (yAC19)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS313_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR	RS sites and lexA binding sites at ARS313 after +-3 nucleosomes. Expression cassette for R Recombinase and lexA (*TEF2* promoter)	Transformation of SbfI digested plasmid K238 into Y0044. Selection on Leu
Y0102 (AC20)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS315_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ; isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0091. Selection on URA
Y0103 (AC21)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0001. Selection on URA
Y0104 (AC22)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0066. Selection on URA
Y0105 (AC23)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS305_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0065. Selection on URA
Y0106 (AC24)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS316_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0069. Selection on URA

	TAP pGAL1-10 RecR;isw2Δ::ura3		
Y0107 (AC25)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS313_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR;isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into Y0069. Selection on URA
Y127 (AC37)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS305_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;ies6Δ::his	Deletion of *IES6* gene	Transformation of SbfI and SacI digested plasmid K303 into Y0065. Selection on histidine
Y128 (AC38)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS313_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;ies6Δ::his	Deletion of *IES6* gene	Transformation of SbfI and SacI digested plasmid K303 into Y0094. Selection on histidine
Y129 (AC39)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; RS_LEXA_NS-3_ARS316_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR;ies6Δ::his	Deletion of *IES6* gene	Transformation of SbfI and SacI digested plasmid K303 into Y0069. Selection on histidine
Y130 (AC40)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4;RS_LEXA_NS-3_ARS315_NS+3_RS; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;ies6Δ::his	Deletion of *IES6* gene	Transformation of SbfI and SacI digested plasmid K303 into Y0091. Selection on histidine
Y135 (AC44)	MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; bar1::kanMX4; Chr I 212kb::LEU2 pTEF2-LEXA-TAP pGAL1-10 RecR ;ies6Δ::his	Deletion of *IES6* gene	Transformation of SbfI and SacI digested plasmid K303 into Y0066. Selection on histidine
Y136 (AC45)	MATa; ade2-1; ura3-1; trp1-1; leu2-3,112; his3-11; can1-100; bar1::his3; isw2Δ::ura3	Deletion of *ISW2* gene	Transformation of amplicon derived from K1 (primers 848/849) into

			Y0084. Selection on URA

Antibodies

HA-tag	1st/2nd	Anti-HA-Peroxidase, high affinity (3F10)	rat	HRP	25ug	IgG1	Sigma-Aldrich
Peroxidase-anti-peroxidase (PAP)	1st/2nd	Peroxidase Anti-Peroxidase Soluble Complex antibody produced in rabbit for detection of TAP-tagged proteins	rabbit	HRP		IgG	Sigma-Aldrich (P1291-500UL)

Primer table

#0016	GATAAGCTTttttgggtcctttgttttcg
#0017	GATGAATTCatttggagggaggagaagga
#0020	TCAGAATTCtcctccagtgggattgctac
#0021	TCACTGCAGgcctagcgggctattacctt
#0040	ttttgggtcctttgttttcgttgtttcagtctggataaattttaagttac tcgatccgatgataagctgtc
#0041	ATTTGGAGGGAGGAGAAGGATAACAGCGACGAAACACCGGACAGATTCCCtgccacctgac gtctaagaa
#0046	GAAAGTAGTTATTACGGCGTCGG
#0047	GGCTCTAGGGTAGTTGCG
#0048	gcgacgcccgacgccgtaataactactttcTCGACCCGAGATCATATC
#0049	caatgagagaaacgcaactaccctagagccCCACGATTTGATGAAAGAATAAC
#0050	GTGTGCTAAGTGTCCTGTTTC
#0051	AATATTGTCTTTGGACGTTTG
#0052	gaacgttccgaaacaggacacttagcacacTCGACCCGAGATCATATC
#0053	tggtttgggcaaacgtccaaagacaatattCCACGATTTGATGAAAGAATAAC
#0074	GAAAAGGGCATGTAATATTG
#0075	CTTGGGGAAGAAGTAACAATGAC
#0080	tcctccagtgggattgctacttctttgttgctgctgcatcctcaacttg tcgatccgatgataagctgtc
#0081	gcctagcgggctattaccttgtaaataccacactatcaatccttaaatgt tgccacctgacgtctaagaa
#0119	aaatcgtggtcgaccggcatgcaagctcccTCGAGGACAGACCACTTATG
#0120	aaccagtggttatatgtacagtagtacgttATCTCTCCGCCTGAATAAG
#0121	aaatcgtggtcgaccggcatgcaagctcccTCGAGAAATACAGAATAGGAAAG
#0122	aaccagtggttatatgtacagtagtacgttATCGTAGCGGTGTTTATC

#0223	Ggcttttcgatcagacttggcatgtgactaatcaagtatggcatgctggt ttttgggtcctttgttttcg
#0224	Tagtaaataacggagactggcgaaccgaatgggcacctgcctctgactgc atttggagggaggagaagga
#0251	TAACTTCAGCACCAAAGCCAACAACTACGACCTATGTCGAGCAACGACTTTCCTCCAGT GGGATTGCTAC
#0252	TTCTTGGCAGTCACATATATGGAAGGTGAATTTAGAGTAGTTTCCTTATAGCCTAGCGG GCTATTACCTT
#0270	TCAgctagcttaattaaAAGACAACAGATTTATTGTA
#0271	TCAggcgcgccCCCGAGGATTATAATTGTTC
#0272	tcaGGCGCGCCtttcgtctcgcgcgtttcgg
#0273	tcaGCTAGCagtagttggaatatcataat
#0274	tcaGCTAGCgagaatttgtattttcaggg
#0275	tcaTTAATTAAccccgttccacaacacaaca
#0319	AAC GAATTC aaaaaccccagcagcagata
#0320	AAC AAGCTT ctcctgggccttgatgatac
#0321	gtggtcgaccggcatgcaagctccctttcctgggcggtttggttac
#0322	gtggttatatgtacagtagtacgttctaactgtaggcgcttttatctcc
#0323	AAC GAATTC GGGAAGGAGATCTGCAGTGT
#0324	AAC AAGCTT CGATTGAAGCCGTGGATGAG
#0330	gtggttatatgtacagtagtacgttGGTTGGGGTATTAAGAGTAC
#0331	gtggtcgaccggcatgcaagctcccTTTTGTATGAAAAACTCATGAATC
#0332	gtggttatatgtacagtagtacgttCTTGGTGCAAGAAGTTGG
#0333	gtggtcgaccggcatgcaagctcccGCCATGCCTAAGAAAGATTG
#0334	gtggttatatgtacagtagtacgttAGGCAGTTTCATCTTCAG
#0338	Agaaagtgcttttggatcgtccggtgaaattgcagtaataccgatagtcc tcgatccgatgataagctgtc
#0346	gaccgataaaaagggtaatagatggCCACGATTTGATGAAAGAATAAC
#0347	ttgatactacaggagctggagatacTCGACCCGAGATCATATC
#0350	tcagcaatcatcattattccCCACGATTTGATGAAAGAATAAC
#0351	atatcgtaaaatttgtgagcTCGACCCGAGATCATATC
#0366	gggtggtatagcccttggtaCCACGATTTGATGAAAGAATAAC
#0367	ataacccctcacctttcaagTCGACCCGAGATCATATC
#0368	CTTGAAAGGTGAGGGGTTATATAC
#0369	TACCAAGGGCTATACCAC
#0457	atactaattgaagagaaagctggtggccaaaataggatattgattgtaga tgccacctgacgtctaagaa
#0458	Gagcttttctttcctctctctttttttttttcttgttacatattcctatat tgccacctgacgtctaagaa
#0769	tacgccaacttaagaccatg
#0770	tctcttttccatggtcatga
#0775	gtggtcgaccggcatgcaagctcccaaaaaaccgcgaagagctcc
#0776	gtggttatatgtacagtagtacgttctcaaccgcagaacccgg
#0777	gtggtcgaccggcatgcaagctcccggttggggtattaagagtacaatgc
#0778	gtggttatatgtacagtagtacgttcgcaattttcttgaacggtttttttc
#0815	aaaaaatccggaacaaaaaaaaccgcTCGACCCGAGATCATATC
#0816	ACCAACTAATTACTGCTCAACCGCACCACGATTTGATGAAAGAATAAC
#0817	acgttattctttcatcaaatcgtggTGCGGTTGAGCAGTAATTAG
#0818	ccacagtgatatgatctcgggtcgaGCGGTTTTTTTGTTCCGG
#0819	actctcactattttgttttttcgacccgagatcatatc
#0820	cgtatatgactgcacaagacccacgatttgatgaaagaataac
#0821	gtcttgtgcagtcatatac
#0822	aaaaacaaaatagtgagagtaatg
#0848	AATCTCACTAAAAGTAACATACAGTACCGATAAATCGAGATTGCAGAGTA tgccacctgacgtctaagaa
#0849	GTTCAATTATCTTAGAATGGATATGAATTAGTTAAAGCGGCTCGACCCAG tcgatccgatgataagctgtc
#0858	TCCATGTCCATGTCCATGTCATCATGGGCCGTGACAAGCGTCGCCGCGCA gccgaataaacttaaaattga

#0859	CCTCGACGGCCTCCAGTTCTTCGACCAACTGTTCGTGATCGTCATCCATT gagcttttctttcctctctct
#0932	attcaaatacgtcatcctat
#0933	actggatacttggtcgtcgt
#1049	gggtaccgggcccccctcgaggtcgacggtatcgataagcttgatatcgcctgcaggagaaaaaaagaatagtaataacaa tttg
#1050	acattatatgacccttctagacacttgtttcgactttctcacg
#1051	tcatcagcgtgagaaagtcgaaacaagtgtctagaagggtcatataatg
#1052	aaatacatacatacatatacaatgcaatcaaaattgtattccacgtc
#1053	ggtgacgtggaatacaattttgattgcattgtatatgtatgtatgtattttag
#1054	ccaccgcggtggcggccgctctagaactagtggatcccccgggctgcagggagctcgttctcccaatgaatggtc
#1158	ccctacaccacagcatacaaaaccgaatcaaaattgtattccacgtc
#1159	ggtgacgtggaatacaattttgattcggttttgtatgctgtgg

9. References

Abdulhay, N. J., McNally, C. P., Hsieh, L. J., Kasinathan, S., Keith, A., Estes, L. S., et al. (2020). Massively multiplex single-molecule oligonucleosome footprinting. *eLife* 9, e59404. doi: 10.7554/eLife.59404.

Adachi, Y., Usukura, J., and Yanagida, M. (1997). A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. *Genes to Cells* 2, 467–479. doi: 10.1046/j.1365-2443.1997.1350333.x.

Alcid, E. A., and Tsukiyama, T. (2014). ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. *Genes Dev.* 28, 2348–2360. doi: 10.1101/gad.250902.114.

Ang, X. L., and Wade Harper, J. (2005). SCF-mediated protein degradation and cell cycle control. *Oncogene* 24, 2860–2870. doi: 10.1038/sj.onc.1208614.

Ansari, A., Cheng, T.-H., and Gartenberg, M. R. (1999). Isolation of Selected Chromatin Fragments from Yeast by Site-Specific Recombinationin Vivo. *Methods* 17, 104–111. doi: 10.1006/meth.1998.0722.

Aparicio, J. G., Viggiani, C. J., Gibson, D. G., and Aparicio, O. M. (2004). The Rpd3-Sin3 Histone Deacetylase Regulates Replication Timing and Enables Intra-S Origin Control in *Saccharomyces cerevisiae*. *Mol Cell Biol* 24, 4769–4780. doi: 10.1128/MCB.24.11.4769-4780.2004.

Aparicio, O. M. (2013). Location, location, location: it's all in the timing for replication origins. *Genes Dev.* 27, 117–128. doi: 10.1101/gad.209999.112.

Araki, H., Nakanishi, N., Evans, B. R., Matsuzaki, H., Jayaram, M., and Oshima, Y. (1992). Site-specific recombinase, R, encoded by yeast plasmid pSR1. *Journal of Molecular Biology* 225, 25–37. doi: 10.1016/0022-2836(92)91023-I.

Ardui, S., Ameur, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Research* 46, 2159–2168. doi: 10.1093/nar/gky066.

Au, T. J., Rodriguez, J., Vincent, J. A., and Tsukiyama, T. (2011). ATP-Dependent Chromatin Remodeling Factors Tune S Phase Checkpoint Activity. *Mol Cell Biol* 31, 4454–4463. doi: 10.1128/MCB.05931-11.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37, W202–W208. doi: 10.1093/nar/gkp335.

Baldi, S., Korber, P., and Becker, P. B. (2020). Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nat Struct Mol Biol* 27, 109–118. doi: 10.1038/s41594-019-0368-x.

Barbaric, S., Luckenbach, T., Schmid, A., Blaschke, D., Hörz, W., and Korber, P. (2007). Redundancy of Chromatin Remodeling Pathways for the Induction of the Yeast PHO5 Promoter in Vivo. *Journal of Biological Chemistry* 282, 27610–27621. doi: 10.1074/jbc.M700623200.

Batrakou, D. G., Heron, E. D., and Nieduszynski, C. A. (2018). Rapid high-resolution measurement of DNA replication timing by droplet digital PCR. *Nucleic Acids Research* 46, e112–e112. doi: 10.1093/nar/gky590.

Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., et al. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14173–14178. doi: 10.1073/pnas.95.24.14173.

Bell, O., Tiwari, V. K., Thomä, N. H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. *Nat Rev Genet* 12, 554–564. doi: 10.1038/nrg3017.

Bell, S. P., and Dutta, A. (2002). DNA Replication in Eukaryotic Cells. *Annu. Rev. Biochem.* 71, 333–374. doi: 10.1146/annurev.biochem.71.110601.135425.

Bell, S. P., and Stillman, B. (1992). AIP.dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357, 7.

Bensimon, A., Simon, A., Chiffaudel, A., Croquette, V., Heslot, F., and Bensimon, D. (1994). Alignment and sensitive detection of DNA by a moving interface. *Science* 265, 2096–2098. doi: 10.1126/science.7522347.

Berbenetz, N. M., Nislow, C., and Brown, G. W. (2010). Diversity of Eukaryotic DNA Replication Origins Revealed by Genome-Wide Analysis of Chromatin Structure. *PLoS Genet* 6, e1001092. doi: 10.1371/journal.pgen.1001092.

Bialic, M., Coulon, V., Drac, M., Gostan, T., and Schwob, E. (2015). “Analyzing the Dynamics of DNA Replication in Mammalian Cells Using DNA Combing,” in *DNA Replication* Methods in Molecular Biology., eds. S. Vengrova and J. Dalgaard (New York, NY: Springer New York), 67–78. doi: 10.1007/978-1-4939-2596-4_4.

Bianco, S., Rodrigue, S., Murphy, B. D., and Gévry, N. (2015). “Global Mapping of Open Chromatin Regulatory Elements by Formaldehyde-Assisted Isolation of Regulatory Elements Followed by Sequencing (FAIRE-seq),” in *DNA-Protein Interactions* Methods in Molecular Biology., eds. B. P. Leblanc and S. Rodrigue (New York, NY: Springer New York), 261–272. doi: 10.1007/978-1-4939-2877-4_17.

Bleichert, F., Botchan, M. R., and Berger, J. M. (2015). Crystal structure of the eukaryotic origin recognition complex. *Nature* 519, 321–326. doi: 10.1038/nature14239.

Blow, J. J., Ge, X. Q., and Jackson, D. A. (2011). How dormant origins promote complete genome replication. *Trends in Biochemical Sciences* 36, 405–414. doi: 10.1016/j.tibs.2011.05.002.

Blow, J. J., Gillespie, P. J., Francis, D., and Jackson, D. A. (2001). Replication Origins in Xenopus Egg Extract Are 5–15 Kilobases Apart and Are Activated in Clusters That Fire at Different Times. *The Journal of Cell Biology* 152, 11.

Bolon, Y.-T., and Bielinsky, A.-K. (2006). The spatial arrangement of ORC binding modules determines the functionality of replication origins in budding yeast. *Nucleic Acids Research* 34, 5069–5080. doi: 10.1093/nar/gkl661.

Boos, D., and Ferreira, P. (2019). Origin Firing Regulations to Control Genome Replication Timing. *Genes* 10, 199. doi: 10.3390/genes10030199.

Bowers, J. L., Randell, J. C. W., Chen, S., and Bell, S. P. (2004). ATP Hydrolysis by ORC Catalyzes Reiterative Mcm2-7 Assembly at a Defined Origin of Replication. *Molecular Cell* 16, 967–978. doi: 10.1016/j.molcel.2004.11.038.

Boyer, L. A., Langer, M. R., Crowley, K. A., Tan, S., Denu, J. M., and Peterson, C. L. (2002). Essential Role for the SANT Domain in the Functioning of Multiple Chromatin Remodeling Enzymes. *Molecular Cell* 10, 935–942. doi: 10.1016/S1097-2765(02)00634-2.

Brahma, S., Udugama, M. I., Kim, J., Hada, A., Bhardwaj, S. K., Hailu, S. G., et al. (2017). INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat Commun* 8, 15616. doi: 10.1038/ncomms15616.

Breier, A. M., Chatterji, S., and Cozzarelli, N. R. (2004). Prediction of *Saccharomyces cerevisiae* replication origins. *Genome Biology*, 14.

Brown, C. R., and Boeger, H. (2014). Nucleosomal promoter variation generates gene expression noise. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17893–17898. doi: 10.1073/pnas.1417527111.

Brown, C. R., Mao, C., Falkovskaia, E., Jurica, M. S., and Boeger, H. (2013). Linking Stochastic Fluctuations in Chromatin Structure and Gene Expression. *PLoS Biol* 11, e1001621. doi: 10.1371/journal.pbio.1001621.

Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology* 109. doi: 10.1002/0471142727.mb2129s109.

Chanou, A., and Hamperl, S. (2021). Single-Molecule Techniques to Study Chromatin. *Front. Cell Dev. Biol.* 9, 699771. doi: 10.3389/fcell.2021.699771.

Chen, P., and Li, G. (2010). Dynamics of the higher-order structure of chromatin. *Protein Cell* 1, 967–971. doi: 10.1007/s13238-010-0130-y.

Chesnokov, I. N., Chesnokova, O. N., and Botchan, M. (2003). A cytokinetic function of *Drosophila* ORC6 protein resides in a domain distinct from its replication activity. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9150–9155. doi: 10.1073/pnas.1633580100.

Cimino, D. G., Howard, B., Gamper, I., Stephen, T., and Hearst, E. J. (1985). Psoralens as Photoactive Probes of Nucleic Acid Structure and Function: Organic Chemistry. Photochem. Biochem. Ann. Rev. Biochem. 54, 1151–1193.

Clapier, C. R., and Cairns, B. R. (2009). The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* 78, 273–304. doi: 10.1146/annurev.biochem.77.062706.153223.

Clapier, C. R., Iwasa, J., Cairns, B. R., and Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18, 407–422. doi: 10.1038/nrm.2017.26.

Clarke, J., Wu, H.-C., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009). Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotech* 4, 265–270. doi: 10.1038/nnano.2009.12.

Collins, N., Poot, R. A., Kukimoto, I., García-Jiménez, C., Dellaire, G., and Varga-Weisz, P. D. (2002). An ACF1–ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat Genet* 32, 627–632. doi: 10.1038/ng1046.

Conaway, R. C., and Conaway, J. W. (2009). The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends in Biochemical Sciences* 34, 71–77. doi: 10.1016/j.tibs.2008.10.010.

Copeland, W. C. (2012). Defects in mitochondrial DNA replication and human disease. *Critical Reviews in Biochemistry and Molecular Biology* 47, 64–74. doi: 10.3109/10409238.2011.632763.

Corona, D. F. V., and Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1677, 113–119. doi: 10.1016/j.bbaexp.2003.09.018.

Costa, A., and Diffley, J. F. X. (2022). The Initiation of Eukaryotic DNA Replication. *Annu. Rev. Biochem.* 91, 107–131. doi: 10.1146/annurev-biochem-072321-110228.

Coudreuse, D., and Nurse, P. (2010). Driving the cell cycle with a minimal CDK control network. *Nature* 468, 1074–1079. doi: 10.1038/nature09543.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2, 292–301. doi: 10.1038/35066075.

Cuperus, G., and Shore, D. (2002). Restoration of Silencing in *Saccharomyces cerevisiae* by Tethering of a Novel Sir2-Interacting Protein, Esc8. *Genetics* 162, 633–645. doi: 10.1093/genetics/162.2.633.

Czajkowsky, D. M., Liu, J., Hamlin, J. L., and Shao, Z. (2008). DNA Combing Reveals Intrinsic Temporal Disorder in the Replication of Yeast Chromosome VI. *Journal of Molecular Biology* 375, 12–19. doi: 10.1016/j.jmb.2007.10.046.

Dammann, R., Lucchini, R., Koller, T., and Sogo, J. M. (1993). Chromatin structures and transcription of rDNA in yeast *Saccharomyces cerevisiae*. *Nucleic Acids Research* 21, 2331–2338. doi: 10.1093/nar/21.10.2331.

Dang, W., and Bartholomew, B. (2007). Domain Architecture of the Catalytic Subunit in the ISW2-Nucleosome Complex. *Mol Cell Biol* 27, 8306–8317. doi: 10.1128/MCB.01351-07.

Das, S. P., Borrman, T., Liu, V. W. T., Yang, S. C.-H., Bechhoefer, J., and Rhind, N. (2015). Replication timing is regulated by the number of MCMs loaded at origins. *Genome Res.* 25, 1886–1892. doi: 10.1101/gr.195305.115.

De Carli, F., Menezes, N., Berrabah, W., Barbe, V., Genovesio, A., and Hyrien, O. (2018). High-Throughput Optical Mapping of Replicating DNA. *Small Methods* 2, 1800146. doi: 10.1002/smtd.201800146.

De Ioannes, P., Leon, V. A., Kuang, Z., Wang, M., Boeke, J. D., Hochwagen, A., et al. (2019). Structure and function of the Orc1 BAH-nucleosome complex. *Nat Commun* 10, 2894. doi: 10.1038/s41467-019-10609-y.

de Moura, A. P. S., Retkute, R., Hawkins, M., and Nieduszynski, C. A. (2010). Mathematical modelling of whole chromosome replication. *Nucleic Acids Research* 38, 5623–5633. doi: 10.1093/nar/gkq343.

Donovan, S., Harwood, J., Drury, L. S., and Diffley, J. F. X. (1997). Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5611–5616. doi: 10.1073/pnas.94.11.5611.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. *Science* 306, 1571–1573. doi: 10.1126/science.1103124.

Dorn, E. S., and Cook, J. G. (2011). Nucleosomes in the neighborhood: New roles for chromatin modifications in replication origin control. *Epigenetics* 6, 552–559. doi: 10.4161/epi.6.5.15082.

Dou, Y., and Gorovsky, M. A. (2000). Phosphorylation of Linker Histone H1 Regulates Gene Expression In Vivo by Creating a Charge Patch. *Molecular Cell* 6, 225–231. doi: 10.1016/S1097-2765(00)00024-1.

Duderstadt, K. E., and Berger, J. M. (2008). AAA+ ATPases in the Initiation of DNA Replication. *Critical Reviews in Biochemistry and Molecular Biology* 43, 163–187. doi: 10.1080/10409230802058296.

Dukaj, L., and Rhind, N. (2021). The capacity of origins to load MCM establishes replication timing patterns. *PLoS Genet* 17, e1009467. doi: 10.1371/journal.pgen.1009467.

Eaton, M. L., Galani, K., Kang, S., Bell, S. P., and MacAlpine, D. M. (2010). Conserved nucleosome positioning defines replication origins. *Genes Dev.* 24, 748–753. doi: 10.1101/gad.1913210.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-Time DNA Sequencing from Single Polymerase Molecules. Science 323, 133-138.

Elsasser, S., Lou, F., Wang, B., Campbell, J. L., and Jong, A. (1996). Interaction between yeast Cdc6 protein and B-type cyclin/Cdc28 kinases. *MBoC* 7, 1723–1735. doi: 10.1091/mbc.7.11.1723.

Fang, D., Lengronne, A., Shi, D., Forey, R., Skrzypczak, M., Ginalski, K., et al. (2017). Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes Dev.* 31, 2405–2415. doi: 10.1101/gad.306571.117.

Fangman, W. L. (1991). Activation of Replication Origins within Yeast Chromosomes. *Annu Rev Cell Biol*. 7:375-402. doi: 10.1146/annurev.cb.07.110191.002111. PMID: 1809350

Fazzio, T. G., Kooperberg, C., Goldmark, J. P., Neal, C., Basom, R., Delrow, J., et al. (2001). Widespread Collaboration of Isw2 and Sin3-Rpd3 Chromatin Remodeling Complexes in Transcriptional Repression. *Mol Cell Biol* 21, 6450–6460. doi: 10.1128/MCB.21.19.6450-6460.2001.

Feng, X., Noguchi, Y., Barbon, M., Stillman, B., Speck, C., and Li, H. (2021). The structure of ORC–Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6. *Nat Commun* 12, 3883. doi: 10.1038/s41467-021-24199-1.

Ferguson, B. M., and Fangman, W. L. (1992). A position effect on the time of replication origin activation in yeast. *Cell* 68, 333–339. doi: 10.1016/0092-8674(92)90474-Q.

Finch, J. T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 73, 1897–1901. doi: 10.1073/pnas.73.6.1897.

Fitzgerald, D. J., DeLuca, C., Berger, I., Gaillard, H., Sigrist, R., Schimmele, K., et al. (2004). Reaction cycle of the yeast Isw2 chromatin remodeling complex. *EMBO J* 23, 3836–3843. doi: 10.1038/sj.emboj.7600364.

Flaus, A. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Research* 34, 2887–2905. doi: 10.1093/nar/gkl295.

Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., et al. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 7, 461–465. doi: 10.1038/nmeth.1459.

Friedman, K. L., Brewer, B. J., and Fangman, W. L. (1997). Replication profile of *Saccharomyces cerevisiae* chromosome VI. *Genes Cells* 2, 667–678. doi: 10.1046/j.1365-2443.1997.1520350.x.

Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V., Brewer, B. J., and Fangman, W. L. (1996). Multiple determinants controlling activation of yeast replication origins late in S phase. *Genes Dev.* 10, 1595–1607. doi: 10.1101/gad.10.13.1595.

Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D., et al. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* 8, 358–366. doi: 10.1038/ncb1382.

Ganapathi, M., Palumbo, M. J., Ansari, S. A., He, Q., Tsui, K., Nislow, C., et al. (2011). Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Research* 39, 2032–2044. doi: 10.1093/nar/gkq1161.

Garzón, J., Ursich, S., Lopes, M., Hiraga, S., and Donaldson, A. D. (2019). Human RIF1-Protein Phosphatase 1 Prevents Degradation and Breakage of Nascent DNA on Replication Stalling. *Cell Reports* 27, 2558-2566.e4. doi: 10.1016/j.celrep.2019.05.002.

Gerhold, C. B., and Gasser, S. M. (2014). INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends in Cell Biology* 24, 619–631. doi: 10.1016/j.tcb.2014.06.004.

Gibney, E. R., and Nolan, C. M. (2010). Epigenetics and gene expression. *Heredity* 105, 4–13. doi: 10.1038/hdy.2010.54.

Goldmark, J. P., Fazzio, T. G., Estep, P. W., Church, G. M., and Tsukiyama, T. (2000). The Isw2 Chromatin Remodeling Complex Represses Early Meiotic Genes upon Recruitment by Ume6p. *Cell* 103, 423–433. doi: 10.1016/S0092-8674(00)00134-3.

Grewal, S. I. S., and Klar, A. J. S. (1996). Chromosomal Inheritance of Epigenetic States in Fission Yeast During Mitosis and Meiosis. *Cell* 86, 95–101. doi: 10.1016/S0092-8674(00)80080-X.

Griesenbeck, J., Boeger, H., Strattan, J. S., and Kornberg, R. D. (2003). Affinity Purification of Specific Chromatin Segments from Chromosomal Loci in Yeast. *Mol Cell Biol* 23, 9275–9282. doi: 10.1128/MCB.23.24.9275-9282.2003.

Griesenbeck, J., Boeger, H., Strattan, J. S., and Kornberg, R. D. (2004). Purification of defined chromosomal domains. *Methods Enzymol* 375, 170–178. doi: 10.1016/s0076-6879(03)75011-3.

Grüne, T., Brzeski, J., Eberharter, A., Clapier, C. R., Corona, D. F. V., Becker, P. B., et al. (2003). Crystal Structure and Functional Analysis of a Nucleosome Recognition Module of the Remodeling Factor ISWI. *Molecular Cell* 12, 449–460. doi: 10.1016/S1097-2765(03)00273-9.

Guilbaud, G., Rappailles, A., Baker, A., Chen, C.-L., Arneodo, A., Goldar, A., et al. (2011). Evidence for Sequential and Increasing Activation of Replication Origins along

Replication Timing Gradients in the Human Genome. *PLoS Comput Biol* 7, e1002322. doi: 10.1371/journal.pcbi.1002322.

Hamperl, S., Brown, C. R., Garea, A. V., Perez-Fernandez, J., Bruckmann, A., Huber, K., et al. (2014). Compositional and structural analysis of selected chromosomal domains from *Saccharomyces cerevisiae*. *Nucleic Acids Research* 42, e2–e2. doi: 10.1093/nar/gkt891.

Hamperl Stephan (2012) Compositional and structural analysis of selected chromosomal domains from Saccharomyces cerevisiae. urn:nbn:de:bvb:355-epub-265861 , doi: 10.5283/epub.26586

Hanson, C., Shen, C., and Hearst, J. (1976). Cross-linking of DNA in situ as a probe for chromatin structure. *Science* 193, 62–64. doi: 10.1126/science.935855.

Harris, L., Zalucki, O., and Piper, M. (2018). BrdU/EdU dual labeling to determine the cell-cycle dynamics of defined cellular subpopulations. *J Mol Hist* 49, 229–234. doi: 10.1007/s10735-018-9761-8.

Hayano, M., Kanoh, Y., Matsumoto, S., Renard-Guillet, C., Shirahige, K., and Masai, H. (2012). Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes Dev.* 26, 137–150. doi: 10.1101/gad.178491.111.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112. doi: 10.1038/nature07829.

Hennion, M., Arbona, J.-M., Cruaud, C., Proux, F., Tallec, B. L., Novikova, E., et al. (2018). Mapping DNA replication with nanopore sequencing. Genomics doi: 10.1101/426858.

Heun, P., Laroche, T., Raghuraman, M. K., and Gasser, S. M. (2001). The Positioning and Dynamics of Origins of Replication in the Budding Yeast Nucleus. *Journal of Cell Biology* 152, 385–400. doi: 10.1083/jcb.152.2.385.

Igo-Kemenes, T., Hörz, W., and Zachau, H. G. (1982). Chromatin. *Annu. Rev. Biochem.* 51, 89–121.

Ishii, H., Kadonaga, J. T., and Ren, B. (2015). MPE-seq, a new method for the genome-wide analysis of chromatin structure. *Proc Natl Acad Sci USA* 112, E3457–E3465. doi: 10.1073/pnas.1424804112.

Jain, M., Olsen, H. E., Paten, B., and Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol* 17, 239. doi: 10.1186/s13059-016-1103-0.

Jaspersen, S. L., Charles, J. F., and Morgan, D. O. (1999). Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Current Biology* 9, 227–236. doi: 10.1016/S0960-9822(99)80111-0.

Jenuwein, T. (2001). Translating the Histone Code. *Science* 293, 1074–1080. doi: 10.1126/science.1063127.

Kagalwala, M. N., Glaus, B. J., Dang, W., Zofall, M., and Bartholomew, B. (2004). Topography of the ISW2–nucleosome complex: insights into nucleosome spacing and chromatin remodeling. *EMBO J* 23, 2092–2104. doi: 10.1038/sj.emboj.7600220.

Karaoglu, H., Lee, C. M. Y., and Meyer, W. (2005). Survey of Simple Sequence Repeats in Completed Fungal Genomes. *Molecular Biology and Evolution* 22, 639–649. doi: 10.1093/molbev/msi057.

Kasianowicz, J. J., Brandin, E., Branton, D., and Deamer, D. W. (1996). Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences* 93, 13770–13773. doi: 10.1073/pnas.93.24.13770.

Kaykov, A., Taillefumier, T., Bensimon, A., and Nurse, P. (2016). Molecular Combing of Single DNA Molecules on the 10 Megabase Scale. *Sci Rep* 6, 19636. doi: 10.1038/srep19636.

Kleinendorst, R. W. D., Barzaghi, G., Smith, M. L., Zaugg, J. B., and Krebs, A. R. (2021). Genome-wide quantification of transcription factor binding at single-DNA-molecule resolution using methyl-transferase footprinting. *Nat Protoc* 16, 5673–5706. doi: 10.1038/s41596-021-00630-1.

Knott, S. R. V., Peace, J. M., Ostrow, A. Z., Gan, Y., Rex, A. E., Viggiani, C. J., et al. (2012). Forkhead Transcription Factors Establish Origin Timing and Long-Range Clustering in *S. cerevisiae*. *Cell* 148, 99–111. doi: 10.1016/j.cell.2011.12.012.

Knott, S. R. V., Viggiani, C. J., Tavaré, S., and Aparicio, O. M. (2009). Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in *Saccharomyces cerevisiae*. *Genes Dev.* 23, 1077–1090. doi: 10.1101/gad.1784309.

Kobayashi, W., and Kurumizaka, H. (2019). Structural transition of the nucleosome during chromatin remodeling and transcription. *Current Opinion in Structural Biology* 59, 107–114. doi: 10.1016/j.sbi.2019.07.011.

Kornberg, R. D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. *Science* 184, 868–871. doi: 10.1126/science.184.4139.868

Kornberg, R. D., and Lorch, Y. (2020). Primary Role of the Nucleosome. *Molecular Cell* 79, 371–375. doi: 10.1016/j.molcel.2020.07.020.

Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C. L., Pugh, B. F., et al. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* 167, 709-721.e12. doi: 10.1016/j.cell.2016.09.045.

Kunkel, T. A., and Burgers, P. M. (2008). Dividing the workload at a eukaryotic replication fork. *Trends in Cell Biology* 18, 521–527. doi: 10.1016/j.tcb.2008.08.005.

Lacroix, J., Pélofy, S., Blatché, C., Pillaire, M.-J., Huet, S., Chapuis, C., et al. (2016). Analysis of DNA Replication by Optical Mapping in Nanochannels. *Small* 12, 5963–5970. doi: 10.1002/smll.201503795.

Lee, D. G., and Bell, S. P. (1997). Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol Cell Biol* 17, 7159–7168. doi: 10.1128/MCB.17.12.7159.

Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998). Genetic instabilities in human cancers. *Nature* 396, 643–649. doi: 10.1038/25292.

Leonard AC, Méchali M. DNA replication origins. (2013). *Cold Spring Harb Perspect Biol*. Oct 1;5(10):a010116. doi: 10.1101/cshperspect.a010116.

Li, J. J., and Herskowitz, I. (1993). Isolation of *ORC6* , a Component of the Yeast Origin Recognition Complex by a One-Hybrid System. *Science* 262, 1870–1874. doi: 10.1126/science.8266075.

Li, N., Lam, W. H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., et al. (2018). Structure of the origin recognition complex bound to DNA replication origin. *Nature* 559, 217–222. doi: 10.1038/s41586-018-0293-x.

Li, S., Wasserman, M. R., Yurieva, O., Bai, L., O'Donnell, M. E., and Liu, S. (2022). Nucleosome-directed replication origin licensing independent of a consensus DNA sequence. *Nat Commun* 13, 4947. doi: 10.1038/s41467-022-32657-7.

Li, Y., and Tollefsbol, T. O. (2011). "DNA Methylation Detection: Bisulfite Genomic Sequencing Analysis," in *Epigenetics Protocols* Methods in Molecular Biology., ed. T. O. Tollefsbol (Totowa, NJ: Humana Press), 11–21. doi: 10.1007/978-1-61779-316-5_2.

Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y. C., Dunlap, D., et al. (2006). Direct Observation of DNA Distortion by the RSC Complex. *Molecular Cell* 21, 417–425. doi: 10.1016/j.molcel.2005.12.013.

Lian, H.-Y., Robertson, E. D., Hiraga, S., Alvino, G. M., Collingwood, D., McCune, H. J., et al. (2011). The effect of Ku on telomere replication time is mediated by telomere length but is independent of histone tail acetylation. *MBoC* 22, 1753–1765. doi: 10.1091/mbc.e10-06-0549.

Lipford, J. R., and Bell, S. P. (2001a). Nucleosomes Positioned by ORC Facilitate the Initiation of DNA Replication. *Molecular Cell* 7, 21–30. doi: 10.1016/S1097-2765(01)00151-4.

Lipford, J. R., and Bell, S. P. (2001b). Nucleosomes Positioned by ORC Facilitate the Initiation of DNA Replication. *Molecular Cell* 7, 21–30. doi: 10.1016/S1097-2765(01)00151-4.

Liu, Y., Rosikiewicz, W., Pan, Z., Jillette, N., Wang, P., Taghbalout, A., et al. (2021). DNA methylation-calling tools for Oxford Nanopore sequencing: a survey and human epigenome-wide evaluation. *Genome Biol* 22, 295. doi: 10.1186/s13059-021-02510-z.

Lomonosov, M., Anand, S., Sangrithi, M., Davies, R., and Venkitaraman, A. R. (2003). Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein. *Genes Dev.* 17, 3017–3022. doi: 10.1101/gad.279003.

Lõoke, M., Kristjuhan, K., Värv, S., and Kristjuhan, A. (2013). Chromatin-dependent and -independent regulation of DNA replication origin activation in budding yeast. *EMBO Rep* 14, 191–198. doi: 10.1038/embor.2012.196.

Lucchini, R. (2001). Nucleosome positioning at the replication fork. *The EMBO Journal* 20, 7294–7302. doi: 10.1093/emboj/20.24.7294.

Luger, K.,Mäder, W. A., Richmond, K. R., Sargent, F. D., and Richmod, J. T. (1997). Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. Nature 389, 251–260. doi: 10.1038/38444

Luk, E., Ranjan, A., FitzGerald, P. C., Mizuguchi, G., Huang, Y., Wei, D., et al. (2010). Stepwise Histone Replacement by SWR1 Requires Dual Activation with Histone H2A.Z and Canonical Nucleosome. *Cell* 143, 725–736. doi: 10.1016/j.cell.2010.10.019.

Lusser, A., and Kadonaga, J. T. (2003). Chromatin remodeling by ATP-dependent molecular machines. *Bioessays* 25, 1192–1200. doi: 10.1002/bies.10359.

Lynch, K. L., Alvino, G. M., Kwan, E. X., Brewer, B. J., and Raghuraman, M. K. (2019). The effects of manipulating levels of replication initiation factors on origin firing efficiency in yeast. *PLoS Genet* 15, e1008430. doi: 10.1371/journal.pgen.1008430.

Ma, W., Noble, W. S., and Bailey, T. L. (2014). Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc* 9, 1428–1450. doi: 10.1038/nprot.2014.083.

Machanick, P., and Bailey, T. L. (2011). MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* 27, 1696–1697. doi: 10.1093/bioinformatics/btr189.

Maier, B., Bensimon, D., and Croquette, V. (2000). Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proceedings of the National Academy of Sciences* 97, 12002–12007. doi: 10.1073/pnas.97.22.12002.

Mantiero, D., Mackenzie, A., Donaldson, A., and Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast: Limiting replication factors execute temporal programme. *The EMBO Journal* 30, 4805–4814. doi: 10.1038/emboj.2011.404.

Marfella, C. G. A., and Imbalzano, A. N. (2007). The Chd family of chromatin remodelers. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 618, 30–40. doi: 10.1016/j.mrfmmm.2006.07.012.

Marquardt, S., Escalante-Chong, R., Pho, N., Wang, J., Churchman, L. S., Springer, M., et al. (2014). A Chromatin-Based Mechanism for Limiting Divergent Noncoding Transcription. *Cell* 157, 1712–1723. doi: 10.1016/j.cell.2014.04.036.

Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., and Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol* 6, 44. doi: 10.1186/1471-2121-6-44.

McGuffee, S. R., Smith, D. J., and Whitehouse, I. (2013). Quantitative, Genome-Wide Analysis of Eukaryotic Replication Initiation and Termination. *Molecular Cell* 50, 123–135. doi: 10.1016/j.molcel.2013.03.004.

Michalet, X. (1997). Dynamic Molecular Combing: Stretching the Whole Human Genome for High-Resolution Studies. *Science* 277, 1518–1523. doi: 10.1126/science.277.5331.1518.

Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X., and Costa, A. (2019). Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* 575, 704–710. doi: 10.1038/s41586-019-1768-0.

Miranda, T. B., Kelly, T. K., Bouazoune, K., and Jones, P. A. (2010). Methylation-Sensitive Single-Molecule Analysis of Chromatin Structure. *Current Protocols in Molecular Biology* 89. doi: 10.1002/0471142727.mb2117s89.

Miura, H., Takahashi, S., Shibata, T., Nagao, K., Obuse, C., Okumura, K., et al. (2020). Mapping replication timing domains genome wide in single mammalian cells with single-cell DNA replication sequencing. *Nat Protoc* 15, 4058–4100. doi: 10.1038/s41596-020-0378-5.

Miyake, T., Loch, C. M., and Li, R. (2002). Identification of a Multifunctional Domain in Autonomously Replicating Sequence-Binding Factor 1 Required for Transcriptional Activation, DNA Replication, and Gene Silencing. *Mol Cell Biol* 22, 505–516. doi: 10.1128/MCB.22.2.505-516.2002.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., and Wu, C. (2004). ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science* 303, 343–348. doi: 10.1126/science.1090701.

Moreau, J.-L., Lee, M., Mahachi, N., Vary, J., Mellor, J., Tsukiyama, T., et al. (2003). Regulated Displacement of TBP from the PHO8 Promoter In Vivo Requires Cbf1 and the Isw1 Chromatin Remodeling Complex. *Molecular Cell* 11, 1609–1620. doi: 10.1016/S1097-2765(03)00184-9.

Morillon, A., Karabetsou, N., Nair, A., and Mellor, J. (2005). Dynamic Lysine Methylation on Histone H3 Defines the Regulatory Phase of Gene Transcription. *Molecular Cell* 18, 723–734. doi: 10.1016/j.molcel.2005.05.009.

Moyer, S. E., Lewis, P. W., and Botchan, M. R. (2006). Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10236–10241. doi: 10.1073/pnas.0602400103.

Müller, C. A., Boemo, M. A., Spingardi, P., Kessler, B. M., Kriaucionis, S., Simpson, J. T., et al. (2019). Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nat Methods* 16, 429–436. doi: 10.1038/s41592-019-0394-y.

Nieduszynski, C. A., Knox, Y., and Donaldson, A. D. (2006). Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev.* 20, 1874–1879. doi: 10.1101/gad.385306.

Nieminuszczy, J., Schwab, R. A., and Niedzwiedz, W. (2016). The DNA fibre technique – tracking helicases at work. *Methods* 108, 92–98. doi: 10.1016/j.ymeth.2016.04.019.

Olins, A. L., and Olins, D. E. (1974). Spheroid Chromatin Units (ngr Bodies). *Science* 183, 330–332. doi: 10.1126/science.183.4122.330.

On KF, Beuron F, Frith D, Snijders AP, Morris EP, Diffley JF. Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication. (2014) *EMBO* J. Mar 18;33(6):605-20. doi: 10.1002/embj.201387369. Epub 2014 Feb 24.

Papamichos-Chronakis, M., and Peterson, C. L. (2008). The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. *Nat Struct Mol Biol* 15, 338–345. doi: 10.1038/nsmb.1413.

Papamichos-Chronakis, M., Watanabe, S., and Rando, O. J. (2012). Global regulation of H2A.Z localization by the INO80 chromatin remodeling enzyme is essential for genome integrity. *Cell* 24, 200-213.

Patel, A. B., Moore, C. M., Greber, B. J., Luo, J., Zukin, S. A., Ranish, J., et al. (2019). Architecture of the chromatin remodeler RSC and insights into its nucleosome engagement. *eLife* 8, e54449. doi: 10.7554/eLife.54449.

Peng, C., Luo, H., Zhang, X., and Gao, F. (2015). Recent advances in the genome-wide study of DNA replication origins in yeast. *Front. Microbiol.* 6. doi: 10.3389/fmicb.2015.00117.

Perkins, G., and Diffley, J. F. X. (1998). Nucleotide-Dependent Prereplicative Complex Assembly by Cdc6p, a Homolog of Eukaryotic and Prokaryotic Clamp-Loaders. *Molecular Cell* 2, 23–32. doi: 10.1016/S1097-2765(00)80110-0.

Peters, J.-M. (2006). The anaphase promoting complex/ cyclosome: a machine designed to destroy. *Nat Rev Mol Cell Biol*. Sep;7(9):644-56. doi: 10.1038/nrm1988.

Poveda, A. M., Le Clech, M., and Pasero, P. (2010). Transcription and replication: Breaking the rules of the road causes genomic instability. *Transcription* 1, 99–102. doi: 10.4161/trns.1.2.12665.

Pozarowski, P., and Darzynkiewicz, Z. (2004). “Analysis of Cell Cycle by Flow Cytometry,” in *Checkpoint Controls and Cancer* (New Jersey: Humana Press), 301–312. doi: 10.1385/1-59259-811-0:301.

Pryde, F., Jain, D., Kerr, A., Curley, R., and Mariotti, F. R. (2009). H3 K36 Methylation Helps Determine the Timing of Cdc45 Association with Replication Origins. *PLoS ONE* 4, 13.

Ray Chaudhuri, A., Callen, E., Ding, X., Gogola, E., Duarte, A. A., Lee, J.-E., et al. (2016). Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature* 535, 382–387. doi: 10.1038/nature18325.

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., and Diffley, J. F. X. (2009). Concerted Loading of Mcm2–7 Double Hexamers around DNA during DNA Replication Origin Licensing. *Cell* 139, 719–730. doi: 10.1016/j.cell.2009.10.015.

Reynolds, A. E., and McCARROLL, R. M. (1989). Time of Replication of ARS Elements along Yeast Chromosome III. *Mol. Cell. Biol.* 9, 7.

Rhind, N., and Gilbert, D. M. (2013). DNA Replication Timing. *Cold Spring Harbor Perspectives in Biology* 5, a010132–a010132. doi: 10.1101/cshperspect.a010132.

Robinson, P. J. J., Fairall, L., Huynh, V. A. T., and Rhodes, D. (2006). EM measurements define the dimensions of the “30-nm” chromatin fiber: Evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6506–6511. doi: 10.1073/pnas.0601212103.

Rossi, M. J., Kuntala, P. K., Lai, W. K. M., Yamada, N., Badjatia, N., Mittal, C., et al. (2021). A high-resolution protein architecture of the budding yeast genome. *Nature* 592, 309–314. doi: 10.1038/s41586-021-03314-8.

Saha, A., Wittmeyer, J., and Cairns, B. R. (2002). Chromatin remodeling by RSC involves ATP-dependent DNA translocation. *Genes Dev.* 16, 2120–2134. doi: 10.1101/gad.995002.

Saha, A., Wittmeyer, J., and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* 7, 437–447. doi: 10.1038/nrm1945.

Sato, M., Gotow, T., You, Z., Komamura-Kohno, Y., Uchiyama, Y., Yabuta, N., et al. (2000). Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex. *Journal of Molecular Biology* 300, 421–431. doi: 10.1006/jmbi.2000.3865.

Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436, 138–141. doi: 10.1038/nature03686.

Schmidt, J. M., Yang, R., Kumar, A., Hunker, O., Seebacher, J., and Bleichert, F. (2022). A mechanism of origin licensing control through autoinhibition of S. cerevisiae ORC·DNA·Cdc6. *Nat Commun* 13, 1059. doi: 10.1038/s41467-022-28695-w.

Schones, D. E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., et al. (2008). Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell* 132, 887–898. doi: 10.1016/j.cell.2008.02.022.

Schwab, R. A., Blackford, A. N., and Niedzwiedz, W. (2010). ATR activation and replication fork restart are defective in FANCM-deficient cells. *EMBO J* 29, 806–818. doi: 10.1038/emboj.2009.385.

Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003). Involvement of Actin-Related Proteins in ATP-Dependent Chromatin Remodeling. *Molecular Cell* 12, 147–155. doi: 10.1016/S1097-2765(03)00264-8.

Shimada, K., Oma, Y., Schleker, T., Kugou, K., Ohta, K., Harata, M., et al. (2008). Ino80 Chromatin Remodeling Complex Promotes Recovery of Stalled Replication Forks. *Current Biology* 18, 566–575. doi: 10.1016/j.cub.2008.03.049.

Shipony, Z., Marinov, G. K., Swaffer, M. P., Sinnott-Armstrong, N. A., Skotheim, J. M., Kundaje, A., et al. (2020). Long-range single-molecule mapping of chromatin accessibility in eukaryotes. *Nat. Methods* 17, 319–327. doi: 10.1038/s41592-019-0730-2.

Simpson, R. T. (1990). Nucleosome positioning can affect the function of a cis- acting DNA element in vivo. *Nature* **343**, 387–389. doi: 10.1038/343387a0.

Sogo, J. M. (2002). Fork Reversal and ssDNA Accumulation at Stalled Replication Forks Owing to Checkpoint Defects. *Science* 297, 599–602. doi: 10.1126/science.1074023.

Sogo, J. M., Ness, P. J., Widmer, R. M., Parish, R. W., and Koller, Th. (1984). Psoralen-crosslinking of DNA as a probe for the structure of active nucleolar chromatin. *Journal of Molecular Biology* 178, 897–919. doi: 10.1016/0022-2836(84)90318-8.

Soriano, I., Morafraile, E. C., Vázquez, E., Antequera, F., and Segurado, M. (2014). Different nucleosomal architectures at early and late replicating origins in Saccharomyces cerevisiae. *BMC Genomics* 15, 791. doi: 10.1186/1471-2164-15-791.

Steger, D. J., Haswell, E. S., Miller, A. L., Wente, S. R., and O'Shea, E. K. (2003). Regulation of Chromatin Remodeling by Inositol Polyphosphates. *Science* 299, 114–116. doi: 10.1126/science.1078062.

Stergachis, A. B., Debo, B. M., Haugen, E., Churchman, L. S., and Stamatoyannopoulos, J. A. (2020). Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science* 368, 1449–1454. doi: 10.1126/science.aaz1646.

Stevenson, J. B., and Gottschling, D. E. (1999). Telomeric chromatin modulates replication timing near chromosome ends. *Genes & Development* 13, 146–151. doi: 10.1101/gad.13.2.146.

Stinchcomb, D. T., Mann, C., Selker, E., and Davis, R. W. (1981). “DNA sequences that allow the replication and segregation of yeast chromosomes,” in The Initiation of DNA Replication, ed. D. S. Ray, (New York, NY: Academic Press), 473–488. doi: 10.1016/b978-0-12-583580-0.50036-4

Strahl, B. D., and Allis, C. D. (2000). The Language of Covalent Histone Modifications. *Nature* 403, 41–45. doi: 10.1038/474125

Strohner, R., Wachsmuth, M., Dachauer, K., Mazurkiewicz, J., Hochstatter, J., Rippe, K., et al. (2005). A “loop recapture” mechanism for ACF-dependent nucleosome remodeling. *Nat Struct Mol Biol* 12, 683–690. doi: 10.1038/nsmb966.

Sun, H. B., Shen, J., and Yokota, H. (2000). Size-Dependent Positioning of Human Chromosomes in Interphase Nuclei. *Biophysical Journal* 79, 184–190. doi: 10.1016/S0006-3495(00)76282-5.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev.* 17, 1153–1165. doi: 10.1101/gad.1065903.

Tanabe, H., Habermann, F. A., Solovei, I., Cremer, M., and Cremer, T. (2002). Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 504, 37–45. doi: 10.1016/S0027-5107(02)00077-5.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445, 328–332. doi: 10.1038/nature05465.

Tang, L., Nogales, E., and Ciferri, C. (2010). Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Progress in Biophysics and Molecular Biology* 102, 122–128. doi: 10.1016/j.pbiomolbio.2010.05.001.

Tazumi, A., Fukuura, M., Nakato, R., Kishimoto, A., Takenaka, T., Ogawa, S., et al. (2012). Telomere-binding protein Taz1 controls global replication timing through its localization near late replication origins in fission yeast. *Genes Dev.* 26, 2050–2062. doi: 10.1101/gad.194282.112.

Técher, H., Koundrioukoff, S., Azar, D., Wilhelm, T., Carignon, S., Brison, O., et al. (2013). Replication Dynamics: Biases and Robustness of DNA Fiber Analysis. *Journal of Molecular Biology* 425, 4845–4855. doi: 10.1016/j.jmb.2013.03.040.

Thangavel, S., Berti, M., Levikova, M., Pinto, C., Gomathinayagam, S., Vujanovic, M., et al. (2015). DNA2 drives processing and restart of reversed replication forks in human cells. *Journal of Cell Biology* 208, 545–562. doi: 10.1083/jcb.201406100.

Theis, J. F., and Newlon, C. S. (1997). The *ARS309* chromosomal replicator of *Saccharomyces cerevisiae* depends on an exceptional ARS consensus sequence. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10786–10791. doi: 10.1073/pnas.94.20.10786.

Thoma F, Koller T. Influence of histone H1 on chromatin structure. (1977) *Cell.* Sep;12(1):101-7. doi: 10.1016/0092-8674(77)90188-x.

Topal, S., Van, C., Xue, Y., Carey, M. F., and Peterson, C. L. (2020). INO80C Remodeler Maintains Genomic Stability by Preventing Promiscuous Transcription at Replication Origins. *Cell Reports* 32, 108106. doi: 10.1016/j.celrep.2020.108106.

Trachtulcov, P., Janatov, I., Kohlwein, S. D., and Ha?ek, J. (2000). Saccharomyces cerevisiae geneISW2 encodes a microtubule-interacting protein required for premeiotic DNA replication. *Yeast* 16, 35–47. doi: 10.1002/(SICI)1097-0061(20000115)16:1

Travers, A., Hiriart, E., Churcher, M., Caserta, M., and Di Mauro, E. (2010). The DNA Sequence-dependence of Nucleosome Positioning *in vivo* and *in vitro*. *Journal of Biomolecular Structure and Dynamics* 27, 713–724. doi: 10.1080/073911010010524942.

Tremethick, D. J. (2007). Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber. *Cell* 128, 651–654. doi: 10.1016/j.cell.2007.02.008.

Tsukiyama, T., Palmer, J., Landel, C. C., Shiloach, J., and Wu, C. (1999). Characterization of the Imitation Switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes & Development* 13, 686–697. doi: 10.1101/gad.13.6.686.

Tuduri, S., Tourrière, H., and Pasero, P. (2010). Defining replication origin efficiency using DNA fiber assays. *Chromosome Res* 18, 91–102. doi: 10.1007/s10577-009-9098-y.

Tyagi, M., Imam, N., Verma, K., and Patel, A. K. (2016). Chromatin remodelers: We are the drivers!! *Nucleus* 7, 388–404. doi: 10.1080/19491034.2016.1211217.

Tye, B. K. (1999). MCM Proteins in DNA Replication. *Annu. Rev. Biochem*. 68:649–686

Tyler, J. K., Adams, C. R., Chen, S.-R., Kobayashi, R., Kamakaka, R. T., and Kadonaga, J. T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* 402, 555–560. doi: 10.1038/990147.

Udugama, M., Sabri, A., and Bartholomew, B. (2011). The INO80 ATP-Dependent Chromatin Remodeling Complex Is a Nucleosome Spacing Factor. *Mol Cell Biol* 31, 662–673. doi: 10.1128/MCB.01035-10.

Vary, J. C., Gangaraju, V. K., Qin, J., Landel, C. C., Kooperberg, C., Bartholomew, B., et al. (2003). Yeast Isw1p Forms Two Separable Complexes In Vivo. *Mol. Cell. Biol.* 23, 12.

Vincent, J. A., Kwong, T. J., and Tsukiyama, T. (2008). ATP-dependent chromatin remodeling shapes the DNA replication landscape. *Nat Struct Mol Biol* 15, 477–484. doi: 10.1038/nsmb.1419.

Vindigni, A., and Lopes, M. (2017). Combining electron microscopy with single molecule DNA fiber approaches to study DNA replication dynamics. *Biophysical Chemistry* 225, 3–9. doi: 10.1016/j.bpc.2016.11.014.

Vujanovic, M., Krietsch, J., Raso, M. C., Terraneo, N., Zellweger, R., Schmid, J. A., et al. (2017). Replication Fork Slowing and Reversal upon DNA Damage Require PCNA Polyubiquitination and ZRANB3 DNA Translocase Activity. *Molecular Cell* 67, 882-890.e5. doi: 10.1016/j.molcel.2017.08.010.

Wang, W., Klein, K., Proesmans, K., Yang, H., Marchal, C., Zhu, X., et al. (2020). Genome-Wide Mapping of Human DNA Replication by Optical Replication Mapping Supports a Stochastic Model of Eukaryotic Replication Timing. *bioRxiv*, 2020.08.24.263459. doi: 10.1101/2020.08.24.263459.

Wang, Y., Wang, A., Liu, Z., Thurman, A. L., Powers, L. S., Zou, M., et al. (2019). Single-molecule long-read sequencing reveals the chromatin basis of gene expression. *Genome Res.* 29, 1329–1342. doi: 10.1101/gr.251116.119.

Weinreich, M., Liang, C., and Stillman, B. (1999). The Cdc6p nucleotide-binding motif is required for loading Mcm proteins onto chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 96, 441–446. doi: 10.1073/pnas.96.2.441.

Weiβ, M., Chanou, A., Schauer, T., Tvardovskiy, A., Meiser, S., König, A.-C., et al. (2023). Single-copy locus proteomics of early- and late-firing DNA replication origins identifies a role of Ask1/DASH complex in replication timing control. *Cell Reports* 42, 112045. doi: 10.1016/j.celrep.2023.112045.

Wellinger, R. E., and Sogo, J. M. (1998). In Vivo Mapping of Nucleosomes Using Psoralen–DNA Crosslinking and Primer Extension. *Nucleic Acids Res.* 26,1544–1545. doi: 10.1093/nar/26.6.1544

Wilmes, G. M., and Bell, S. P. (2002). The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 101–106. doi: 10.1073/pnas. 012578499

Wittner, M., Hamperl, S., Stöckl, U., Seufert, W., Tschochner, H., Milkereit, P., et al. (2011). Establishment and Maintenance of Alternative Chromatin States at a Multicopy Gene Locus. *Cell* 145, 543–554. doi: 10.1016/j.cell.2011.03.051.

Woodward, A. M., Göhler, T., Luciani, M. G., Oehlmann, M., Ge, X., Gartner, A., et al. (2006). Excess Mcm2–7 license dormant origins of replication that can be used under conditions of replicative stress. *Journal of Cell Biology* 173, 673–683. doi: 10.1083/jcb.200602108.

Wuite, G. J. L., Smith, S. B., Young, M., Keller, D., and Bustamante, C. (2000). Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* 404, 103–106. doi: 10.1038/35003614.

Xue, Y., Van, C., Pradhan, S. K., Su, T., Gehrke, J., Kuryan, B. G., et al. (2015). The Ino80 complex prevents invasion of euchromatin into silent chromatin. *Genes Dev.* 29, 350–355. doi: 10.1101/gad.256255.114.

Yabuki, N., Terashima, H., and Kitada, K. (2002). Mapping of early firing origins on a replication profile of budding yeast: Genome-wide mapping of early firing origins. *Genes to Cells* 7, 781–789. doi: 10.1046/j.1365-2443.2002.00559.x.

Yang, S. C., Rhind, N., and Bechhoefer, J. (2010). Modeling genome-wide replication kinetics reveals a mechanism for regulation of replication timing. *Mol Syst Biol* 6, 404. doi: 10.1038/msb.2010.61.

Yang, Y., Sebra, R., Pullman, B. S., Qiao, W., Peter, I., Desnick, R. J., et al. (2015). Quantitative and multiplexed DNA methylation analysis using long-read single-molecule real-time bisulfite sequencing (SMRT-BS). *BMC Genomics* 16, 350. doi: 10.1186/s12864-015-1572-7.

Yao, W., King, D. A., Beckwith, S. L., Gowans, G. J., Yen, K., Zhou, C., et al. (2016). The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Mol Cell Biol* 36, 979–991. doi: 10.1128/MCB.00801-15.

Yen, K., Vinayachandran, V., and Pugh, B. F. (2013). SWR-C and INO80 Chromatin Remodelers Recognize Nucleosome-free Regions Near +1 Nucleosomes. *Cell* 154, 1246–1256. doi: 10.1016/j.cell.2013.08.043.

Zellweger, R., and Lopes, M. (2018). “Dynamic Architecture of Eukaryotic DNA Replication Forks In Vivo, Visualized by Electron Microscopy,” in *Genome Instability* Methods in Molecular Biology., eds. M. Muzi-Falconi and G. W. Brown (New York, NY: Springer New York), 261–294. doi: 10.1007/978-1-4939-7306-4_19.

Zhang, H., Roberts, D. N., and Cairns, B. R. (2005). Genome-Wide Dynamics of Htz1, a Histone H2A Variant that Poises Repressed/Basal Promoters for Activation through Histone Loss. *Cell* 123, 219–231. doi: 10.1016/j.cell.2005.08.036.

Zook, J. M., Hansen, N. F., Olson, N. D., Chapman, L., Mullikin, J. C., Xiao, C., et al. (2020). A robust benchmark for detection of germline large deletions and insertions. *Nat Biotechnol* 38, 1347–1355. doi: 10.1038/s41587-020-0538-8.