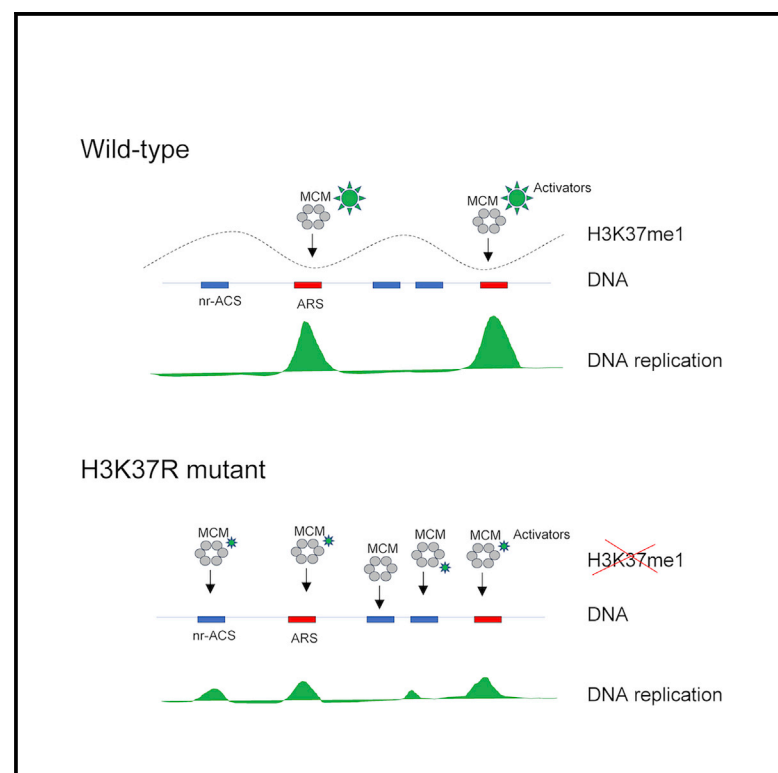


# Methylation of histone H3 at lysine 37 by Set1 and Set2 prevents spurious DNA replication

## Graphical abstract



## Authors

Helena Santos-Rosa,  
Gonzalo Millán-Zambrano,  
Namshik Han, ..., Kostantinos Tzelepis,  
Till Bartke, Tony Kouzarides

## Correspondence

hsrs2@gurdon.cam.ac.uk (H.S.-R.),  
t.kouzarides@gurdon.cam.ac.uk (T.K.)

## In brief

Genomes are the template for DNA replication and transcription, which timely coincide. To minimize detrimental collisions between both molecular machineries, only a subset of all replication origins performs as such. Santos-Rosa et al. show that the histone H3K37me1 prevents the establishment of unconventional origins, protecting the genome from spurious replication events.

## Highlights

- Set1p and Set2p mono-methylate H3K37
- H3K37me1 hinders MCM replicative helicase interaction with chromatin
- Low levels of H3K37me1 at replication origins facilitate origin establishment
- H3K37me1 abrogation results in MCM recruitment and licensing of non-replicative ACS

Article

# Methylation of histone H3 at lysine 37 by Set1 and Set2 prevents spurious DNA replication

Helena Santos-Rosa,<sup>1,8,\*</sup> Gonzalo Millán-Zambrano,<sup>1,2,8</sup> Namshik Han,<sup>1,3,8</sup> Tommaso Leonardi,<sup>1,4</sup> Marie Klimontova,<sup>1</sup> Simona Nasiscionyte,<sup>5</sup> Luca Pandolfini,<sup>1,6</sup> Kostantinos Tzelepis,<sup>1,7</sup> Till Bartke,<sup>5</sup> and Tony Kouzarides<sup>1,9,\*</sup>

<sup>1</sup>The Gurdon Institute and Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK

<sup>2</sup>Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER), 41092 Sevilla, Spain

<sup>3</sup>Milner Therapeutics Institute, University of Cambridge, Cambridge CB2 0AW, UK

<sup>4</sup>Center for Genomic Science Istituto Italiano di Tecnologia (IIT), 20139 Milano, Italy

<sup>5</sup>Institute of Functional Epigenetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany

<sup>6</sup>Istituto Italiano di Tecnologia (IIT), Center for Human Technologies (CHT), 16152 Genova, Italy

<sup>7</sup>Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK

<sup>8</sup>These authors contributed equally

<sup>9</sup>Lead contact

\*Correspondence: [hsrs2@gurdon.cam.ac.uk](mailto:hsrs2@gurdon.cam.ac.uk) (H.S.-R.), [t.kouzarides@gurdon.cam.ac.uk](mailto:t.kouzarides@gurdon.cam.ac.uk) (T.K.)

<https://doi.org/10.1016/j.molcel.2021.04.021>

## SUMMARY

DNA replication initiates at genomic locations known as origins of replication, which, in *S. cerevisiae*, share a common DNA consensus motif. Despite being virtually nucleosome-free, origins of replication are greatly influenced by the surrounding chromatin state. Here, we show that histone H3 lysine 37 mono-methylation (H3K37me1) is catalyzed by Set1p and Set2p and that it regulates replication origin licensing. H3K37me1 is uniformly distributed throughout most of the genome, but it is scarce at replication origins, where it increases according to the timing of their firing. We find that H3K37me1 hinders Mcm2 interaction with chromatin, maintaining low levels of MCM outside of conventional replication origins. Lack of H3K37me1 results in defective DNA replication from canonical origins while promoting replication events at inefficient and non-canonical sites. Collectively, our results indicate that H3K37me1 ensures correct execution of the DNA replication program by protecting the genome from inappropriate origin licensing and spurious DNA replication.

## INTRODUCTION

The eukaryotic genome is compacted into chromatin, whose fundamental structural unit is the nucleosome. Nucleosomes consist of 150 bp of DNA wrapped around a core histone octamer composed of two copies each of histones H2A, H2B, H3, and H4 (Luger et al., 1997). Dynamic remodeling of nucleosomes affects chromatin organization, controlling the association of protein complexes to the genome. In addition, histones are decorated by a plethora of covalent post-translational modifications that can directly affect chromatin structure or regulate the association of effector proteins to chromatin. Thus, the organization of the genome into chromatin provides an instructive regulatory platform for all DNA processes, including replication.

In *S. cerevisiae*, DNA replication initiates at particular sites in the genome known as autonomous replication sequences (ARSs), which are bound by the origin recognition complex (ORC) throughout most of the cell cycle. Initiation of DNA replication can be divided into two distinct steps. First, during the G1 phase, ORC coordinates the recruitment of Cdc6, followed by the Cdt1-MCM replicative DNA helicase (Mcm2–7) onto ARS to form an inactive pre-replicative complex (pre-RC). Second,

as cells enter S phase, DNA replication is initiated by the action of S-phase kinases. Firing of each replication origin requires the loading of two MCM complexes as a head-to-head double hexamer around the DNA. Phosphorylation of these MCM dimers by the S-phase kinases leads to structural changes and recruitment of the activating factors Sld3/Cdc45, GINS, and Mcm10, among others, to assemble a complete replisome and initiate DNA synthesis (for a review, see Bell and Labib, 2016).

In *S. cerevisiae*, origins of replication share a 17 bp AT-rich motif, called ARS consensus sequence (ACS), that is necessary but not sufficient for origin activity. The ACS is followed by a T-rich element, B1, and two A-rich elements, B2 and B3, which are important for proper origin function (for a review, see Bell and Labib, 2016). Although the ACS occurs more than 10,000 times in the yeast genome (Breier et al., 2004), the current estimation is that only 400–600 ARSs are commonly used (Siow et al., 2012). On average, in a population, ARSs fire according to a very coordinated and reproducible temporal program, although in individual cells it is a probabilistic event. During S phase, some origins fire early, some later, while others are passively replicated from neighboring ones. As a general trend, early/efficient ARSs are located near transcriptionally active

open reading frames (ORFs), whereas late/inefficient ARSs are close to the telomeric chromosomal ends. Regardless of their firing time, most ACSs are located in intergenic regions, minimizing collisions between RNA polymerases and components of the replication machinery (Eaton et al., 2010; Mori and Shirahige, 2007).

Numerous mathematical models and experimental work have sought to explain the stochastic nature of ARS firing. The likelihood that an origin will fire is directly proportional to the number of MCM dimers recruited to it, which in turn is influenced by the interaction between the MCM helicase and ACS adjacent nucleosomes (Das et al., 2015; Das and Rhind, 2016). Accordingly, repositioning nucleosomes adjacent to an ARS results in firing defects due to impaired assembly of the pre-replication complex (Lipford and Bell, 2001). Reciprocally, binding of ORC to the ACS defines nucleosome positioning around ARS (Eaton et al., 2010). In addition to recruitment, helicase activation also depends on the interaction between DNA-loaded MCM and adjacent chromatin (Belsky et al., 2015).

Over the past few years, the importance of proximal histone modifications in defining origin use and regulating replication initiation has become increasingly apparent. For instance, H3 and H4 lysine acetylation at origin-proximal nucleosomes promotes origin firing by recruiting Cdc45 (Vogelauer et al., 2002; Unnikrishnan et al., 2010). Moreover, mono-methylation but not di- or tri-methylation of H3K36 at ARS flanking nucleosomes also contributes to MCM helicase activation by Cdc45 (Pryde et al., 2009). Here, we show that H3K37me1 is important for the regulation of origin licensing. Specifically, H3K37me1 hinders Mcm2 interaction with chromatin, maintaining low levels of MCM outside of replication origins and preventing DNA replication events at non-canonical sites.

## RESULTS

### Set2p and Set1p (COMPASS) are required for H3K37me1 *in vivo*

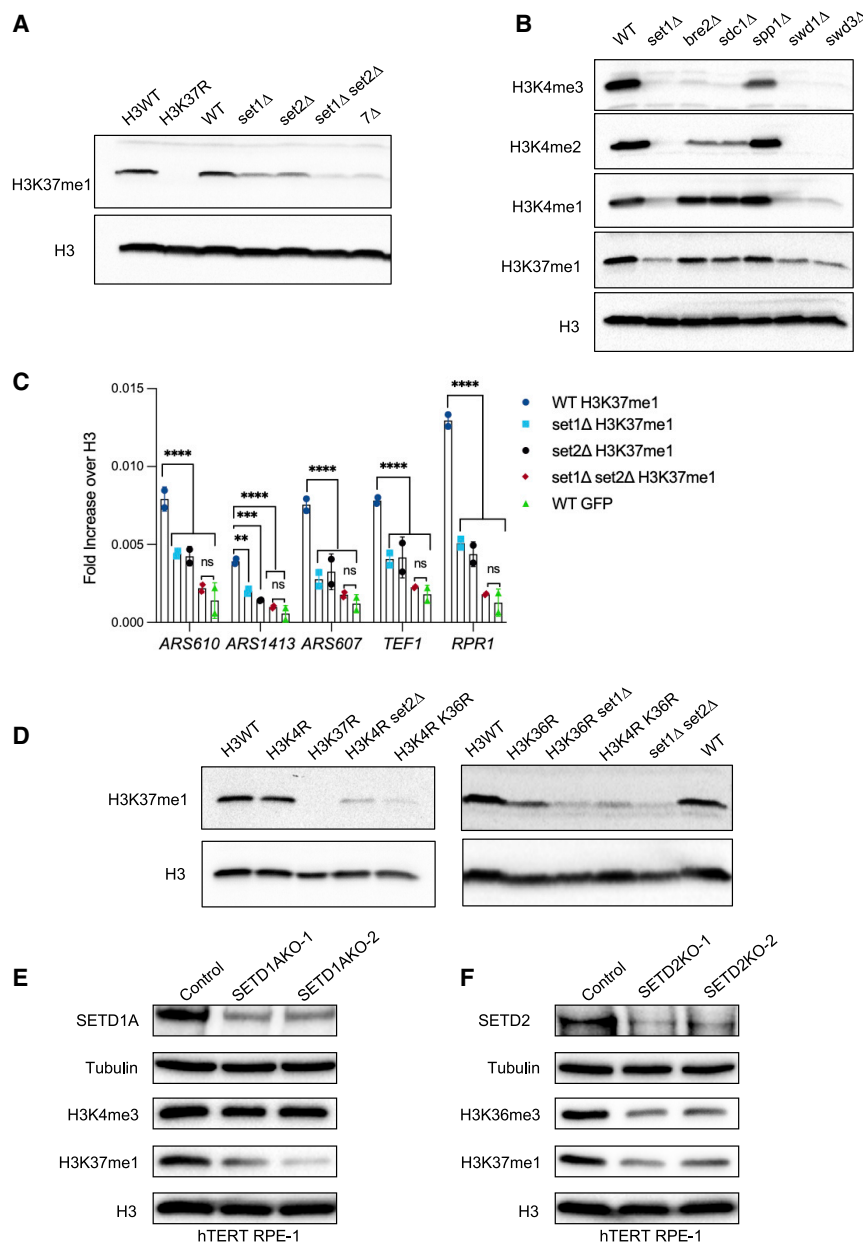
Various mass spectrometry studies have identified H3K37me1 in both yeast and human cells (Li et al., 2017; Ren et al., 2007; Unnikrishnan et al., 2010). However, its function remains unknown. To investigate possible roles of H3K37me1, we raised a specific antibody against H3K37me1. Dot-blot analyses confirmed that the antibody recognizes H3K37me1 peptides, but not unmodified ones, and that it does not cross-react with H3K36me1 (Figure S1A). Importantly, it does recognize H3K37me1 in combination with H3K36me1, the actual context in which H3K37me1 was reported to occur in yeast (Unnikrishnan et al., 2010) (Figure S1A). In addition, ELISA revealed antibody specificity toward H3K37me1-modified peptides several orders of magnitude higher than toward unmodified ones (Figure S1B). Immunoblot studies further showed that the antibody reacts with H3 purified from wild-type (WT) yeast, but not from an isogenic H3K37A mutant strain (Figure S1C). Similarly, chromatin immunoprecipitation (ChIP) experiments revealed that the antibody detects H3 from WT (H3WT) but not from H3K37A mutant chromatin (Figure S1D). Moreover, the antibody strongly detects H3 in yeast total protein extracts, but it shows only very weak reactivity toward unmodified recombinant H3 produced in bacteria (Figures S1E

and S1F). Finally, the antibody also detects H3K37me1 in a wide variety of mammalian cell lines (Figure S1G). Taken together, these results confirm that the antibody specifically recognizes H3K37me1 and that this modification is conserved from yeast to human.

To identify the enzyme(s) responsible for H3K37me1 deposition, we immunoblotted extracts from individual yeast knockout strains of all non-essential proteins harboring an S-adenosylmethionine (SAM) binding domain (Figure S2A). As SAM is the only known methyl-group donor in yeasts, the list includes all *bona fide* and putative methyltransferases (Petrossian and Clarke, 2009). No single deletion completely abolished H3K37me1, but deletion of either *SET1* or *SET2* lysine methyltransferases resulted in a decrease of H3K37me1 levels (Figure S2A). Confirming these findings, *de novo* deletion of *SET1*, *SET2*, both together, and the entire SET family of lysine methyltransferases ( $\Delta$ ) showed that individual deletion of *SET1* or *SET2* decreased H3K37me1, but the signal was further reduced when both were deleted together (Figure 1A). Indeed, *SET1* and *SET2* double deletion reduced the level of H3K37me1 as effectively as deleting the entire SET family ( $\Delta$ ) (Figure 1A). Of note, the residual signal in the *set1 $\Delta$  set2 $\Delta$*  double mutant was comparable with that of unmodified recombinant H3 (Figure S2B). These results strongly suggested that Set1p and Set2p are redundant enzymes in the catalysis of H3K37me1. Furthermore, we found that deletion of *SWD1* and *SWD3*, two members of the Set1 protein complex (COMPASS) required for its activity, phenocopied *SET1* deletion (Figure 1B), further supporting the notion that Set1p is important for H3K37me1 *in vivo*. Consistent results were observed in ChIP experiments. Both, *set1 $\Delta$*  and *set2 $\Delta$*  single mutants showed lower levels of H3K37me1 than the isogenic WT, at all tested locations (Figure 1C). Importantly, the modification was further decreased to background levels in the *set1 $\Delta$  set2 $\Delta$*  double mutant (Figure 1C, compare H3K37me1 with GFP). Thus, simultaneous deletion of *SET1* and *SET2* eliminates H3K37me1 *in vivo*.

Notwithstanding the above, it was possible that the lack of H3K37me1 in the *set1 $\Delta$  set2 $\Delta$*  strain was actually due to a transcriptional defect of an essential methyltransferase. *S. cerevisiae* encodes nine SAM binding proteins that are essential for cell viability (Petrossian and Clarke, 2009). However, we found no significant differences in the relative amounts of their mRNAs in WT and the *set1 $\Delta$  set2 $\Delta$*  double-knockout strains, except for mild increases of *NOP2* and *DIM1* (Figure S2C). Altogether, the above results indicate that both Set2p and Set1p (COMPASS) are responsible for H3K37me1 *in vivo*.

It is well established that Set1p (COMPASS) catalyzes mono-, di-, and tri-methylation of H3K4 (Schneider et al., 2005) and that Set2p catalyzes mono-, di-, and tri-methylation of H3K36 (Strahl et al., 2002). However, none of these modifications are essential for H3K37me1, as mutation of either K4 or K36 to arginine (K4R and K36R, respectively) does not eliminate H3K37me1 (Figure 1D). We suspected that the residual H3K37me1 observed in each of the “R mutants” was actually catalyzed by the enzyme whose known target site was not mutated. To test this hypothesis, we deleted *SET2* in the H3K4R strain and *SET1* in the H3K36R strain. H3K37me1 was further reduced in both cases (Figure 1D). We conclude that although K4 and K36 methylation



**Figure 1. Set2 and Set1 (COMPASS) are required for H3K37me1 *in vivo***

(A and B) Immunoblot analysis of total protein extracts from isogenic yeast strains as specified. Proteins were separated using SDS-PAGE in 16% acrylamide gels.

(C) H3K37me1 levels at different genomic locations, shown by ChIP-qPCR. Chromatin from isogenic strains was immunoprecipitated using anti-H3K37me1, anti-H3, and anti-GFP (negative control) antibodies. Statistical analysis was performed using two-way ANOVA multiple comparisons and Tukey's multiple-comparison test ( $\alpha = 0.05$ ; \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ , and \*\*\*\* $p \leq 0.0001$ . Error bars represent mean  $\pm$  SD of two independent experiments.

(D) Immunoblot analysis of total protein extracts from isogenic yeast strains as specified. Proteins were separated using SDS-PAGE in 16% acrylamide gels.

(E and F) Immunoblot analysis of total protein extracts from (E) SETD1A or (F) SETD2 CRISPR knockout cell pools. Human RPE1 cells stably expressing Cas9 were transfected with two guide RNAs targeting different exons of each gene or a non-targeting control. Protein extracts, prepared 7 days after transfection, were separated using SDS-PAGE in 12% acrylamide or 3%–8% Tris-Acetate gels.

reduced H3K36me3 levels and, importantly, H3K37me1 (Figure 1F). Together, these data support the notion that the mammalian SETD1A and SETD2 enzymes are orthologs of Set1p and Set2p regarding H3K37 methylation.

### Set2p and Set1p (COMPASS) methylate H3K37 *in vitro*

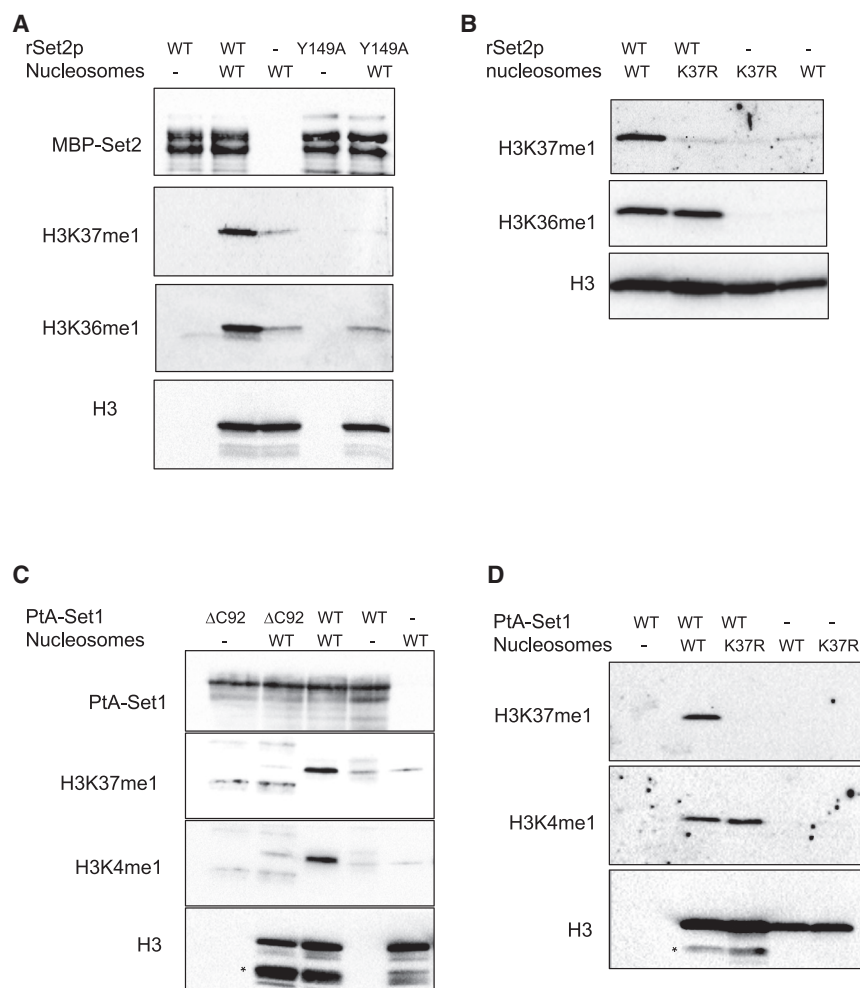
Thereafter, we addressed whether Set1p (COMPASS) and Set2p display H3K37 methyltransferase activity *in vitro*. Recombinant Set2p is active as a methyltransferase (Strahl et al., 2002), so we produced in *E. coli* full-length WT Set2p and two different mutants, targeting conserved

are dispensable, Set1p needs a K at position 4, and Set2p a K at position 36, to facilitate H3K37 methylation.

Given that H3K37me1 is conserved throughout evolution, we explored the identity of the enzymes responsible for this modification in higher eukaryotes. SETD1A and SETD1B, which are required for H3K4 methylation, are the closest homologs of Set1p (Shilatifard, 2012), while SETD2, which is responsible for H3K36me3 *in vivo*, is the closest homolog of Set2p (McDaniel and Strahl, 2017). We used CRISPR-Cas9 technology to knock out SETD1A and SETD2 in hTERT RPE-1 cells targeting two different exons per gene. Consistent with previous reports, knocking out SETD1A did not affect H3K4me3 (Tajima et al., 2015, 2019). However, it clearly decreased the levels of H3K37me1 (Figure 1E). As expected, knocking out SETD2

residues within the catalytic domain (Set2 Y149A and Set2 N198Q). WT Set2p, but neither of the mutants, methylated H3 in a radioactive methyltransferase reaction using  $^3\text{H}$ -SAM as a methyl donor (Figure S3A). However, such an approach does not reveal the site of methylation within H3. Hence, we performed reactions on nucleosomes reconstituted with either recombinant WT (H3WT) or H3K37R mutant proteins and analyzed them by immunoblot using H3K36me1- and H3K37me1-specific antibodies. First, WT Set2p, but not Set2Y149A, methylated H3K37 and H3K36 in nucleosomes harboring H3WT (Figure 2A), indicating that methylation of both of these K residues depends on the Set2p catalytic domain. Then, we assessed Set2p activity on WT versus H3K37R mutant nucleosomes. Set2p methylated H3K36 in both nucleosome preparations (Figure 2B). However,





**Figure 2. Set2 and Set1 (COMPASS) methylate H3K37 *in vitro***

(A) *In vitro* methyltransferase assay. Recombinant wild-type Set2p and Set2Y149Ap mutant proteins were incubated with wild-type (WT) H3.1 nucleosomes in the presence of SAM. Reactions were separated using SDS-PAGE in 16% acrylamide gels.

(B) *In vitro* methyltransferase assay. Equal amount of recombinant wild-type Set2p was incubated with H3.1WT or H3.1K37R nucleosomes in the presence of SAM. Reactions were separated using SDS-PAGE in 16% acrylamide gels.

(C) *In vitro* methyltransferase assay. Wild-type PtA-Set1p and PtA-set1ΔC92p yeast purified complexes were incubated with wild-type H3.1 nucleosomes in the presence of SAM. Reactions were separated using SDS-PAGE in 16% acrylamide gels. Asterisk indicates Δ tail-H3.

(D) *In vitro* methyltransferase assay. An equal amount of wild-type PtA-Set1p complex was incubated with H3.1WT or H3.1K37R nucleosomes in the presence of SAM. Reactions were separated using SDS-PAGE in 16% acrylamide gels. Asterisk indicates Δ tail-H3.

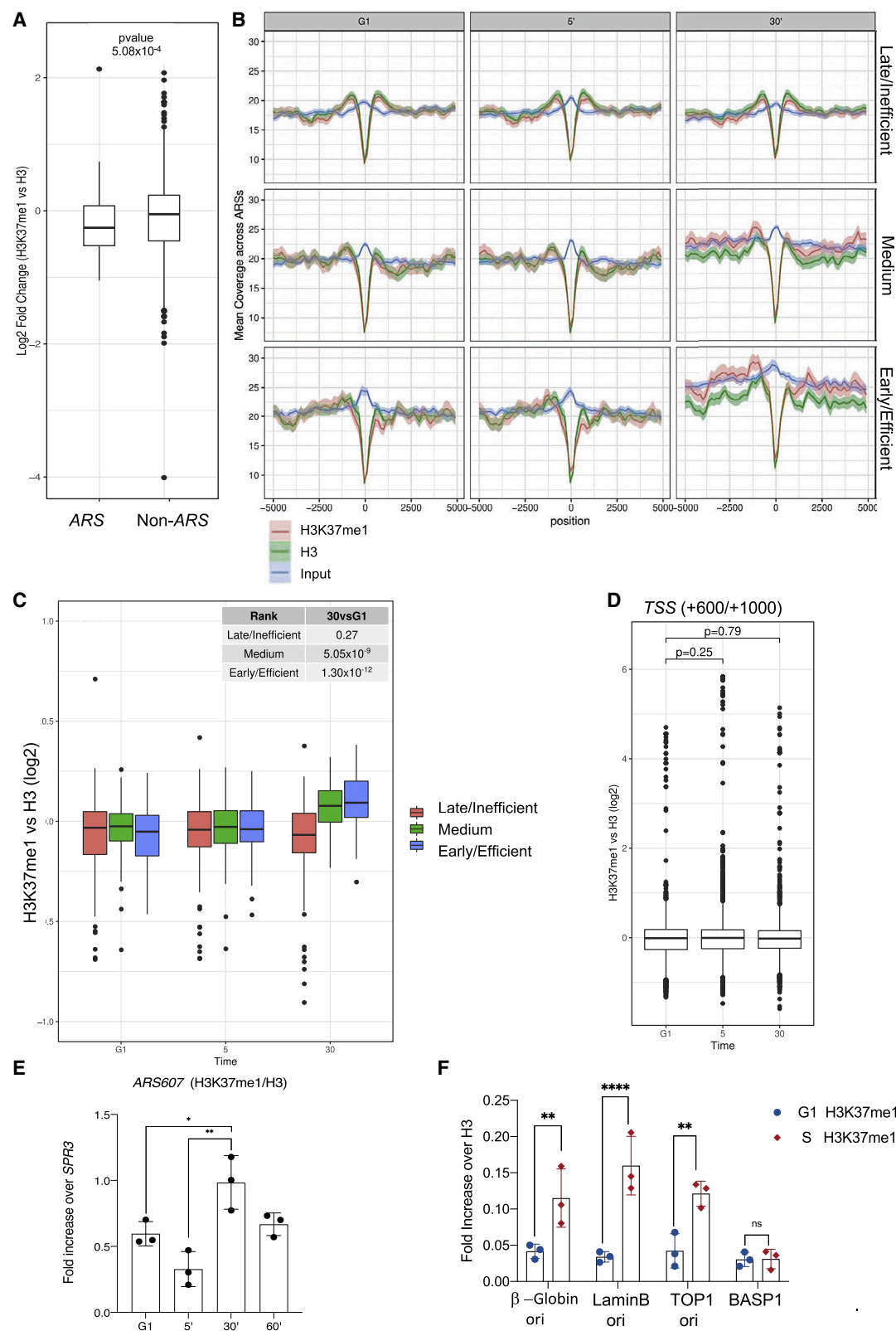
### H3K37me1 correlates with replication origin firing

To investigate the function of H3K37me1, we analyzed its genome-wide distribution using ChIP followed by high-throughput DNA sequencing (ChIP-seq). H3K37me1 was broadly distributed throughout the genome of asynchronous yeast cells,

it could only methylate H3K37 in H3WT-containing nucleosomes and not in H3K37R mutant nucleosomes (Figure 2B). Thus, Set2p mono-methylates H3K37 *in vitro*.

As recombinantly produced Set1p is not active *in vitro* (Takahashi et al., 2011), we purified the Set1 complex (COMPASS) from yeasts expressing either a protein A-tagged Set1p WT allele (PtA-Set1) or a mutant lacking the C-terminal 92 amino acids (PtA-Set1ΔC92), which corresponds to the post-SET domain (Figure S3B). As previously shown (Santos-Rosa et al., 2002), *in vitro* methyltransferase reactions using H3 and <sup>3</sup>H-SAM as donor revealed that PtA-Set1ΔC92 is not active, in contrast to WT Set1p (Figure S3C). We then performed *in vitro* reactions using PtA-Set1 and PtA-Set1ΔC92 on WT nucleosomes (Figure 2C, top panel) and analyzed them by immunoblotting with H3K37me1- and H3K4me1-specific antibodies. We detected H3K37me1 and H3K4me1 only in the reactions catalyzed by the WT enzyme (Figure 2C, middle panels). Furthermore, H3K37me1 methylation by WT PtA-Set1p was lost when H3K37R-containing nucleosomes were assayed (Figure 2D, top panel). These results show that Set1p (COMPASS) mono-methylates H3K37 *in vitro*. Altogether, the above results indicate that both Set2p and Set1p (COMPASS) are indeed H3K37 mono-methyltransferases.

with no apparent enrichment at specific locations. However, previous mass spectrometry studies identified H3K37me1 at nucleosomes flanking the plasmid-borne replication origin *ARS1* (Unnikrishnan et al., 2010), so we focused our attention on origins of replication. Surprisingly, we found that ARSs were significantly H3K37 hypo-methylated compared with the rest of the genome (Figure 3A), suggesting that H3K37me1 is specifically restrained from these locations. To explore whether H3K37me1 deposition at ARSs is cell cycle regulated, we arrested WT cells in G1 using  $\alpha$  factor and then released them into the cycle. We prepared ChIP-seq libraries from G1-, G1/S-, and S-phase synchronized cells, which correspond to 0, 5, and 30 min after release, respectively (Figure S4A). Origins were classified as early/efficient, medium, and late/inefficient on the basis of their firing efficiency according to our DNA replication profiling dataset (see Figure S6C). As expected, early origins showed a clear increase in DNA copy number during S phase, whereas medium origins showed a moderate increase, and late origins showed no change at all (Figure 3B, 30' input). Notably, we observed a pronounced increase of H3K37me1, relative to H3, during S phase specifically at early origins, while medium and late origins showed a milder increase or no change, respectively (Figures 3B and 3C, 30'). We chose "non-ARS



(legend on next page)

control region” as the window +600 to 1,000 bp from transcription start sites, on the basis of the ORF average size in *S. cerevisiae* (1.2 kb) and the premise that most *ARS* are located outside ORFs. In contrast to the origins of replication, non-*ARS* showed a constant H3K37me1/H3 ratio throughout the experiment (Figure 3D; Figure S4B). We independently confirmed these results at different locations using qPCR (Figure 3E; Figure S4C). From these experiments, we conclude that H3K37me1 is under-represented at origins of replication in G1, but its presence at these regions is cell cycle regulated and correlates with the timing of their firing.

Recently it was reported that H3K37 methylation is involved in gametogenesis in *S. pombe* (Shen et al., 2019). Thus, we analyzed H3K37me1 in a sporulation time course using the SK1 diploid strain (Figure S5A). We found that, in fact, H3K37me1 is also upregulated at the onset of premeiotic DNA replication in *S. cerevisiae* (Figure S5B), suggesting that H3K37me1 correlates with DNA replication not only during mitosis, but also during meiosis.

To address whether H3K37me1 is also cell cycle regulated at human origins of replication, we synchronized non-transformed but immortalized retinal pigment epithelial cells (hTERT RPE-1) at the G1/S boundary and released them into the cell cycle (Figure S5C). We then performed ChIP on G1/S- and S-phase synchronized cells and analyzed H3K37me1 at different well-characterized replication origins. We observed that H3K37me1 levels significantly increased with respect to H3 at origins of replication, but not at an unrelated region, during S phase (Figure 3F). This result suggests that the distribution and cell cycle regulation of this modification are conserved in mammalian cells.

### H3K37me1 regulates DNA replication origin firing

Set1p and Set2p are involved in many cellular processes due to their K4 and K36 methyltransferase activities, respectively (for a review, see Freitag, 2017). This prevented us from studying the function of H3K37 methylation by mutating the responsible enzymes. Instead, we probed the function of this modification by mutating the modified residue. Mutation of H3K37 to arginine retains the residue's positive charge but prevents its methylation by Set1p and Set2p. Thus, we compared the DNA replication profiles of H3WT and isogenic H3K37R mutant by performing genome-wide immunoprecipitation (IP) of nascent BrdU-labeled DNA (Lengronne et al., 2001). Once transformed with the thymi-

dine kinase (TK) and nucleoside transporter (hENT1), both H3WT and H3K37R strains incorporated BrdU with similar efficiencies (Figure S6A). Also, we immunoprecipitated DNA only from cells incubated with BrdU (Figure S6B), and the signal was higher at *ARS*s compared with other regions of the genome (Figure S6B), testifying to the specificity of our IPs.

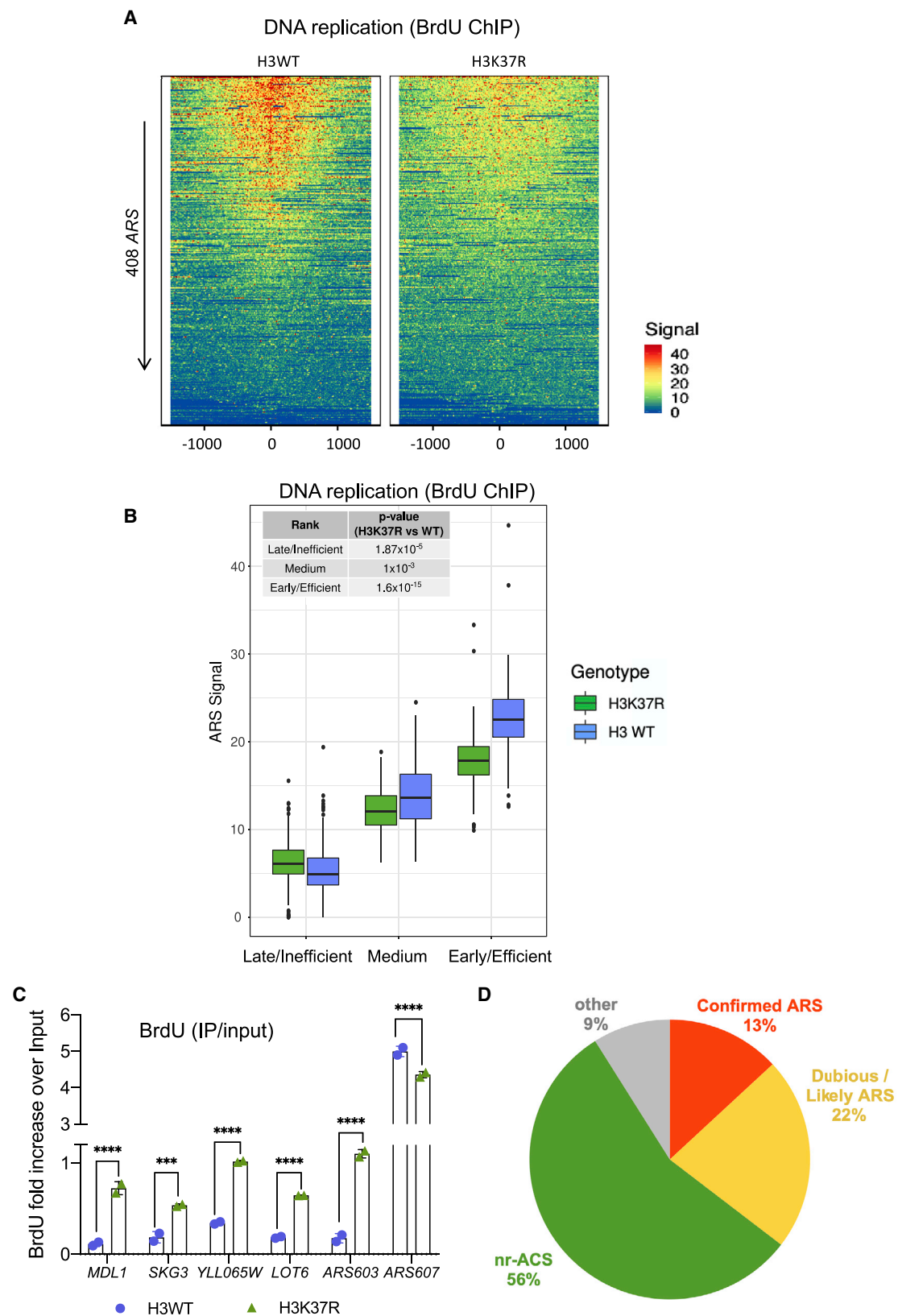
We then compared H3WT and H3K37R BrdU incorporation at replication origins. *ARS*s were ranked top to bottom according to their firing efficiencies in the H3WT strain. Remarkably, we observed an evident general decrease in the replication efficiency of the H3K37R mutant (Figure 4A; Figure S7A). Next, we sorted all replication origins into early/efficient, medium, and late/inefficient *ARS*s as before (Figure S6C). We found that early/efficient and medium replication origins were less active in the H3K37R mutant strain than in the WT strain (Figure 4B). Interestingly, the late/inefficient group, which includes the “dormant replication origins,” showed an increased BrdU signal in the H3K37R mutant strain with respect to the H3WT strain (Figure 4B), indicating that H3K37me1 is required for the correct execution of the DNA replication program.

Surprisingly, comparison of the H3WT and H3K37R DNA replication profiles outside of canonical *ARS* revealed numerous genomic locations displaying stronger or exclusive BrdU labeling in the mutant strain (Figure S7B). These replication events occur several kilobases away from the nearest firing *ARS*, supporting the notion that they result from *de novo* firing rather than from interference with nearby replicating *ARS* and/or differences in the speed of the replication fork (Figure S7C). Importantly, BrdU-IP followed by qPCR using specific primers confirmed that DNA replication initiates substantially better at these distinct locations in the H3K37R mutant (Figure 4C; Figure S7D), suggesting that H3K37me1 may serve to restrict the initiation of DNA replication to bona fide origins of replication.

To further investigate the nature of the H3K37R unique events, we first calculated, as a proxy, the average width of the replication peaks of the middle *ARS* group (2.0957 kb). We then assumed that all H3K37R unique peaks occurring within a 2 kb size window initiated from a single event. Using these stringent criteria, we identified 326 replication events, 91% of which occurred at previously predicted “ACS matches” (Breier et al., 2004). Specifically, H3K37R unique peaks included (1) confirmed late/inefficient *ARS*s, which were not firing in the WT strain under our experimental conditions (13%); (2) dubious/likely *ARS*

### Figure 3. Histone H3K37me1 correlates with replication origin firing

(A) Boxplot representing H3K37me1 enrichment over H3 ( $\log_2$ ) at replication origins (*ARS*  $\pm$  2 kb) versus non-*ARS* (window +600 to +1,000 bp from transcription start sites).  
(B) Coverage plot showing the mean normalized ChIP-signal ( $\pm$ SEM) for H3, H3K37me1, and input at *ARS*  $\pm$  5 kb (total number of *ARS*s = 408). *ARS*s were grouped into three categories according to firing efficiency/time.  
(C and D) Boxplot showing the mean H3K37me1 enrichment over H3 ( $\log_2$ ) at *ARS*  $\pm$  2 kb (left panel) and at non-*ARS* control regions (TSS +600 to +1,000 bp, right panel). p values were calculated using the Mann-Whitney-Wilcoxon test.  
(E) ChIP-qPCR H3K37me1 levels at *ARS*607. Chromatin from the time points shown in Figure S4A was immunoprecipitated using anti-H3K37me1 and anti-H3 antibodies. H3K37me1/H3 levels were normalized to a non-*ARS* region (*SPR3*). Statistical analysis was performed using two-way ANOVA multiple comparisons and Tukey's multiple-comparison test ( $\alpha$  = 0.05); \*p  $\leq$  0.05, \*\*p  $\leq$  0.01, and \*\*\*p  $\leq$  0.001. Error bars represent the mean  $\pm$  SD of three independent experiments.  
(F) ChIP-qPCR H3K37me1 enrichment at different origins of replication in G1/S- and S-phase synchronized human RPE1 cells as shown in Figure S4D. Statistical analysis was performed using two-way ANOVA corrected for the comparisons using the Holm-Sidak method ( $\alpha$  = 0.05); \*p  $\leq$  0.05, \*\*p  $\leq$  0.01, and \*\*\*p  $\leq$  0.001. Error bars represent the mean  $\pm$  SD of three independent experiments.



(legend on next page)



(22%); and (3) non-replicative ACS (nr-ACS), which, under our experimental conditions, were able to drive significant replication in the H3K37R mutant (56%) (Figure 4D). All together, these results are consistent with a model in which H3K37me1 restricts origin firing potential to conventionally used origins of replication, preventing non-canonical ones from firing.

### H3K37me1 modulates MCM association to chromatin

It has been reported that Mcm2, a subunit of the MCM replicative helicase, physically interacts with nucleosomes through association of its N terminus to histone H3 (Ishimi et al., 1998). Thus, we considered the possibility that methylation of H3K37 might affect this interaction. As a preliminary analysis, we performed *in vitro* pull-down experiments using H3 biotinylated peptides (unmodified and methylated at K37, K36, or both) and recombinant Mcm2<sub>aa1-aa480</sub>-His<sub>6</sub>, which has been shown to mediate the Mcm2 to H3 interaction (Ishimi et al., 1998). We found that Mcm2 has a slight preference to bind to H3 when K36 is mono-methylated (Figure 5A; Figures S8A and S8B). This is in agreement with *in vivo* data showing that K36me1, but not K36me2 or K36me3, is present at ARSs upon firing (Pryde et al., 2009). Notably, the presence of one methyl group at H3K37 compromises the interaction between the H3 and Mcm2 (Figure 5A; Figures S8A and S8B), suggesting that H3K37me1 creates an unfavorable environment for the interaction between MCM and chromatin.

Next, we investigated the *in vivo* recruitment of the MCM complex to chromatin in the absence of H3K37me1. To do so, we HA-tagged Mcm2 (as a proxy for the binding of the MCM complex) in H3WT and H3K37R cells. Both strains expressed similar levels of Mcm2-HA (Figure S8C). In addition, we could immunoprecipitate Mcm2-HA at ARSs in the HA-tagged, but not in its isogenic untagged strain, attesting to the specificity of our ChIPs (Figure S8D). Then, we performed genome-wide Mcm2-HA ChIP-seq in G1-arrested H3WT and H3K37R cells. We “spiked” our yeast cultures with an identical number of *S. pombe* cells expressing Mcm3-HA-tagged, for use as an “external normalizer” (Orlando et al., 2014). Importantly, we found that lack of H3K37me1 resulted in a significant increase in the association of MCM to all ARS groups (Figure 5B; Figure S8E) and, notably, all over the genome (Figure 5C; Figure S8F). ChIP-qPCR confirmed the increase in MCM recruitment, which was particularly pronounced at those specific genomic locations that triggered DNA replication in the H3K37R mutant but not in the WT strain (Figure 5D). Importantly, however, we found that histone occupancy was not affected by mutating H3K37 (Figure S8G). These results, in agreement with the *in vitro* data, support the

notion that H3K37me1 negatively regulates the association of MCM to chromatin.

Concurring with the above, H3K37me1 and MCM anticorrelate at the efficient origin ARS607, but do not at the inactive ARS1333, in time-course experiments (Figure S9A). In fact, we found a global anti-correlation between H3K37me1 and MCM at active ARSs (Figure S9B, early/efficient origins). We then investigated whether H3K37me1 may also affect MCM dynamics upon G1 release. We observed that H3K37R mutation led to retention and/or better detection of MCM at origins of replication, particularly at late/inefficient ones (Figure S9C). This result is intriguing, as it is well established that MCM recruitment to ARS is restricted to late M/G1 phase (Diffley, 2011). We did not find significant differences between H3WT and H3K37R chromatin states upon G1 release (as measured by H3 occupancy; Figure S9D). Yet we cannot exclude that MCM structural changes occurring during G1 → S transition in the H3K37R mutant may increase epitope accessibility and detection of the already G1-recruited helicase. Alternatively, lack of H3K37me1 may result in longer retention of inactive MCM complexes, particularly at late/inefficient ARS, increasing the probability of those to encounter a helicase traveling with the replication forks from adjacent origins. Altogether, these results are consistent with a model in which H3K37me1 modulates the association and dynamics of the replicative helicase to chromatin.

### Depletion of H3K37me1 results in competition between conventional and non-replicative ARS for rate-limiting MCM activators

Our results show that lack of H3K37me1 increases MCM chromatin binding across the genome but reduces DNA replication from most commonly used ARSs. It has been reported that recruitment of limiting helicase activators (Sld2, Sld3, Sld7, Dpb11, Dbf4, and Cdc45) define origin efficiency and firing time (Mantiero et al., 2011; Tanaka et al., 2011). Thus, we hypothesized that competition between canonical ARSs and novel replication sites for these rate-limiting factors might underlie H3K37R phenotypes. To test this, we first performed ChIP analysis of Cdc45 in cells arrested in G1 and released into S phase in the presence of hydroxyurea. We found lower levels of Cdc45 at efficient ARSs, such as ARS607 or ARS305 (Figure 6A, left panel), in the H3K37R mutant strain compared with the WT strain but increased Cdc45 recruitment to the late/inefficient ARS112 and also to genomic locations that supported DNA replication exclusively in the H3K37R mutant (Figure 6A, right panel). These results suggest that in the absence of H3K37me1, limiting

### Figure 4. H3K37me1 regulates DNA replication origin firing

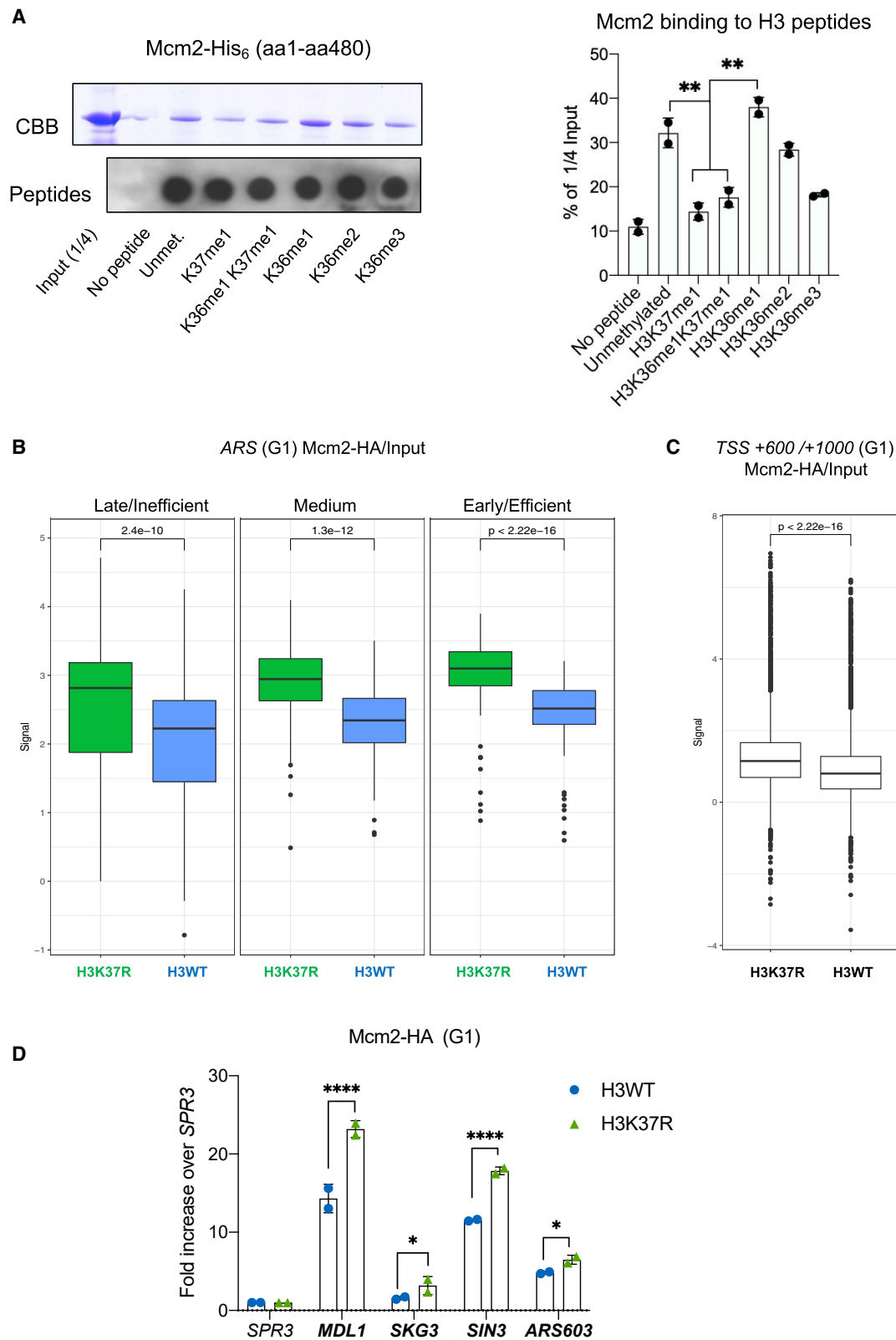
(A) Heatmap showing the distribution of BrdU incorporation at ARS in wild-type H3 and H3K37R mutant cells. Origins were aligned from highest to lowest BrdU signal in the wild-type strain and centered at ARS ACS (ARS consensus sequence). For visualization purposes, the heatmaps' color scale was saturated at the 99th percentile of the distribution of signal intensities. The plot represents the average of four independent experiments.

(B) Boxplot showing the distribution of mean BrdU signal at the replication origins (ARS ± 1 kb) shown in (A). The p values were calculated using the Mann-Whitney-Wilcoxon test. ARSs were classified into early/efficient, medium, and late/inefficient following the BrdU distribution shown in Figure S6C.

(C) ChIP-qPCR experiments showing incorporation of BrdU in H3WT and H3K37R mutant cells. Statistical analysis was performed using two-way ANOVA corrected for the comparisons using the Holm-Sidak method (alpha = 0.05); \*p ≤ 0.05, \*\*p ≤ 0.01, \*\*\*p ≤ 0.001, and \*\*\*\*p ≤ 0.0001. Error bars represent the mean ± SD of two independent experiments.

(D) Pie chart showing proportion of H3K37R unique replication events occurring at different genomic locations.





(legend on next page)

helicase activators are diverted from canonical *ARS*s to unconventional and non-replicative *ACS*s.

Next, we engineered H3WT and H3K37R strains to express additional copies of Sld2, Sld3, Sld7, Dpb11, Dbf4, and Cdc45 under the control of the strong inducible *GAL1-10* promoter (Mantiero et al., 2011; Tanaka et al., 2011; STAR Methods). Importantly, immunoblot analysis revealed that both strains induced the expression of all factors to the same extent (Figure S10A). We performed BrdU IP under non-inducing conditions (glucose, OFF) and over-expressing conditions (galactose, ON) and analyzed the immunoprecipitated DNA with primers specific to efficient (*ARS607*) and inefficient (*ARS603*) origins. In agreement with previous reports (Mantiero et al., 2011; Tanaka et al., 2011), overexpression of MCM activators (galactose) resulted in stronger replication from inefficient and, to a certain extent, also from efficient *ARS*s in the WT strain (Figure S10B, compare H3WT Glu with H3WT Gal).

Thus, we analyzed DNA replication genome-wide. As in previous experiments, *ARS*s were ranked top to bottom according to their firing efficiencies in the H3WT strain. We found that overexpression of the limiting helicase activators suppressed the overall firing deficiency otherwise occurring in the H3K37R strain (Figure 6B), suggesting that competition for limiting MCM activators underlies the mutant's DNA replication defects. We sorted *ARS*s into early/efficient, medium, and late/inefficient, as before. Notably, overexpression of MCM activators largely rescued DNA replication from early/efficient origins (Figure 6B; Figure S10C). However, under the same conditions, late/inefficient *ARS*s still fired more in the H3K37R mutant than in the WT strain (Figure 6B; Figure S10C). Similar results were obtained using ChIP-qPCR analysis of *ARS607* (early/efficient) and *ARS603* (late/inefficient), respectively (Figure S10B). These results argue that the increased activity of inefficient *ARS* in the H3K37R mutant is not merely a response to palliate the deficient firing from conventional ones. In fact, "ACS sites" that remained largely inactive in WT cells, even under excess helicase activators, fired in the absence of H3K37me1 (Figure 6C).

Altogether, the above experiments are consistent with a model in which the absence of H3K37me1 causes a redistribution of origin firing toward unconventional and non-replicative *ACS*s, diverting limiting helicase activators from canonical *ARS*s.

## DISCUSSION

In late mitosis/early G1 of each cell cycle, *ARS*s are bound by ORC and the replicative helicase (MCM) in a process known as

origin licensing, necessary but not sufficient for an origin to fire. Commitment to fire occurs upon MCM activation by recruitment of rate-limiting factors such as Sld2, Sld3, Sld7, Dpb11, Dbf4, and Cdc45 at the G1/S transition (for a review, see Bell and Labib, 2016; Riera et al., 2017). ORC binding to the *ACS* and recruitment of the replicative helicase depend on the organization of adjacent nucleosomes, whose epigenetic state can influence origin specificity, licensing, and helicase activation (Azmi et al., 2017; Belsky et al., 2015; Soriano et al., 2014; Soudet and Stutz, 2019). In particular, histone acetylation and H3K36 mono-methylation promote *ARS* firing and recruitment of the helicase activating protein Cdc45 (Knott et al., 2009; Pryde et al., 2009; Unnikrishnan et al., 2010; Vogelauer et al., 2002). Also, H4K20 mono-methylation plays a role, at the exit of mitosis, in establishing the necessary chromatin compaction to limit DNA replication licensing (Shoaib et al., 2018).

Mcm2, a component of the replicative helicase, directly binds to a region of H3 spanning amino acids 26–67 (Ishimi et al., 1998). Our results suggest that regulation of this interaction by H3K37me1 is important for origin licensing. We showed that H3K37me1 is uniformly distributed throughout the vast majority of the genome during G1 but is particularly sparse at origins of replication, where its absence might improve ORC binding and/or stabilize MCM-ORC interactions on chromatin, thus facilitating MCM loading. In agreement with this possibility, depletion of H3K37me1 results in increased MCM chromatin binding across the entire genome, particularly at *ACS* locations where cooperativity between ORC and nucleosomes may occur. Perhaps surprisingly, we observed that this increase in MCM chromatin association is accompanied by a general decrease in firing from canonical *ARS*s. Although the *S. cerevisiae* genome harbors approximately 10,000 *ACS* motifs (Breier et al., 2004), only about 400 are effectively used during standard growth conditions (Eaton et al., 2010; Xu et al., 2006). Notably, we found that mutation of H3K37 causes a redistribution of origin firing potential from *ARS* efficiently used toward inefficient and non-canonical *ACS*, suggesting that H3K37me1-dependent regulation of MCM chromatin loading (and perhaps activation) is important for the proper execution of the origin firing program. In line with this, recruitment of limiting helicase activators (Sld2 Sld3, Sld7, Dpb11, Dbf4, and Cdc45) define origin efficiency and firing time (Mantiero et al., 2011; Tanaka et al., 2011). Our results suggest that competition for these rate-limiting factors, between canonical and non-canonical *ARS*s, underlies the DNA replication defects observed in the absence of H3K37me1.

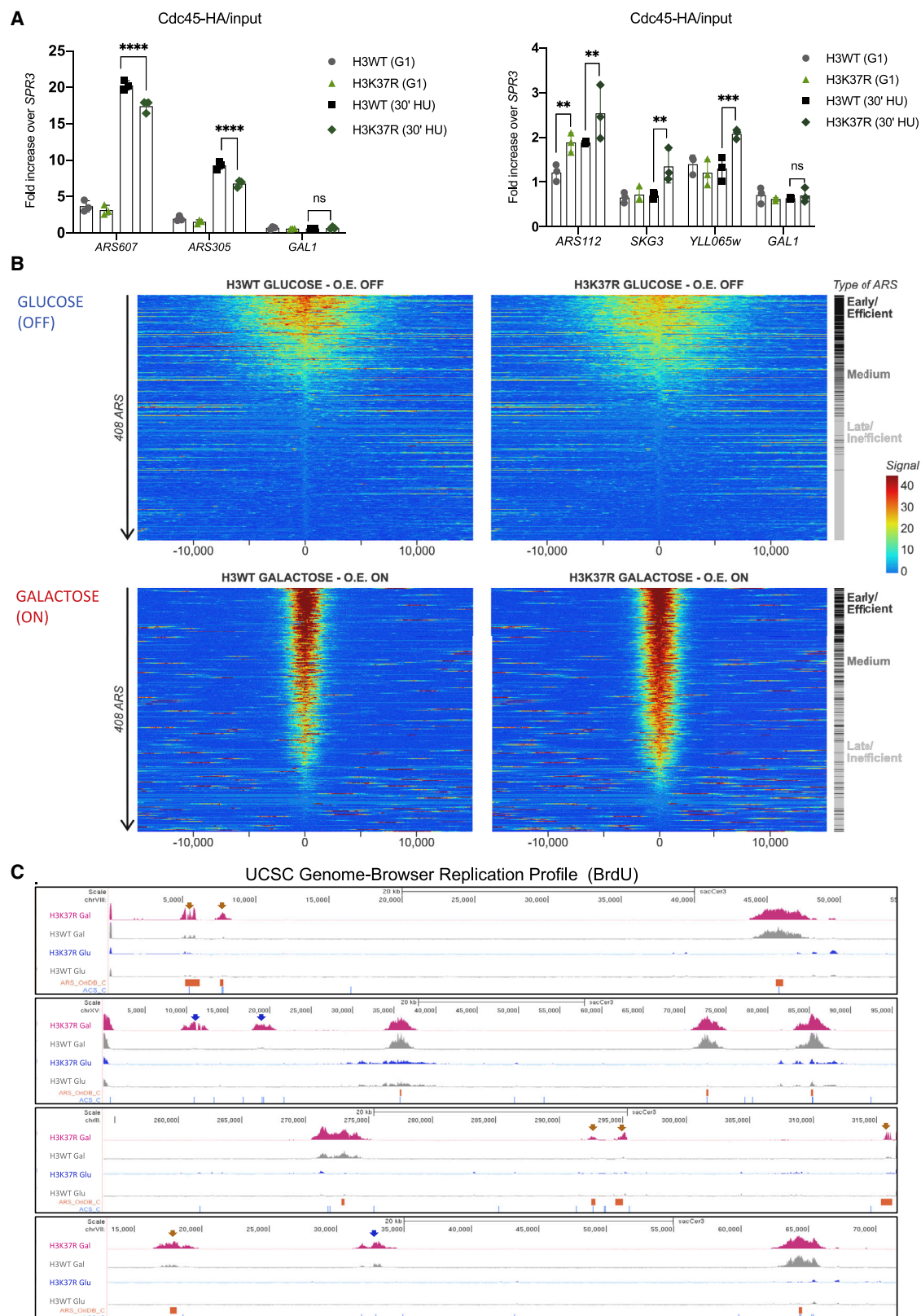
### Figure 5. H3K37me1 regulates origin establishment

(A) Left: Mcm2-His<sub>6</sub> *in vitro* binding to biotinylated H3 peptides, modified as indicated. Input and peptide-bound Mcm2 were resolved using SDS-PAGE in 8% acrylamide and detected by Coomassie brilliant blue (CBB) staining. One-fifth of each bead slurry was spotted onto polyvinylidene fluoride (PVDF) membrane, and biotin-peptides were detected using anti-biotin antibody. Right: average ImageJ quantification of assays shown at left and in Figure S8C. Binding is represented as percentage of the signal corresponding to one-quarter of the input.

(B) Boxplot of Mcm2-HA occupancy at *ARS* ± 1 kb. Blue, H3WT; green, H3K37R.

(C) Boxplot of Mcm2 occupancy at non-*ARS* (TSS +600 to +1,000 bp) in H3WT and H3K37R mutant. The plots in (B) and (C) represent the average of biological triplicates.

(D) ChIP-qPCR analysis of Mcm2-HA levels at "H3K37R unique" replication sites in G1-arrested cells. Data were normalized to the non-*ARS* region (*SPR3*). Statistical analysis was performed using two-way ANOVA corrected for the comparisons using the Holm-Sidak method (alpha = 0.05); \*p ≤ 0.05, \*\*p ≤ 0.01, \*\*\*p ≤ 0.001, and \*\*\*\*p ≤ 0.0001. Error bars represent the mean ± SD of two independent experiments.



(legend on next page)

Altogether, our results support a model in which the absence of H3K37me1 makes chromatin broadly more permissive to interactions with the replicative helicase, facilitating MCM loading and origin licensing at cryptic genomic ACS, so rendering such normally non-replicative locations competent to fire. Hence, we propose a role for H3K37me1 in safeguarding a proper origin firing program by protecting the genome from unscheduled association of MCM to chromatin, thereby preventing spurious initiation of DNA replication from cryptic origin locations.

The identification of Set1p and Set2p as H3K37 methyltransferases was quite unexpected, as these enzymes were discovered because of their ability to methylate H3K4 and H3K36, respectively. However, Set1p also methylates the kinetochore protein Dam1, while the Set2p human homolog SETD2, methylates tubulin during mitosis (Park et al., 2016; Zhang et al., 2005). Thus, the panel of Set1p and Set2p substrates is broader than first anticipated. Importantly, several reports implicate the two enzymes in cell cycle regulation and DNA replication. First, Set1p and Set2p enzymatic activities are required for cell cycle progression via transcriptional and non-transcriptional pathways (Biswas et al., 2008; Park et al., 2016; Rizzardi et al., 2012; Zhang et al., 2005). Second, H3K36me1 and H3K4me2 are specifically enriched at ARSs, where they contribute to firing, although the precise mechanisms are not fully understood (Biswas et al., 2008; Pryde et al., 2009; Rizzardi et al., 2012). Moreover, although Set1p and Set2p are recruited to chromatin by interacting with RNA polymerase II, COMPASS physically interacts with Orc2p (Kan et al., 2008), and a Set2 mutant lacking its phospho-CTD interaction domain still binds to chromatin (Youldell et al., 2008), arguing that both enzymes can be recruited to chromatin in a transcription-independent manner. Hence, although Set1p and Set2p are involved in a plethora of cellular functions, a more direct role in controlling DNA replication is increasingly supported.

Although methylation of H3K4 and/or H3K36 is not a prerequisite for H3K37me1, both lysines appear important for Set1p and Set2p activity, respectively. Although the positioning of H3K36 directly adjacent to H3K37 might explain its relevance to Set2p's substrate recognition, it is intriguing that mutation of H3K4 affects Set1p activity toward K37. There is no reason to suspect that Set1p uses distinct mechanisms to catalyze H3K37 and H3K4 mono-methylation, as the requirement of COMPASS members for methylation is the same for both sites. It was recently shown that COMPASS operates as a dimer *in vivo*

(Choudhury et al., 2019), so one possibility is that the enzyme is recruited by binding of one COMPASS to H3K4, while the second module catalyzes methylation of H3K37.

During the course of this work, it was reported that *S. pombe* Set7 is a H3K37 methyltransferase involved in gametogenesis (Shen et al., 2019). *Pombe* Set7 shares a similar degree of homology with all *S. cerevisiae* SET family members, as it does with human counterparts, so there is no clear ortholog. We found that H3K37me1 is also upregulated at the onset of premeiotic DNA replication in *S. cerevisiae*. In this regard, it is notable that deletion of *SET1* leads to sporulation defects due to impaired initiation of premeiotic DNA replication (Sollier et al., 2004). Interestingly, this Set1p function requires its catalytic domain but not H3K4 methylation (Sollier et al., 2004). Moreover, *set1Δ* sporulation defect is exacerbated by simultaneous deletion of *SET2* (Morohashi et al., 2005). These data support a model in which H3K37me1 deposition by Set1 and Set2 regulates DNA replication not only during mitosis but also during meiosis.

We found that H3K37me1 and its cell cycle regulation are evolutionary conserved, raising the possibility that H3K37me1 may participate in the selection of replication origins in higher eukaryotes (for a review, see Cayrou et al., 2011; Ding and MacAlpine, 2011). This might be particularly relevant in human cancer, in which deregulated intragenic origin firing generates DNA replication stress that contributes critically to genomic instability (Macheret and Halazonetis, 2018). Given that SETD1A and SETD2 are implicated in human malignancies, the unraveling of a function for these enzymes in coordinating origin firing may be of therapeutic importance.

### Limitations of study

Histone-modifying enzymes are involved in the regulation of diverse DNA-templated processes, often modifying more than one histone residue or even additional non-histone substrates. Hence, phenotypes resulting from enzyme mutation cannot always be attributed to the lack of a particular histone mark. Here, we mutated histone H3K37 to arginine to investigate the role of H3K37me1 in DNA replication. While retaining the residue's positive charge, H3K37R mutation would prevent not only methylation but also any other possible post-translational modification at lysine 37, which might also contribute to the observed phenotypes. Note, however, that no other post-translational modification has been identified so far at this particular site.

### Figure 6. Overexpression of MCM activators suppresses H3K37R replication defects

(A) ChIP-qPCR analysis of Cdc45-HA levels at efficient ARS (left panel) and "H3K37R unique" replication sites (right panel). H3WT and H3K37R cells expressing Cdc45-HA were arrested in G1 and released into HU-containing medium for 30 min. Data were normalized to a non-ARS region (*SPR3*). Statistical analysis was performed using two-way ANOVA corrected for the comparisons using the Holm-Sidak method ( $\alpha = 0.05$ ); \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ , and \*\*\*\* $p \leq 0.0001$ . Error bars represent the mean  $\pm$  SD of three independent experiments. Note the scales in the left and right panels.

(B) Heatmap showing the distribution of BrdU incorporation in wild-type and H3K37R mutant under non-overexpression (GLUCOSE/OFF) and overexpression (GALACTOSE/ON) of MCM activators. ARSs were aligned from highest to lowest BrdU signal in the wild-type, centered at ACS and classified into early/efficient, medium, and late/inefficient following the distribution shown in Figure S6C. For visualization purposes, the heatmaps' color scale was saturated at the 99th percentile of the distribution of signal intensities. The plot represents the average of three independent experiments.

(C) Representative Genome Browser snapshots of BrdU incorporation in H3WT and isogenic H3K37R mutant at different genomic locations. ARS (OriDB) are represented as orange boxes; non-replicative ACS matches are represented as blue lines. H3K37R exclusive firing event are indicated by the corresponding color-coded arrows. The plots represent the average of biological triplicates per strain and condition.



## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Yeast histone purification
  - Western blot and dot-blot
  - Flow cytometry analysis
  - Recombinant protein purification
  - PtA-Set1p complex (COMPASS) purification
  - Reconstitution of recombinant nucleosomes
  - *In vitro* methyltransferase assays
  - *In vitro* peptide pull-downs
  - Yeast chromatin immunoprecipitation
  - *S. pombe* spike
  - BrdU-DNA labeling and immunoprecipitation
  - Library preparation
  - Analysis of BrdU IP data
  - Co-localization of H3K37R unique peaks on ARS or ACS
  - ARS annotation and efficiency estimates
  - Analysis of H3K37me1 ChIP-seq
  - Analysis of Mcm2 ChIP-seq
  - Human cell chromatin immunoprecipitation
  - Direct enzyme-linked immunosorbent assay (ELISA)
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2021.04.021>.

## ACKNOWLEDGMENTS

Our gratitude to F. Puddu, S. Siniosoglou, P. Zegerman, P. Tessarz, and A.J. Bannister for critically reading the manuscript and technical help. We are indebted to Hiroyuki Araki for the generous gift of the antibodies to detect the MCM activators. The Kouzarides laboratory is supported by Cancer Research UK (grants RG72100 and RG96894) and core support from the Wellcome Trust (core grant WT203144) and Cancer Research UK (grant C6946/A24843). For the purpose of open access, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. G.M.-Z. was funded by a European Molecular Biology Organization (EMBO) long-term fellowship (ALTF907-2014) and Asociación Española Contra el Cáncer (POSTD18021MILL). K.T. was supported by a Sir Henry Wellcome Fellowship (grant RG94424). T.B. was supported by the European Research Council (ERC; StG 309952) and the Helmholtz Association.

## AUTHOR CONTRIBUTIONS

Conceptualization and project administration, H.S.-R.; methodology, H.S.-R., G.M.-Z., and N.H.; investigation, H.S.-R., G.M.-Z., and M.K.; software, formal analysis, and data curation, N.H., T.L., and L.P.; writing – review & editing, H.S.-R., G.M.-Z., T.B., and T.K.; resources, S.N. and T.B.; general supervision, H.S.-R. and T.K.; funding, T.K.

## DECLARATION OF INTERESTS

T.K. is a co-founder and shareholder of Abcam Plc and a co-founder of Storm Therapeutics Ltd. (Cambridge, UK). T.L. is a consultant for Storm Therapeutics Ltd.

Received: May 13, 2020

Revised: February 9, 2021

Accepted: April 21, 2021

Published: May 11, 2021

## REFERENCES

- Azmi, I.F., Watanabe, S., Maloney, M.F., Kang, S., Belsky, J.A., MacAlpine, D.M., Peterson, C.L., and Bell, S.P. (2017). Nucleosomes influence multiple steps during replication initiation. *eLife* 6, e22512.
- Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470–484.
- Bell, S.P., and Labib, K. (2016). Chromosome duplication in *Saccharomyces cerevisiae*. *Genetics* 203, 1027–1067.
- Belsky, J.A., MacAlpine, H.K., Lubelsky, Y., Hartemink, A.J., and MacAlpine, D.M. (2015). Genome-wide chromatin footprinting reveals changes in replication origin architecture induced by pre-RC assembly. *Genes Dev.* 29, 212–224.
- Biswas, D., Takahata, S., Xin, H., Dutta-Biswas, R., Yu, Y., Formosa, T., and Stillman, D.J. (2008). A role for Chd1 and Set2 in negatively regulating DNA replication in *Saccharomyces cerevisiae*. *Genetics* 178, 649–659.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Breier, A.M., Chatterji, S., and Cozzarelli, N.R. (2004). Prediction of *Saccharomyces cerevisiae* replication origins. *Genome Biol.* 5, R22.
- Broad Institute (2019). Picard. <https://broadinstitute.github.io/picard/>.
- Cayrou, C., Coulombe, P., Vigneron, A., Stanojic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R., and Méchali, M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Res.* 21, 1438–1449.
- Choudhury, R., Singh, S., Arumugam, S., Roguev, A., and Stewart, A.F. (2019). The Set1 complex is dimeric and acts with Jhd2 demethylation to convey symmetrical H3K4 trimethylation. *Genes Dev.* 33, 550–564.
- Das, S.P., and Rhind, N. (2016). How and why multiple MCMs are loaded at origins of DNA replication. *BioEssays* 38, 613–617.
- Das, S.P., Borrmann, 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.
- Diffley, J.F.X. (2011). Quality control in the initiation of eukaryotic DNA replication. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 3545–3553.
- Ding, Q., and MacAlpine, D.M. (2011). Defining the replication program through the chromatin landscape. *Crit. Rev. Biochem. Mol. Biol.* 46, 165–179.
- Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375, 23–44.
- 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.
- Freitag, M. (2017). Histone methylation by SET domain proteins in fungi. *Annu. Rev. Microbiol.* 71, 413–439.
- Ishimi, Y., Komamura, Y., You, Z., and Kimura, H. (1998). Biochemical function of mouse minichromosome maintenance 2 protein. *J. Biol. Chem.* 273, 8369–8375.
- Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., and Knop, M. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new



fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* 21, 947–962.

Kan, J., Zou, L., Zhang, J., Wu, R., Wang, Z., and Liang, C. (2008). Origin recognition complex (ORC) mediates histone 3 lysine 4 methylation through cooperation with Spp1 in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 283, 33803–33807.

Kane, S.M., and Roth, R. (1974). Carbohydrate metabolism during ascospore development in yeast. *J. Bacteriol.* 118, 8–14.

Kirmizis, A., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bähler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* 449, 928–932.

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.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* 9, e1003118.

Lengronne, A., Pasero, P., Bensimon, A., and Schwob, E. (2001). Monitoring S phase progression globally and locally using BrdU incorporation in TK(+) yeast strains. *Nucleic Acids Res.* 29, 1433–1442.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Li, Q.Q., Hao, J.J., Zhang, Z., Krane, L.S., Hammerich, K.H., Sanford, T., Trepel, J.B., Neckers, L., and Agarwal, P.K. (2017). Proteomic analysis of proteome and histone post-translational modifications in heat shock protein 90 inhibition-mediated bladder cancer therapeutics. *Sci. Rep.* 7, 201.

Lipford, J.R., and Bell, S.P. (2001). Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* 7, 21–30.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.

Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* 44, e45.

Macheret, M., and Halazonetis, T.D. (2018). Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* 555, 112–116.

Mantiero, D., Mackenzie, A., Donaldson, A., and Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J.* 30, 4805–4814.

McDaniel, S.L., and Strahl, B.D. (2017). Shaping the cellular landscape with Set2/SETD2 methylation. *Cell. Mol. Life Sci.* 74, 3317–3334.

Mori, S., and Shirahige, K. (2007). Perturbation of the activity of replication origin by meiosis-specific transcription. *J. Biol. Chem.* 282, 4447–4452.

Morohashi, N., Mitchell, A.P., and Shimizu, M. (2005). Effect of histone methyltransferase gene mutations on sporulation in *S. cerevisiae*. *Nucleic Acids Symp. Ser. (Oxf.)* (49), 325–326.

Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., and Guenther, M.G. (2014). Quantitative ChIP-seq normalization reveals global modulation of the epigenome. *Cell Rep.* 9, 1163–1170.

Park, I.Y., Powell, R.T., Tripathi, D.N., Dere, R., Ho, T.H., Blasius, T.L., Chiang, Y.C., Davis, I.J., Fahey, C.C., Hacker, K.E., et al. (2016). Dual chromatin and cytoskeletal remodeling by SETD2. *Cell* 166, 950–962.

Petrossian, T.C., and Clarke, S.G. (2009). Multiple motif scanning to identify methyltransferases from the yeast proteome. *Mol. Cell. Proteomics* 8, 1516–1526.

Pryde, F., Jain, D., Kerr, A., Curley, R., Mariotti, F.R., and Vogelauer, M. (2009). H3 k36 methylation helps determine the timing of cdc45 association with replication origins. *PLoS ONE* 4, e5882.

Puddu, F., Herzog, M., Selivanova, A., Wang, S., Zhu, J., Klein-Lavi, S., Gordon, M., Meirman, R., Millan-Zambrano, G., Ayestaran, I., et al. (2019). Genome architecture and stability in the *Saccharomyces cerevisiae* knockout collection. *Nature* 573, 416–420.

Ren, C., Liu, S., Ghoshal, K., Hsu, P.-H., Jacob, S.T., Marcucci, G., and Freitas, M.A. (2007). Simultaneous metabolic labeling of cells with multiple amino acids: localization and dynamics of histone acetylation and methylation. *Proteomics Clin. Appl.* 1, 130–142.

Riera, A., Barbon, M., Noguchi, Y., Reuter, L.M., Schneider, S., and Speck, C. (2017). From structure to mechanism-understanding initiation of DNA replication. *Genes Dev.* 31, 1073–1088.

Rizzardi, L.F., Dorn, E.S., Strahl, B.D., and Cook, J.G. (2012). DNA replication origin function is promoted by H3K4 di-methylation in *Saccharomyces cerevisiae*. *Genetics* 192, 371–384.

Rossmann, M.P., and Stillman, B. (2013). Immunoblotting histones from yeast whole-cell protein extracts. *Cold Spring Harb. Protoc.* 2013, 625–630.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C.T., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411.

Schneider, J., Wood, A., Lee, J.S., Schuster, R., Dueker, J., Maguire, C., Swanson, S.K., Florens, L., Washburn, M.P., and Shilatifard, A. (2005). Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol. Cell* 19, 849–856.

Shen, Y., Mevius, D.E.H.F., Caliendo, R., Carozzini, B., Roh, Y., Kim, J., Kim, S., Ha, S.C., Morishita, M., and di Luccio, E. (2019). Set7 Is a H3K37 methyltransferase in *Schizosaccharomyces pombe* and is required for proper gametogenesis. *Structure* 27, 631–638.e8.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* 81, 65–95.

Shoaib, M., Walter, D., Gillespie, P.J., Izard, F., Fahrenkrog, B., Lleres, D., Lerdrup, M., Johansen, J.V., Hansen, K., Julien, E., et al. (2018). Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing. *Nat. Commun.* 9, 3704.

Siow, C.C., Nieduszynska, S.R., Müller, C.A., and Nieduszynski, C.A. (2012). OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Res.* 40, D682–D686.

Sollier, J., Lin, W., Soustelle, C., Suhre, K., Nicolas, A., Géli, V., and de La Roche Saint-André, C. (2004). Set1 is required for meiotic S-phase onset, double-strand break formation and middle gene expression. *EMBO J.* 23, 1957–1967.

Soriano, I., Morafraila, 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.

Soudet, J., and Stutz, F. (2019). Regulation of gene expression and replication initiation by non-coding transcription: a model based on reshaping nucleosome-depleted regions: influence of pervasive transcription on chromatin structure. *BioEssays* 41, e1900043.

Statham, A.L., Strbenac, D., Coolen, M.W., Stirzaker, C., Clark, S.J., and Robinson, M.D. (2010). Repitools: an R package for the analysis of enrichment-based epigenomic data. *Bioinformatics* 26, 1662–1663.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.-W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2002).

Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol. Cell. Biol.* **22**, 1298–1306.

Tajima, K., Yae, T., Javadi, S., Tam, O., Comaills, V., Morris, R., Wittner, B.S., Liu, M., Engstrom, A., Takahashi, F., et al. (2015). SETD1A modulates cell cycle progression through a miRNA network that regulates p53 target genes. *Nat. Commun.* **6**, 8257.

Tajima, K., Matsuda, S., Yae, T., Drapkin, B.J., Morris, R., Boukhali, M., Niederhoffer, K., Comaills, V., Dubash, T., Nieman, L., et al. (2019). SETD1A protects from senescence through regulation of the mitotic gene expression program. *Nat. Commun.* **10**, 2854.

Takahashi, Y.H., Westfield, G.H., Oleskie, A.N., Trievel, R.C., Shilatifard, A., and Skiniotis, G. (2011). Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc. Natl. Acad. Sci. USA* **108**, 20526–20531.

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

Tessarz, P., and Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15**, 703–708.

Thomas, B.J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* **56**, 619–630.

Unnikrishnan, A., Gafken, P.R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. *Nat. Struct. Mol. Biol.* **17**, 430–437.

Viggiani, C.J., and Aparicio, O.M. (2006). New vectors for simplified construction of BrdU-Incorporating strains of *Saccharomyces cerevisiae*. *Yeast* **23**, 1045–1051.

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

Xu, W., Aparicio, J.G., Aparicio, O.M., and Tavaré, S. (2006). Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in *S. cerevisiae*. *BMC Genomics* **7**, 276.

Yoshida, K., Bacal, J., Desmarais, D., Padioleau, I., Tsaponina, O., Chabes, A., Pantescio, V., Dubois, E., Parrinello, H., Skrzypczak, M., et al. (2014). The histone deacetylases sir2 and rpd3 act on ribosomal DNA to control the replication program in budding yeast. *Mol. Cell* **54**, 691–697.

Youde, M.L., Kizer, K.O., Kisseleva-Romanova, E., Fuchs, S.M., Duro, E., Strahl, B.D., and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol. Cell. Biol.* **28**, 4915–4926.

Zhang, K., Lin, W., Latham, J.A., Riefler, G.M., Schumacher, J.M., Chan, C., Tatchell, K., Hawke, D.H., Kobayashi, R., and Dent, S.Y.R. (2005). The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* **122**, 723–734.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-H3 K37me1	This study/Abcam	Ab215728
Anti-H3	Abcam	RRID:AB_302613
Anti-H3 K36me1	Abcam	RRID:AB_306964
Anti-H3 K36me2	Abcam	RRID:AB_1280939
Anti-H3 K36me3	Abcam	RRID:AB_306966
Anti-H3 K4me1	Abcam	RRID:AB_306847
Anti-H3 K4me2	Abcam	RRID:AB_2560996
Anti-H3 K4me3	Abcam	RRID:AB_306649
Anti-SETD2	Abcam	RRID:AB_2185782
Anti-SETD1A	Abcam	RRID:AB_1951955
Anti-TUBULIN	SIGMA	RRID:AB_477593
Anti-HA	Abcam	RRID:AB_307019
Anti-GFP	Abcam	RRID:AB_303395
Anti-BrdU	BD Bioscience	RRID:AB_395993
Anti-ORC6	Stillman B.	N/A
Anti-Biotin	SIGMA	RRID:AB_258625
Anti-PAP	SIGMA	RRID:AB_1079562
<b>Bacterial strains</b>		
BL21 (DE3)	NEB	Cat.No. C2571
<b>Yeast strains</b>		
W303	<i>Mat a ade2-1 can1-100 his3-11,15 leu2-3, 112 trp1-1 ura3-1</i>	Thomas and Rothstein, 1989
SK1	<i>Mat a/α HO gal2 cup<sup>S</sup> can1<sup>R</sup> BIO</i>	Kane and Roth, 1974
ID 15459	<i>h90 ade6-216 leu1-32 lys1-131 ura4-d18 mcm3::MCM3-GFP-HA-KAN MX</i>	NRBP; <a href="https://yeast.nig.ac.jp/yeast/">https://yeast.nig.ac.jp/yeast/</a>
HSR36	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2-HHF2</i>	Tessarar and Kouzarides, 2014
HSR364	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2 K37A-HHF2</i>	This study
HSR436	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K37R-HHF2</i>	This study
HSR660	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K36R-HHF2</i>	This study
HSR662	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K4R-HHF2</i>	This study
HSR688	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K4R-HHF2 set2::HIS3</i>	This study
HSR691	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K36R-HHF2 set1::3ARU</i>	This study
HSR416	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2-HHF2 MCM2-HA<sub>6</sub>::HYGRO</i>	This study
HSR453	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2-HHF2 HIS3::GPD-TK<sub>7</sub></i>	This study

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HSR454	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K37R-HHF2 HIS3::GPD-TK<sub>7</sub></i>	This study
HSR438	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K37R-HHF2 MCM2-HA<sub>6</sub>::HYGRO</i>	This study
HSR707	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2-HHF2 CDC45-HA<sub>6</sub>::HYGRO</i>	This study
HSR708	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 TRP1-HHT2K37R-HHF2 CDC45-HA<sub>6</sub>::HYGRO</i>	This study
HSR645	<i>Mat a ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 set1::3ARU</i>	This study
HSR643	<i>Mat a ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 set2::HIS3</i>	This study
HSR646	<i>Mat a ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura31 set1::3ARU set2::HIS3</i>	This study
HSR495 (7Δ)	<i>Mat a ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura31 set1N1016Q::LEU2 set2::HIS3 set3::NAT set4::KAN MX set5::TRP1 set6::URA3 set7::ADE2</i>	This study
HSR727	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 LEU2- Gal1-10 CDC45/SLD7 TRP1- Gal1-10 SLD2/DPB11 URA3-Gal1-10 SLD3/DBF4 ADE2-HHT2-HHF2 HIS3::GPD-TK<sub>7</sub></i>	This study
HSR729	<i>Mat a hht2,hhf2::natMX4, hht1,hhf1::kanMX4 LEU2- Gal1-10 CDC45/SLD7 TRP1- Gal1-10 SLD2/DPB11 URA3-Gal1-10 SLD3/DBF4 ADE2-HHT2K37R-HHF2 HIS3::GPD-TK<sub>7</sub></i>	This study
Plasmid	Description	Source
pMAL-SET2	<i>Ptac-malE-SET2</i>	This study
pMAL-Set2 N198Q	<i>Ptac-malE-set2N198Q</i>	This study
pMAL-Set2 Y149A	<i>Ptac-malE-set2Y149A</i>	This study
PtA-SET1	PNOP1-PtA-SET1-LEU2	<a href="#">Santos-Rosa et al., 2002</a>
PtA-SET1 ΔC92	PNOP1-PtA-set11ΔC92-LEU2	<a href="#">Santos-Rosa et al., 2002</a>
pET21-hH4	pET21b(+)- Wild Type human H4	<a href="#">Bartke et al., 2010</a>
pET21-hH2A	pET21b(+)- Wild Type human H2A	<a href="#">Bartke et al., 2010</a>
pET21-hH3.1	pET21b(+)- Wild Type human H3.1	<a href="#">Bartke et al., 2010</a>
pET21-hH2B	pET21b(+)- Wild Type human H2B	<a href="#">Bartke et al., 2010</a>
pUC19-16x601	pUC19 containing 16 repeats of the 601 nucleosome positioning sequence	<a href="#">Bartke et al., 2010</a>
pET21-hH3.1 K37R	pET21b(+)-human H3.1 K37R	This study
p402-H3WT	pRS402- HHT2-HHF2	This study
p402-H3K37R	pRS402- HHT2K37R-HHF2	This study
pET30-Mcm2 (1-480)	pET30z(+)-MCM2 (aa1-aa408)-His6	This study
p403-BrdU-Inc	<i>HIS3::GDP-TK/ADH1-hENT1</i>	<a href="#">Viggiani and Aparicio, 2006</a>
bpZ46	pRS305-Gal1-10-Cdc45/Sld7	<a href="#">Mantiero et al., 2011</a>
bpZ109	pRS304-Gal1-10-Sld2/Dpb11	<a href="#">Mantiero et al., 2011</a>
bpZ287	pRS306-Gal1-10-Sld3/Dbf4	<a href="#">Mantiero et al., 2011</a>
Chemicals, peptides, and recombinant proteins	Source	Cat. no.
Biotinylated H3 K37me1 peptide (aa28-aa48)	GeneCust	N/A
Biotinylated H3 K36me1 peptide (aa28-aa48)	GeneCust	N/A

(Continued on next page)

### Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biotinylated H3 K36me1K37me1 peptide (aa28-aa48)	GeneCust	N/A
Biotinylated H3 unmodified peptide (aa28-aa48)	GeneCust	N/A
Biotinylated H3 K36me2 peptide (aa28-aa48)	GeneCust	N/A
Biotinylated H3 K36me3 peptide (aa28-aa48)	GeneCust	N/A
Sulfolink Coupling Resin	ThermoFisher	20402
Protease inhibitors	Roche	11697498001
SAM	NEB	B9003S
<sup>3</sup> H-SAM	PerkinElmer	NET155V250UC
5-Bromo-2'-deoxyuridine	Sigma	10280879001
$\alpha$ -factor	GenScript	RP01002
Sytox Green	Invitrogen	S7020
Amylose Resin High Flow	NEB	E8022L
IgG Sepharose 6 Fast Flow	GE Healthcare	17-0969-01
Ni Sepharose 6 Fast Flow	GE Healthcare	17-5318-01
Dynabeads M-280	DYNAL	112.02
Calf histones	Roche	10223565001
Data and code availability	Source	Identifier
BrdUseq	This study	ArrayExpress: E-MTAB-9097
BrdUseq overexpression limiting factors	This study	ArrayExpress: E-MTAB-9999
K37me1 ChIPseq	This study	ArrayExpress: E-MTAB-9058
MCM2 ChIPseq wt versus H3K37R	This study	ArrayExpress: E-MTAB-9070
MCM2 ChIPseq time course	This study	ArrayExpress: E-MTAB-9071
Mendeley Original Data	This study	<a href="https://doi.org/10.17632/vrsxrh7xg.1">https://doi.org/10.17632/vrsxrh7xg.1</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Tony Kouzarides ([t.kouzarides@gurdon.cam.ac.uk](mailto:t.kouzarides@gurdon.cam.ac.uk)).

### Materials availability

All unique reagents generated in the study will be available without restrictions on request to Lead Contact.

### Data and code availability

The accession numbers for the sequencing data reported in this paper are:

Arrayexpress: E-MTAB-9097  
 ArrayExpress: E-MTAB-9999  
 ArrayExpress: E-MTAB-9058  
 ArrayExpress: E-MTAB-9070  
 ArrayExpress: E-MTAB-9071

Unprocessed imaging data are deposited on Mendeley Data: <https://doi.org/10.17632/vrsxrh7xg.1>.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

All *Saccharomyces cerevisiae* strains used in this study are derivatives of W303 background except for those used in the screen for H3K37 methyltransferases (Figure S3A), which are BY4743 and curated isogenic deletion strain collection (Puddu et al., 2019). Methyltransferases not properly disrupted or missing from the collection were deleted in BY4741 for this study and compared to the isogenic WT.



Integrations and deletions were performed using one-step PCR-based methods (Janke et al., 2004). Histone point mutants were shuffled by counter-selection on 5-FOA. Yeast strains overexpressing the helicase activators were constructed by genomic integration of the following plasmids (Mantiero et al., 2011): bpZ46 (pRS305 Gal1-10 Cdc45/Sld7, KasI digested); bpZ109 (pRS304 Gal1-10 Sld2/Dpb11, HindIII digested); bpZ287 (pRS306 Sld3/Dbf4, NdeI digested).

Yeast genotypes are included in the Key Resources Table. All strains were routinely grown in YPAD at 30°C.

hTERT immortalized RPE-1 cells were cultured in Dulbecco's Modified Eagle's Medium/Nutrient Mixture F-12 Ham (D8437; Sigma-Aldrich) supplemented with 10% (vol/vol) fetal bovine serum (Sigma-Aldrich), 100U/ml penicillin, 100μg/ml streptomycin (Sigma-Aldrich) and 2mM L-glutamine.

## METHOD DETAILS

### Yeast histone purification

Yeast cells were grown in 1l of YPD for 3-4 generations to OD<sub>600</sub> = 1. Cells were then collected, washed with water and frozen in liquid nitrogen. The cell pellet was resuspended in SP buffer (1M sorbitol; 50mM potassium phosphate pH 6.8, 14mM β-mercaptoethanol) and spheroplasted by zymolyase digestion. Nuclei were then prepared by Douncing in Lysis buffer (18% Ficoll-400 [w/v]; 20mM potassium phosphate pH 6.8; 1mM MgCl<sub>2</sub>; 0.5mM EDTA) supplemented with both protease and phosphatase inhibitors (Roche). After spinning down in a benchtop centrifuge, supernatant was recovered and nuclei were pelleted by spinning at 50,000 g in a Beckman SW-41 Ti rotor. Nuclei were then resuspended in NP buffer (0.34M sucrose; 20mM Tris-HCl pH 7.4; 50mM KCl; 5mM MgCl<sub>2</sub>) supplemented with both protease and phosphatase inhibitors (Roche) and pelleted by spinning at 30,000 g in a Beckman SW-41 Ti rotor. Nuclei were washed three times with Buffer A (10mM Tris-HCl pH 8.0; 0.5% NP-40 [v/v]; 75mM NaCl) supplemented with both protease and phosphatase inhibitors (Roche). Histones were then extracted in Buffer B (10mM Tris-HCl pH 8; 400mM NaCl; 0.2M H<sub>2</sub>SO<sub>4</sub>) and TCA precipitated.

### Western blot and dot-blot

Total protein extracts were prepared as previously described (Rossmann and Stillman, 2013). Unless anything different is specified, yeast cultures were grown in YPAD to OD<sub>600nm</sub>:0.6-0.75. 7 ODs of cells were spun, washed with PBS and resuspended in 200 μl of Laemmli buffer, in the presence of 300 μl acid washed glass beads. Cells were disrupted by mechanical vortex and subsequent boiling at 95°C [3x (2min vortex+2min boil)]. Extracts were spun at full speed for 5 min and the supernatant, containing solubilized proteins, were transferred to fresh Eppendorf tubes. Protein extracts were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the specified acrylamide % gels. Biotinylated peptides were separated in 17% Tris-Tricine gels and transferred to PVDF membranes as described (Kirmizis et al., 2007). Dot-blot analysis of peptides was performed by spotting serial dilutions onto a PVDF membrane, air drying them and processed following manufacturer's instructions. Western blots were developed within the linear range of exposure using a ChemiDoc MP imaging system (Bio-Rad).

### Flow cytometry analysis

Cell pellets were resuspended in ice-cold 70% [v/v] ethanol and incubated at 4°C overnight. Cells were then washed with 50mM Tris-HCl pH 8.0 and incubated with 0.4mg/ml RNase A at 37°C overnight. After treatment with pepsin for 30 min at 37°C, cells were resuspended in 1μM Sytox Green solution and analyzed using BD FACSCalibur flow cytometer.

### Recombinant protein purification

Plasmids used for recombinant expression are described in Key Resources Table. Expression of MBP-SET2, MBP-SET2N198Q and MBP-SET2Y149A fusions in BL21 (DE3) cells was induced with 0.3mM IPTG over night at 18°C. Cells were then lysed by sonication in TBS + 1% Triton X-100 supplemented with protease inhibitors, and clarified lysate was incubated with Amylose Resin for 2h at 4°C. Immobilized fusion proteins were washed extensively with TBS + 1% Triton X-100 and then with TBS. Immobilized recombinant proteins were then used for methyltransferase reactions.

Expression of MCM2 (aa1-aa480)-His6 fusions in BL21 (DE3) cells was induced with 1mM IPTG over night at 18°C. Cells were then lysed by sonication in TBS + 1% Triton X-100 +10mM imidazole supplemented with protease inhibitors, and clarified lysate was incubated with Ni-Sepharose beads for 2h at 4°C. Immobilized fusion proteins were washed extensively with TBS + 1% Triton X-100, then with TBS+20mM imidazole and eluted with TBS+400mM imidazole. Full length Mcm2 was dialyzed O/N at 4°C against TBS+10% glycerol. Mcm2 (aa1-aa480) was further purified by HPLC. Fractions containing Mcm2(1-480) were pooled together and its concentration measured by Bradford assay.

### PtA-Set1p complex (COMPASS) purification

Purification of the fusion proteins and complex interacting factors, onto IgG-Sepharose resin, were performed as previously described (Santos-Rosa et al., 2002). WT yeast expressing either PtA-Set1p or PtA-set1ΔC92 under the control of the PNOP promoter, were grown in SDC-Leu for plasmid selection, until OD<sub>600nm</sub>:1. Cell pellets corresponding to 10 l of culture/strain were washed with PBS, converted into spheroplasts and lysed in 300 mL lysis buffer (1% Triton X-100, 150 mM KCl, 5mM MgCl<sub>2</sub>, 20 mM Tris-HCl, pH 8.0) by performing several strokes with a Dounce homogenizer. The solubilized fraction was centrifuged for 20 min at 20000 rpm in

a SS34 rotor and the supernatant was loaded onto a 30 mm diameter column packed with 0.5ml IgG-Sepharose (Pharmacia, Germany) and equilibrated with washing buffer (0.5% Tween 20, 150 mM NaCl, 20 mM Tris-HCl, pH 8.0). Flow Through was discarded and the column was washed with 100 mL washing buffer and equilibrated with 2ml of methyltransferase buffer *without DTT* (50 mM Tris-HCl pH 8.5; 20mM KCl; 10mM MgCl<sub>2</sub>; 250mM sucrose; 0.2%  $\beta$ -maltoside). The beads bound ProtA fusion protein-complexes were then used for methyltransferase reactions (80  $\mu$ l of slurry per reaction).

### Reconstitution of recombinant nucleosomes

Recombinant human core histone proteins H2A, H2B, H3.1, and H4 were expressed in *E. coli* BL21(DE3)/RIL cells from pET21b(+) (Novagen) vectors and purified by denaturing gel filtration and ion exchange chromatography essentially as described (Dyer et al., 2004). The H3.1 K36R mutant was generated from the wild-type pET21-hH3.1 expression plasmid by site directed mutagenesis. Histone octamers (WT and H3 K37R mutant) were refolded from the purified histones, purified by size exclusion chromatography on a Superdex 200 gel filtration column (GE Healthcare), and then assembled into nucleosomes with biotinylated 601 DNA via salt deposition dialysis as previously described (Bartke et al., 2010; Dyer et al., 2004). Biotinylated nucleosomal DNA containing the 601 nucleosome positioning sequence was prepared from a pUC19 vector containing 16 tandem repeats of the 601 sequence as described (Bartke et al., 2010; Lowary and Widom, 1998).

### In vitro methyltransferase assays

Methylation reactions were performed for 1h at 30°C in JB buffer (50 mM Tris-HCl pH 8.5; 20mM KCl; 10mM MgCl<sub>2</sub>; 0.5mM DTT; 250mM sucrose; 0.2%  $\beta$ -maltoside) with either 32  $\mu$ M SAM (NEB) or 0.1MBq <sup>3</sup>H-SAM, using 35 ng of MBP-Set2 (full length WT) or MBP-Set2 Y149A (full length YA) enzyme and 250ng nucleosomes (equivalent to 62.5 ng human H3.1 WT or human H3.1 K37R). Reactions on histones were separated by SDS-PAGE in 16% polyacrylamide gels, whereas reactions on peptides were resolved in 17% polyacrylamide Tricine gels. For radioactive assays, gels were stained with Coomassie brilliant blue to visualize proteins and peptides, dried and exposed to film. For non-radioactive assays, gels were transferred to nitrocellulose membranes and analyzed by western blot as described above. Set1 catalyzed methyltransferase reactions were performed equally, but replacing the recombinant enzyme by 80  $\mu$ l of beads slurry containing PtA-Set1p or PtA-set1 $\Delta$ C92 purified as described above.

### In vitro peptide pull-downs

15 $\mu$ g of each biotinylated H3 peptide were diluted in 0.5ml of TBS+1% [v/v] NP-40 and incubated with 150 $\mu$ l of Dynabeads MyOne Streptavidin T1 for 60 minutes at room temperature. Beads were washed once with TBS/500mM NaCl+1% [v/v] NP-40, once with TBS/250mM NaCl+1% [v/v] NP-40, twice with TBS+1% [v/v] NP-40 and resuspended in 150 $\mu$ l TBS+1% [v/v] NP-40. 5 $\mu$ l beads slurry was spotted onto PVDF membrane and stained with Coomassie Blue to monitor binding for each peptide. For the pull downs, 50  $\mu$ l Mcm2(aa1-aa480)-His6, corresponding to 5  $\mu$ g of protein, were incubated with 20 $\mu$ l beads in a total volume of 300 $\mu$ l TBS/200mM NaCl+1% [v/v] NP-40 during 1 hour at room temperature on a bench wheel. Beads were washed 4 times with binding buffer and bound material eluted by boiling the beads in 25  $\mu$ l SDS Laemmli buffer, resolved by SDS-PAGE and visualized by Coomassie staining.

### Yeast chromatin immunoprecipitation

Yeast strains were grown in YPAD medium at 30°C and synchronized to G1 by addition of  $\alpha$ -factor (10 $\mu$ g/ml final concentration). 100ml of yeast cultures per ChIP experiment were collected at the indicated times. Cells were cross-linked with 1% formaldehyde for 15 minutes at room temperature, and the reaction was quenched with 125mM glycine. Cells were resuspended in ChIP SDS buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH 8.0) supplemented with protease inhibitors (Roche) and disrupted with glass beads by using a FastPrep instrument (MP Biomedicals). Chromatin was sonicated (Bioruptor Pico, Diagenode; 9 cycles, 30 s on/off) to yield an average DNA fragment of 300-500 base pairs, and diluted 10 times in ChIP IP buffer (0.01% SDS, 1.1% Triton X-100, 167mM NaCl, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.0) supplemented with protease inhibitors (Roche) prior to overnight immunoprecipitation on rotation at 4°C. Next day, 50  $\mu$ l of protein A Dynabeads were added, and samples were incubated again on rotation at 4°C for 2 hours. Dynabeads were washed twice with the following buffers: TSE 150 (1% Triton X-100, 0.1% SDS, 150mM NaCl, 2mM EDTA, 20mM Tris-HCl pH 8.0), TSE 500 (1% Triton X-100, 0.1% SDS, 500mM NaCl, 2mM EDTA, 20mM Tris-HCl pH 8.0) and LiCl buffer (0.25M LiCl, 1% NP-40, 1% deoxycholate, 1mM EDTA, 10mM Tris-HCl pH 8.0). DNA was eluted for 30 min at 37°C in elution buffer (100mM NaHCO<sub>3</sub>, 1% SDS), and cross-linking was reverted by overnight incubation at 65°C. Samples were treated with 0.5mg/ml of RNase A at 37°C for 2 h. DNA was purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Relative DNA amounts were determined by qPCR using Fast SYBR Green Master Mix (Applied Biosystems). Primer pairs used for amplification are available upon request. For each strain and/or condition, three independent colonies were grown and processed. The mean values  $\pm$  SEM derived from three biological replicates were plotted using Prism (GraphPad Software, Inc.).

### S. pombe spike

*S. cerevisiae* Mcm2-HA ChIPs were spiked with *S. pombe* expressing a Mcm3-HA allele as independent internal normalizer. Crosslink treatment (1% formaldehyde, 15 minutes) and sonication conditions (9 cycles 30 s ON/30 s OFF) yielded similar fragment size for both

yeasts strains as monitored by tape station (256-261 bp, below). Thus, a volume of *S.pombe*, corresponding to 5% of each *S.cerevisiae* culture, was added to each WT and H3K37R culture prior to formaldehyde treatment. All samples (containing *S.pombe* and *S.cerevisiae* cells) were processed for ChIP in parallel, following our standard procedure.

### BrdU-DNA labeling and immunoprecipitation

Immunoprecipitation of BrdU nascent DNA was adapted from Lengronne et al., 2001. Per each strain, 500ml of culture was synchronized in G1 by addition of  $\alpha$ -factor (10 $\mu$ g/ml final concentration) for 1h and 30'. Half culture (250ml) remained untreated (-BrdU) whereas 5ml BrdU (25mg/ml stock) was added to the other half (+BrdU). Both flasks were incubated for a further 30min at 30°C, then cells were collected by centrifugation and resuspended into 30°C warmed YPAD medium containing 25ml HU (2M stock) and incubated for 1.25 hours. BrdU (+) and (-) cultures were immediately transferred to ice/water bath and 2.5ml NaN<sub>3</sub> (10% stock) was added to each one. After 10min, cells were collected by centrifugation, washed twice with TBS and stored at -20°C. Incorporation of BrdU into DNA was monitored by dot blot with an anti-BrdU antibody as described (Lengronne et al., 2001).

For BrdU Immunoprecipitation, each of the above cell pellets was shaken five times for 2 min in NIB buffer (17% [v/v] glycerol, 50mM MOPS buffer, 150mM potassium acetate, 2mM MgCl<sub>2</sub>, 500 $\mu$ M spermidine, 150 $\mu$ M spermine at pH 7.2) with zirconium beads on a Vibrax (VXR basic, Ika) at 4°C. Genomic DNA extraction, from the above cultures was performed using QIAGEN Genomic DNA Isolation Kit and QIAGEN genomic-tip 100/G (Cat. No. 19060 and Cat. No.10243). DNA was fragmented using sonication (~200- to 500-base-pair [bp] size range), boiled at 95°C for 10 minutes, placed on ice for 10 minutes and precipitated O/N at 4°C with 5  $\mu$ g mouse anti-BrdU antibody coupled to Dynabeads M-280 Sheep anti-mouse IgG. Immunoprecipitated material was washed and eluted as described for ChIP.

Specificity of the IPs was tested comparing samples +BrdU and -BrdU at replication origins (ARS) and comparing samples + BrdU at ARS versus non replication origins.

For the immunoprecipitation of BrdU nascent DNA under overexpression of the helicase activators, 3 independent colonies of HSR727 and HRS729 were grown in YPA-Raffinose to OD600nm = 0.4 and then synchronized in G1 by addition of  $\alpha$ -factor (10 $\mu$ g/ml final concentration) for 1.5 hours. To half of a culture (100ml), glucose (2% final concentration) and 1.3ml BrdU (50mg/ml stock) was added. To the other half, galactose (2% final concentration) and 1.3ml BrdU (50mg/ml stock) was added. Both flasks were incubated for a further 30min at 30°C, then cells were collected by centrifugation and resuspended into 100ml of 30°C pre-warmed YPA-2%glucose or YPA-2%galactose medium containing 25ml HU (2M stock) plus 1.2ml BrdU (50mg/ml stock) and incubated for 70 minutes. YPA-Glucose and YPA-Galactose cultures were immediately transferred to ice/water bath and 2.5ml NaN<sub>3</sub> (10% stock) added to each one. After 10min, cells were collected by centrifugation, washed twice with TBS and store at -20°C. Incorporation of BrdU onto DNA was monitored by dot blot analysis with an anti-BrdU antibody as described (Lengronne et al., 2001).

Quantitative PCR was performed with primers specific to each genomic location. All sequences are available upon request.

### Library preparation

Library preparations followed the NEXTflex Rapid DNA-Seq protocol (2015; Bioo Scientific) with a few adjustments. Illumina adaptors were substituted with Bioo NEXTflex adaptors. Agilent Technologies High Sensitivity D1000 was used for determining the fragment size on the Agilent 4200 TapeStation. Libraries were quantified using Invitrogen Qubit dsDNA HS Assay. Equal molarity of each bar-coded library was pulled together to multiplex. Single-end 50 bp Next-generation sequencing was performed on a HiSeq 4000 (Illumina).

### Analysis of BrdU IP data

Raw reads were processed with Trimmomatic (v0.39) to remove adapters (Bolger et al., 2014), retaining reads longer than 20nt. Reads were then mapped to the reference genome (sacCer3) with BWA (-n 3 -k 2 -R 300) (Li and Durbin, 2009) and filtered to remove multimappers based on the alignments quality (samtools view -q 10) (Li et al., 2009). Finally, PCR duplicates were removed with Picard MarkDuplicates (Broad Institute, 2019).

In order to estimate the library fragment size we used the correlateReads function of the csaw package (Lun and Smyth, 2016) to estimate the cross-correlation coefficients. We then selected as a fragment size the distance that maximized the loess-smoothed cross-correlation.

Coverage plots were produced by computing the read coverage of annotated ARS (see below) with the featureScores function of the Repitools package (Statham et al., 2010), using a window size of 150 and a smoothing size equal to the fragment length.

To identify enriched binding regions of BrdU-IP-seq outside canonical ARS, we followed the method in Yoshida et al. (Yoshida et al., 2014) but improved the bin size to a single base pair resolution. We then normalized the binned tag densities of IP sample by the densities of the matched input sample. MACS2 with default setting was used for peak calling for the histone mutation samples versus wild-type samples, and p value  $\leq 0.05$  was used to identify confident peaks.

### Co-localization of H3K37R unique peaks on ARS or ACS

To determine how many H3K37R unique peaks are occurring at ARS, we downloaded ARS table from OriDB (Siow et al., 2012) and used the first “status (ARS)” column to generate (1) “Confirmed ARS” and (2) “Dubious/Likely ARS” in Figure 6F. We used nr-ACS list from Eaton et al. (Eaton et al., 2010) for (3) “nr-ACS.” We downloaded ACS genomic location data from OriScan database (Breier et al., 2004) and lift-over the reference genome to SacCer3 for (4) “ACS like.” The (5) “Other” indicate any peaks that falls outside of any regions of 1–4.

### ARS annotation and efficiency estimates

To define ARS efficiency, we averaged the coverage of experimental replicates and then further summarized the coverage of each ARS by averaging all the coverage of all windows within 1kb of the ARS center. We then fitted a Gaussian Mixture Model to find the parameters of two normal distributions, representing inactive and active origins respectively. We defined as “Late/Inefficient” all ARS having a coverage less than 2 standard deviations above the mean of the first normal distribution. The remaining ARS were then equally split into the “Medium” and “Early/Efficient” categories based on their coverage (see Figure S5C).

The annotation of ARS was obtained from OriDB. All ARS with “Confirmed” status in OriDB were consider, with the exception of ARS1216.5 which was excluded due to its overlap with repetitive rDNA and its extreme coverage in the Input samples.

### Analysis of H3K37me1 ChIP-seq

Raw reads were pre-processed and mapped as outlined for the BrdU libraries. To calculate the K37me1 around ARS, BAM files were converted to GRanges (Lawrence et al., 2013) objects and replicates were average using the mergeReplicates function of Repitools (Statham et al., 2010). We then used the featureScores function of Repitools package to calculate the coverage of 150nt windows around ARS, using as a smoothing distance the library fragmented size estimated as explained before.

To test for significant difference at 5min and 30min versus G1, we averaged the coverage of all windows within 2kb of the ARD mid-point, and then applied the Wilcoxon rank sum test on the base 2 logarithm of the coverage ratio between H3K37me1 and H3.

To calculate the coverage at TSSs, we repeated the same procedure as above, but calculating coverage of regions within 2.5kb of SGD annotated TSSs, using a window size of 50 and the same smoothing distance as before.

### Analysis of Mcm2 ChIP-seq

Reads were adaptor-trimmed as outlined before and then mapped with bwa to a combined reference database consisting of both *S. cerevisiae* and *S. pombe* genomes. Uniquely mapping reads were filtered as outlined before and the resulting BAM files were then split by reference chromosomes with the split utility of bamtools. After splitting, BAM files were combined again into two new BAM files containing alignments belonging to *S. cerevisiae* and *S. pombe*, respectively. To calculate Mcm2 coverage around ARS, we first converted the BAM files containing *S. cerevisiae* alignments to GRanges objects, and then used the featureScores function of the Repitools package to calculate the coverage of 50nt windows around ARS, using as a smoothing distance the library fragmented size estimated as explained before. The resulting window-coverage counts were automatically normalized by library size by the featureScores function; therefore, to normalize by the coverage of the *S. pombe* spike-in chromatin the window counts were divided by the total number of reads mapped to *S. pombe* and multiplied by the total number of reads mapped to *S. cerevisiae* (this latter step is required in order to revert the automatic normalization performed by featureScores). Downstream analysis and statistical testing were then done as outlined for the H3K37me1 ChIP-seq experiment.

### Human cell chromatin immunoprecipitation

Exponentially growing cells were plated in 150mm dishes at 30%–40% confluence and incubated at 37°C overnight. Next day, 2mM thymidine was added and cells were incubated for 18h at 37°C. Thymidine was removed by washing with PBS and cells were incubated with fresh medium for 9h at 37°C before a second round of 2mM thymidine block. Cells were then washed with PBS and released into the cycle. Cells were cross-linked with 1% formaldehyde for 10 minutes at 37°C, and the reaction was quenched with 125mM glycine for 5 minutes at room temperature. Cells were washed twice with cold PBS and scraped in cold PBS supplemented with protease inhibitors (Roche). Cells were lysed in in ChIP lysis buffer I (5mM PIPES KOK pH 8; 85mM KCl; 0.5% NP-40) for 10 minutes on ice, spun down and resuspended in ChIP lysis buffer II (1% SDS, 10mM EDTA, 50mM Tris-HCl pH 8.0) supplemented with protease inhibitors (Roche). Chromatin was sonicated (Bioruptor, Diagenode; 20 cycles, 30 s on/off) to yield an average DNA fragment of 300–500 base pairs, and diluted 10 times in ChIP IP buffer (0.01% SDS, 1.1% Triton X-100, 167mM NaCl, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.0) supplemented with protease inhibitors (Roche) prior to overnight immunoprecipitation with rotation at 4°C. Next day, 50  $\mu$ L of protein A Dynabeads were added, and samples were incubated again with rotation at 4°C for 2 hours. Dynabeads were washed twice with the following buffers: TSE 150 (1% Triton X-100, 0.1% SDS, 150mM NaCl 2mM EDTA, 20mM Tris-HCl pH 8.0), TSE 500 (1% Triton X-100, 0.1% SDS, 500mM NaCl, 2mM EDTA, 20mM Tris-HCl pH 8.0) and LiCl buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1mM EDTA, 10mM Tris-HCl pH 8.0). DNA was eluted for 30min at 37°C in elution buffer (100mM NaHCO<sub>3</sub>, 1% SDS), and cross-linking was reverted by overnight incubation at 65°C. Samples were treated with 0.5mg/ml of RNase A at 37°C for 2 h. DNA was purified using the QIAGEN PCR purification kit. Relative DNA amounts were determined by qPCR using SYBR Green Master Mix (Applied Biosystems). Primer pairs used for amplification are available upon request.

### **Direct enzyme-linked immunosorbent assay (ELISA)**

Serial dilutions of biotinylated unmodified and K37me1 peptides were immobilized on streptavidin-coated 96-well plates and the ELISA was performed using a 1:5000 dilution of the anti-H3K37me1 antibody. The ELISA was developed using horseradish peroxidase (HRP)-coupled secondary anti-rabbit antibody and SuperSignal ELISA Pico and quantified in using a ClarioStar plate reader.

### **QUANTIFICATION AND STATISTICAL ANALYSIS**

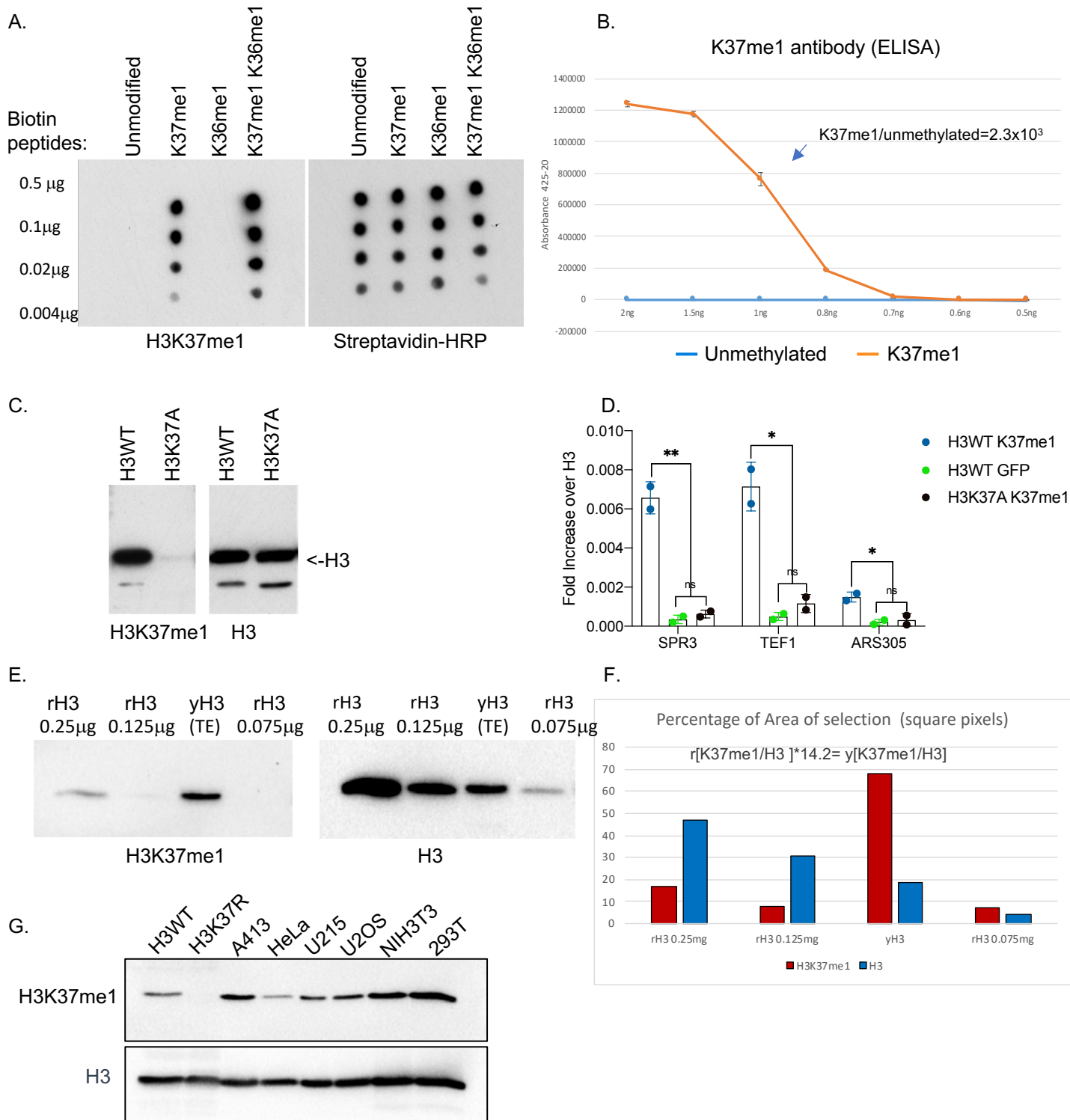
Procedures for quantification and statistical analysis of each individual experiment is described in the corresponding figure legends.



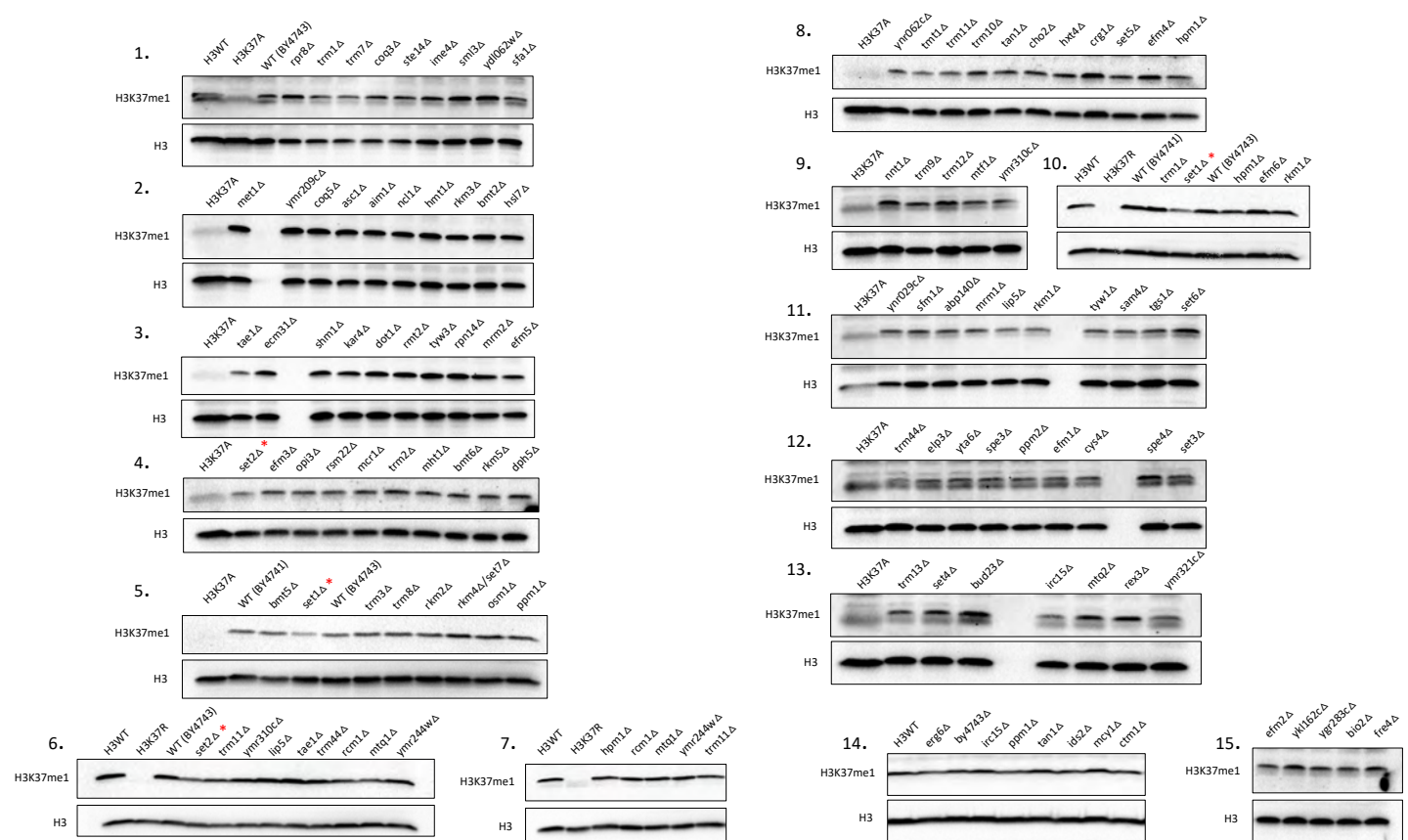
**Supplemental information**

**Methylation of histone H3 at lysine 37 by Set1  
and Set2 prevents spurious DNA replication**

**Helena Santos-Rosa, Gonzalo Millán-Zambrano, Namshik Han, Tommaso Leonardi, Marie Klimontova, Simona Nasiscionyte, Luca Pandolfini, Kostantinos Tzelepis, Till Bartke, and Tony Kouzarides**

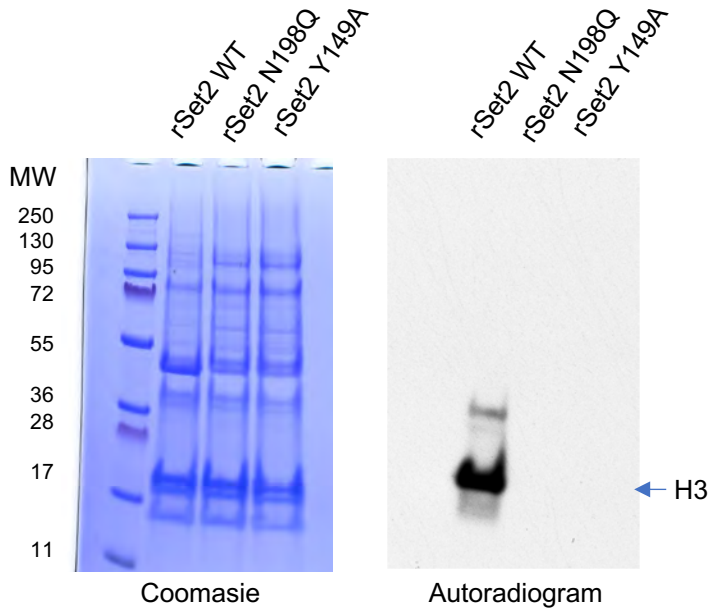


**Figure S1. Characterization of H3K37me1 monoclonal antibody. Related to Figure 1.** Dot-blot analysis using H3K37me1 specific antibody. 5-fold serial dilutions of different biotinylated H3 peptides (aa28-aa48), as indicated, were spotted onto a PVDF membrane. The membrane was subsequently blotted with Streptavidin-HRP for a loading control. **(B)** Direct enzyme-linked immunosorbent assay (ELISA) against the indicated peptides. **(C)** Immunoblot analysis of purified yeast histones. H3WT and isogenic H3K37A mutant cells were grown to exponential phase. Purified yeast histones were separated by SDS-PAGE in 16% acrylamide gels. Blots were probed with anti-H3K37me1 antibody and then re-probed with an anti-H3 antibody as indicated. **(D)** ChIP qPCR experiments showing H3K37me1 levels at different genomic locations. H3WT and H3K37A strains were crosslinked and chromatin was immunoprecipitated (IP) using anti-H3K37me1, anti-H3 or anti-GFP antibodies. Statistical analysis was performed using multiple t test corrected for the comparisons using the Holm-Sidak method (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 2 independent experiments. **(E)** Immunoblot analysis of unmodified recombinant H3 (rH3) versus yeast total extracts (yH3, TE). Blots were probed with anti-H3K37me1 antibody and then re-probed with anti-H3 antibody as indicated. **(F)** Quantification of immunoblots shown in (E). **(G)** Immunoblot analysis of total protein extracts from different mammalian cell lines as indicated. Proteins were separated by SDS-PAGE in 16% acrylamide gels. Blots were probed with anti-H3K37me1 antibody and then re-probed with anti-H3 antibody as indicated. Wild-type H3 and H3K37R mutant were used as controls.

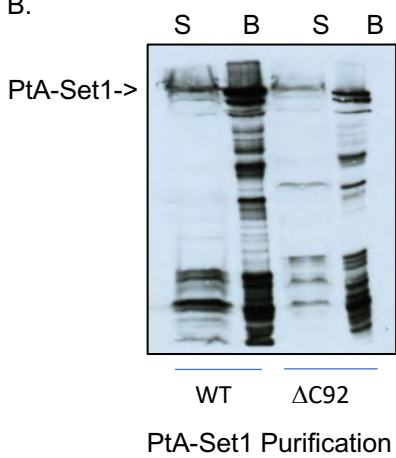


**Figure S2. Set1 (COMPASS) and Set2 are responsible for H3 K37me1 *in vivo*. Related to Figure 1. (A)** Immunoblot analysis of total protein extracts from wild-types (BY4741 and BY4743) and isogenic deletion mutants as specified. Proteins were separated by SDS-PAGE in 16% acrylamide gels. Blots were probed with anti-H3K37me1 antibody and then re-probed with anti-H3 antibody as indicated. H3WT and isogenic H3K37A/R mutants were used as controls. *set1Δ* and *set2Δ* are highlighted \*. **(B)** Immunoblot analysis of unmodified recombinant H3 (rH3) versus yeast wild-type and *set1Δset2Δ* total extracts. Blots were probed with anti-H3K37me1 antibody and then re-probed with anti-H3 antibody as indicated. **(C)** RT-PCR showing mRNA levels of essential methyltransferases in wild-type (W303) and *set1Δset2Δ* isogenic strain. The housekeeping transcript *RTG2* was used for internal normalization. Statistical analysis was performed using multiple t-test corrected for the comparisons using the Holm-Sidak method (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 3 biological replicates.

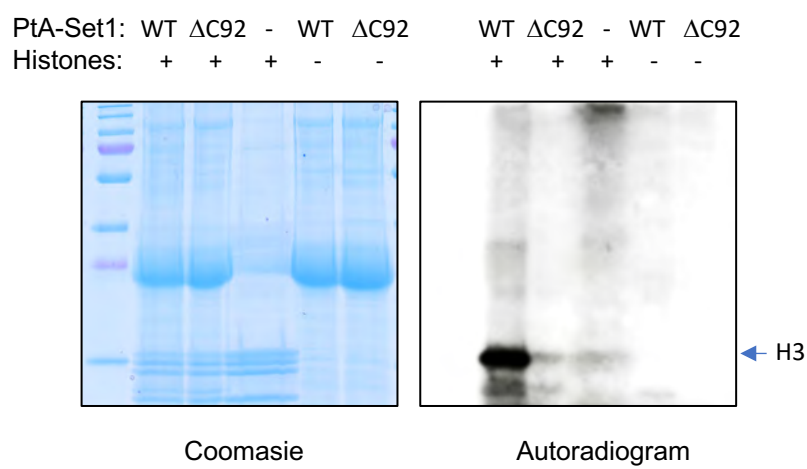
A.



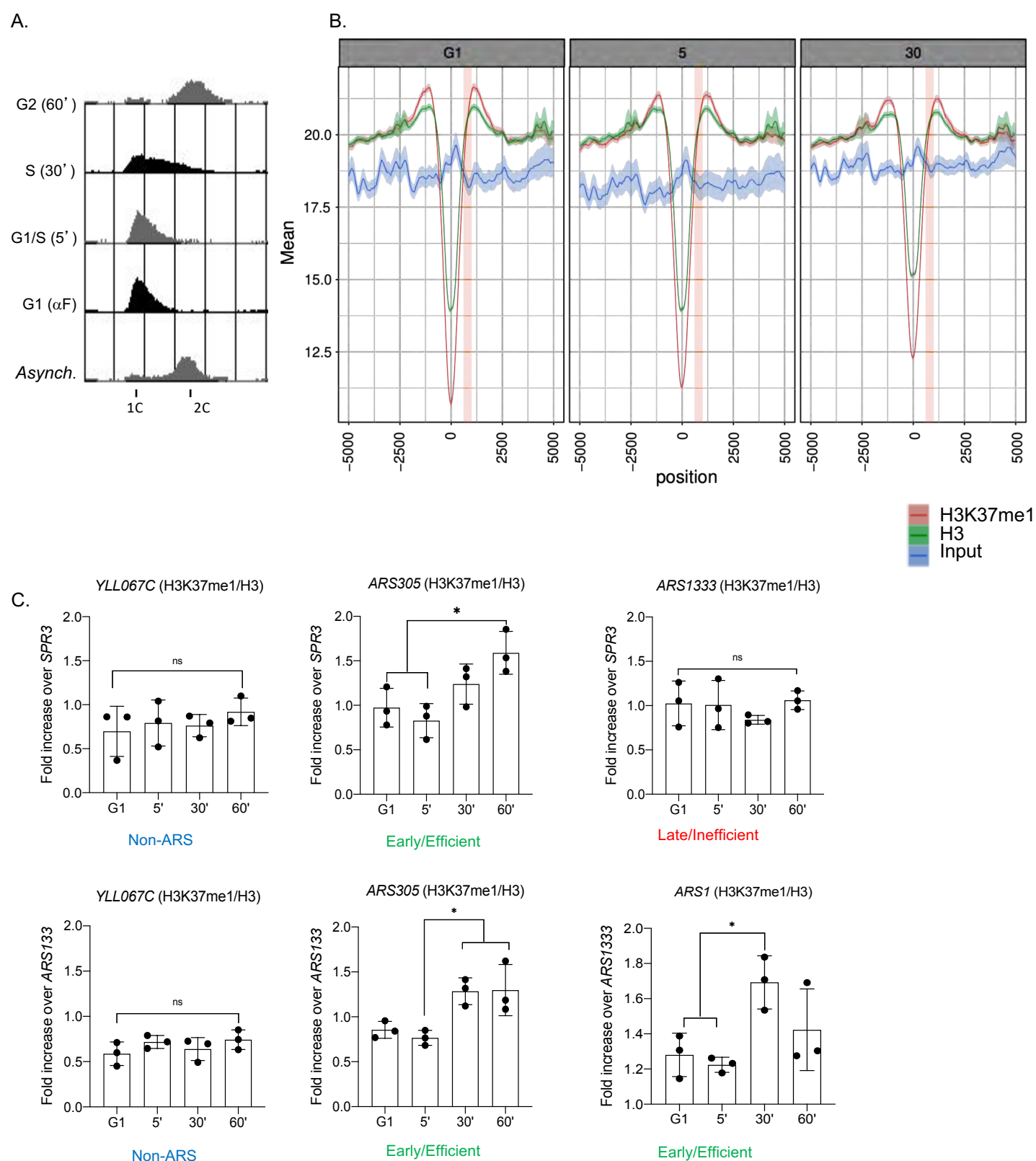
B.



C.



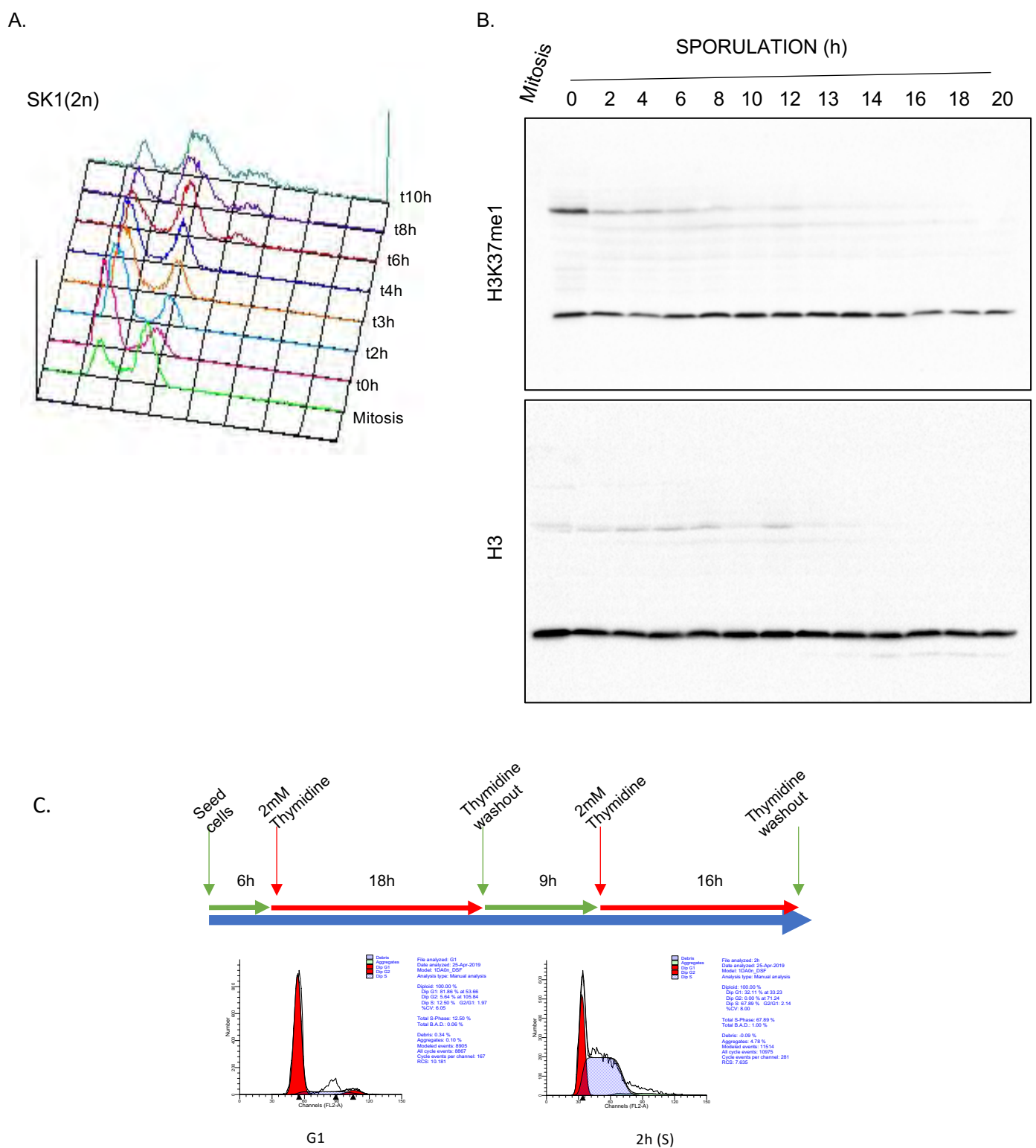
**Figure S3. Set1 (COMPASS) and Set2 *in vitro* activity. Related to Figure 2.** (A)  $^3\text{H}$ -SAM dependent methyltransferase reactions on calf thymus histones, catalysed by recombinant wild-type (WT) Set2p, Set2N198Qp and Set2Y149Ap. Reactions were separated in a NuPAGE 4-12% gel. Left panel: Coomassie Blue staining. Right panel: Autoradiogram. H3 is highlighted. (B) IgG chromatography purification of WT PtA-Set1p and PtA-set1ΔC92p. Complexes were resolved by SDS-PAGE in 16% acrylamide gels. Set1p and set1ΔC92p were detected by immunoblot using anti-PAP antibody. S: yeast soluble input, B: beads bound fraction. (C)  $^3\text{H}$ -SAM dependent methyltransferase reactions on calf thymus histones, catalysed by equal amounts of wild-type PtA-Set1p and PtA-set1ΔC92p yeast purified complexes. Reactions were separated in NuPAGE 4-12% gel. Left panel: Coomassie Blue staining. Right panel: Autoradiogram. H3 is highlighted.



**Figure S4. H3 K37me1 is cell cycle regulated at replication origins.**

**Related to Figure 3. (A)** Flow cytometry analysis. Wild-type cells were arrested in G1 using  $\alpha$ -factor, released into the cell cycle and samples were collected at the specified times. **(B)** Coverage plot showing the mean normalised ChIP-signal ( $\pm$  s.e.m.) for H3, H3K37me1 and Input at Transcription Start Sites  $\pm$  5kb. The time points (G1, 5' and 30') correspond to those in Fig.3B. **(C)** ChIP qPCR experiments showing H3K37me1 levels at different genomic locations. Wild-type cells were arrested in G1 using  $\alpha$ -factor and released into the cell cycle. Chromatin from G1, 5' 30' and 60' time points was immunoprecipitated using anti-H3K37me1 and anti-H3 antibodies. H3K37me1 enrichment over H3 was normalized to a non-ARS region (*SPR3*) and to a late/inefficient ARS (*ARS1333*). Statistical analysis was performed using One-way ANOVA multiple comparisons using Tukey's multiple comparison test (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 3 independent experiments

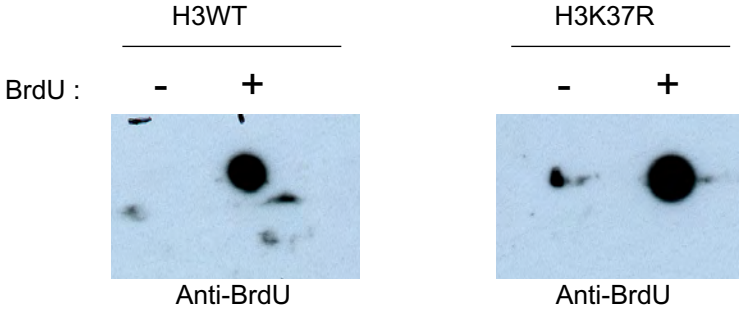




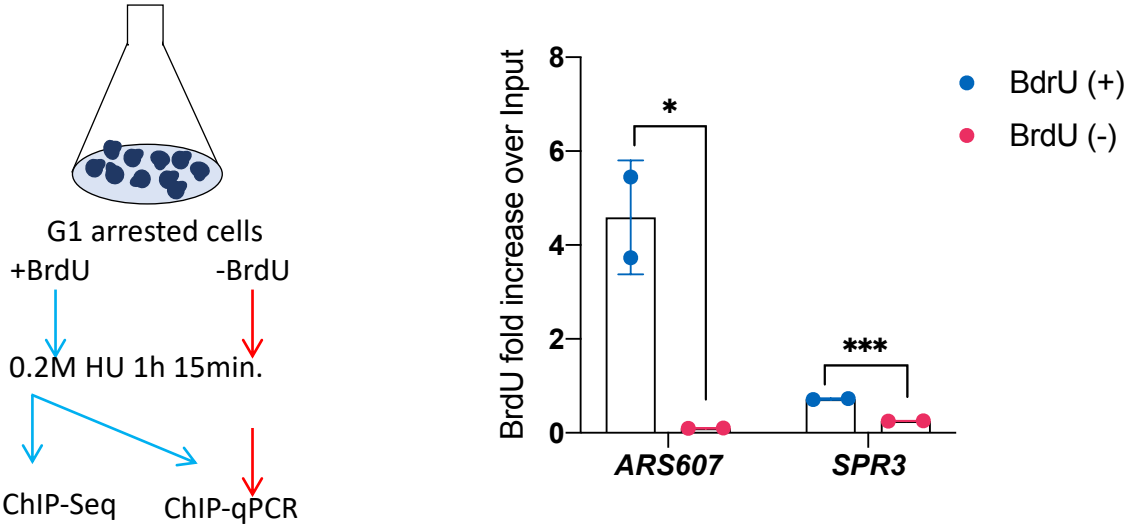
**Figure S5. H3 K37me1 increases during premeiotic DNA replication. Related to Figure 3.** SK1 diploid cells were G1 enriched by nutrient starvation and released to sporulation medium. Samples for FACS and protein total extracts were collected at the specified times. **(A)** Flow cytometry analysis. **(B)** Proteins were separated by SDS-PAGE in 16% acrylamide gels and probed with anti-H3K37me1 antibody and then re-probed with an anti-H3 antibody as indicated. **(C)** Experimental design for double thymidine arrest and flow cytometry analysis of RPE1 cell cultures used for G1→S H3K37me1 chromatin immunoprecipitation experiments shown in Figure 3F.

A.

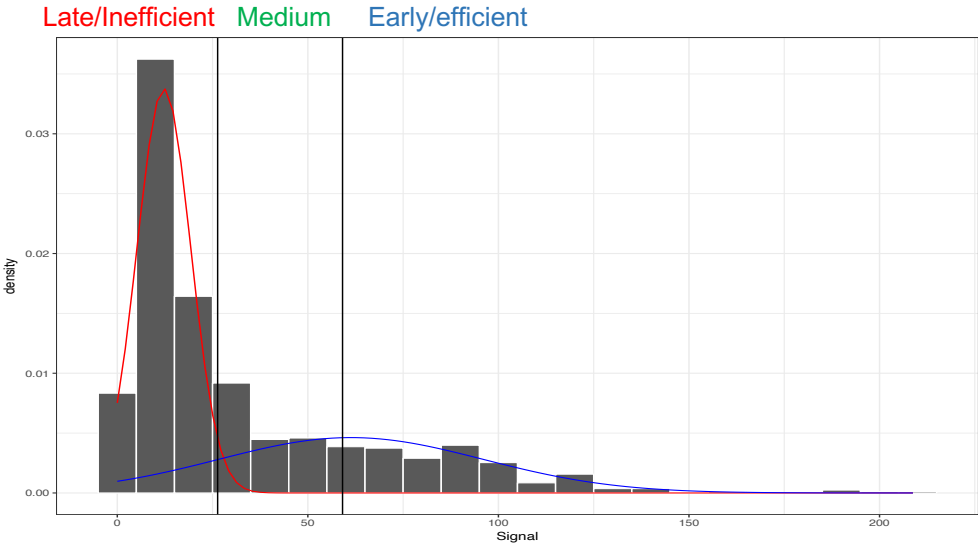
Thymidine Kinase (TK) + Nucleoside transporter (hENT1)



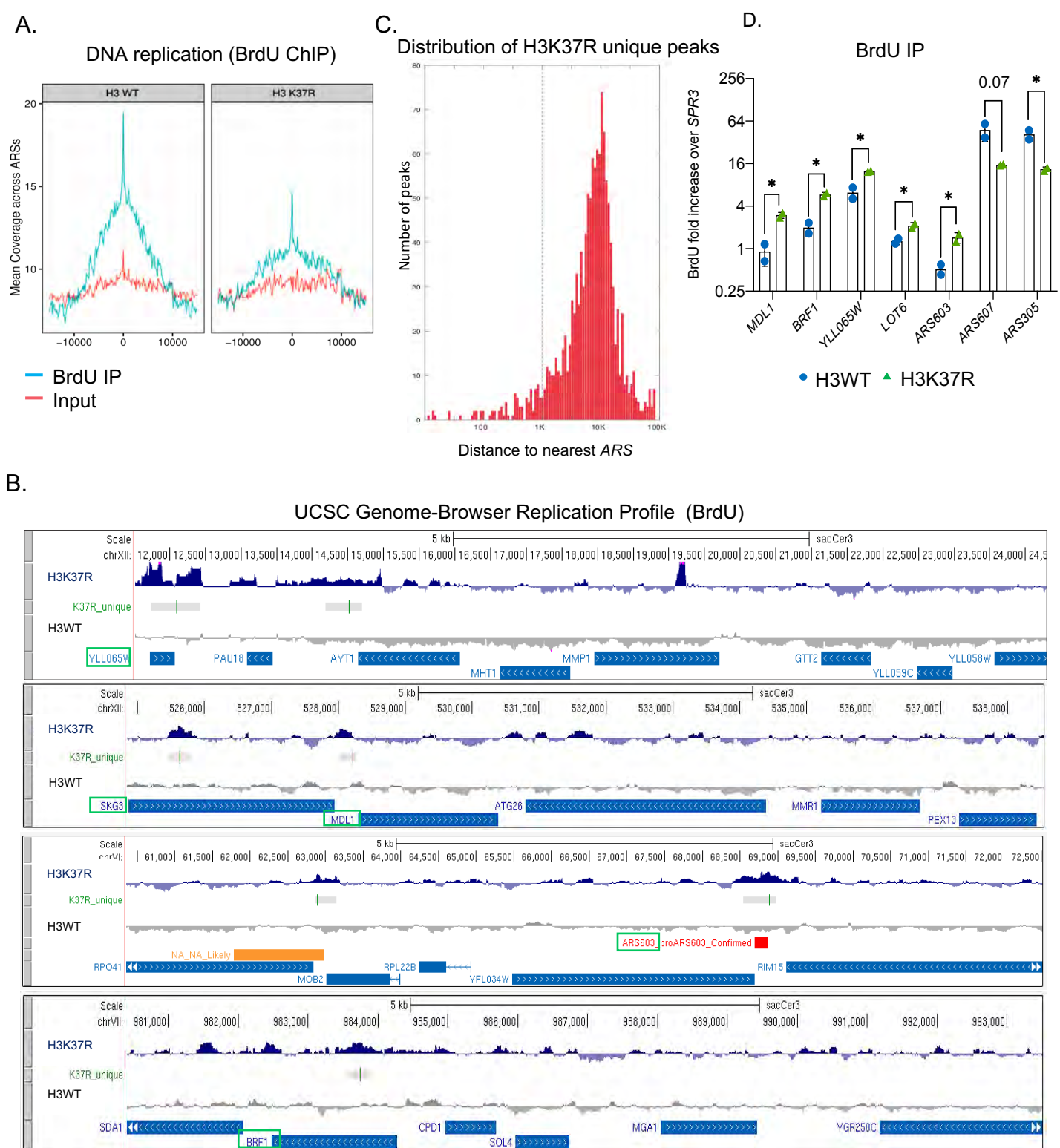
B.



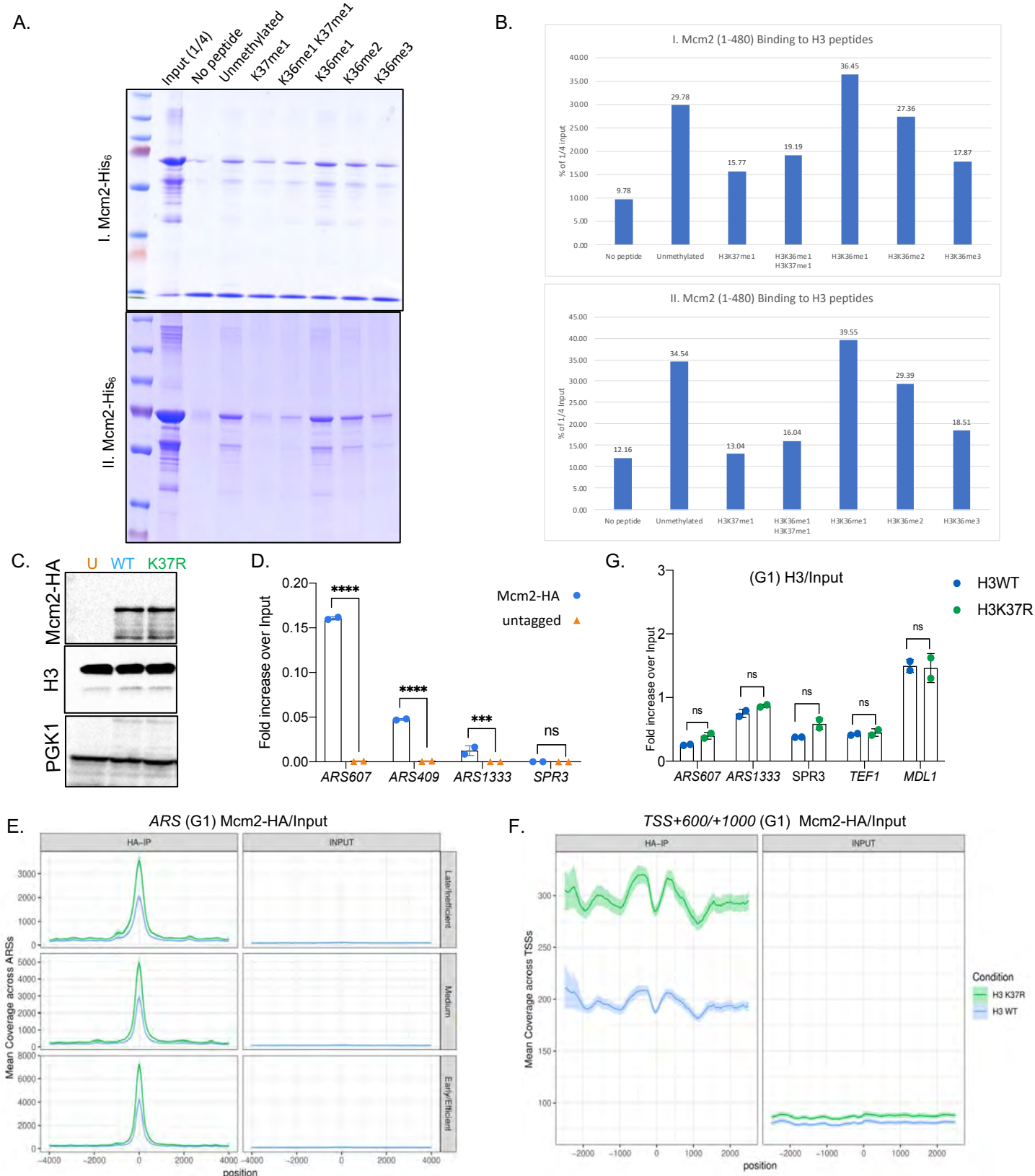
C.



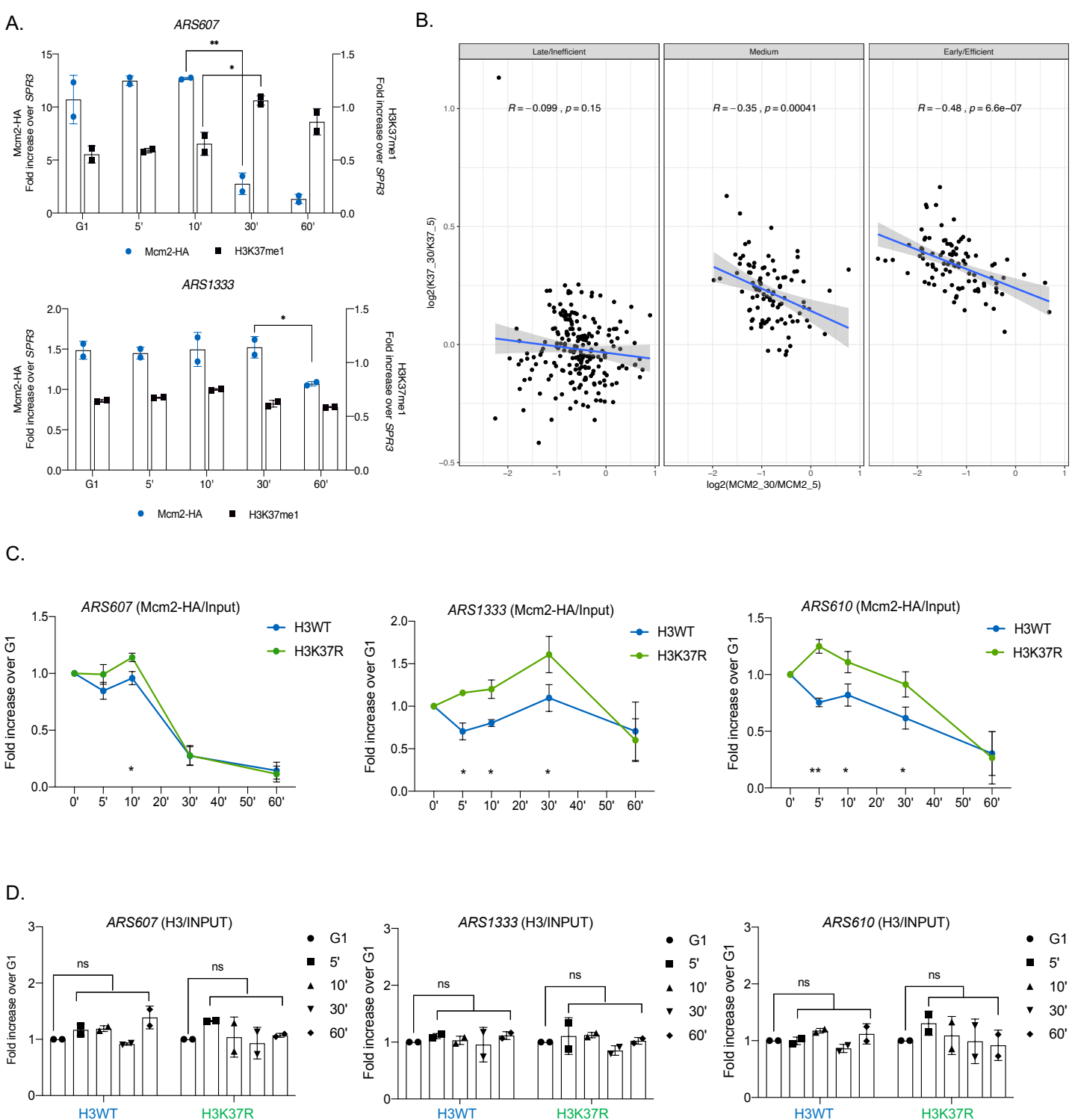
**Figure S6. DNA replication profiling of H3WT and H3K37R mutant. Related to Figure 4.** (A) BrdU incorporation of H3WT and isogenic H3K37R mutant strains. Equal amounts of DNA from H3WT and H3K37R cells incubated in the absence (-) or presence (+) of BrdU was immunoblotted with anti-BrdU antibody. (B) **Left:** Experimental design for DNA replication profiling experiments. **Right:** Equal amounts of DNA from H3WT cells incubated in the absence (-) or presence (+) of BrdU was immunoprecipitated using anti-BrdU antibody. IP material was quantified by qPCR using specific primers as indicated. Statistical analysis was performed using multiple t-test corrected for the comparisons using the Holm-Sidak method (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 2 independent experiments. (C) Distribution of BrdU IP signal intensity at replication origins. y-axis shows the probability densities, indicating the frequencies of ARSs observed at different BrdU intensity levels (x-axis: Signal), as estimated by the kernel density function. Red and Blue lines represent the components of a Gaussian Mixture Model corresponding, respectively, to non-active and active ARS in a H3 wild-type strain under the experimental conditions described in (B).



**Figure S7. Lack of H3K37me1 results in genome-wide spurious replication events. Related to Figure 4. (A)** Coverage plot of the mean signal across all *ARS* shown in (Fig.4A) centred at ACS (*ARS* Consensus Sequence). Input (red); IP (blue). **(B)** Representative genome browser snapshots of BrdU incorporation in H3WT and isogenic H3K37R mutant cells at different chromosomal locations. H3K37R statistically significant unique peaks are highlighted (green). Locations chosen for further analysis are labelled by green boxes. The plots represent the average of 4 independent experiments. **(C)** Distribution of H3K37R unique peaks relative to the nearest confirmed *ARS*. Note the logarithmic scale on *x-axis*. **(D)** BrdU incorporation in H3WT and H3K37R mutant cells analysed by qPCRs at different genomic locations with specific primers. The data correspond to the experiment shown in Fig.4C normalized to a non-*ARS* region (*SPR3*) instead to “IP fold increase over Input”. Error bars represent the mean  $\pm$  SD of 2 biological replicates. Statistical analysis was performed using multiple t test without correction for multiple comparisons (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ , \*\*\* -  $P \leq 0.001$ , \*\*\*\* -  $P \leq 0.0001$ . Note the logarithmic scale on *y-axis*.



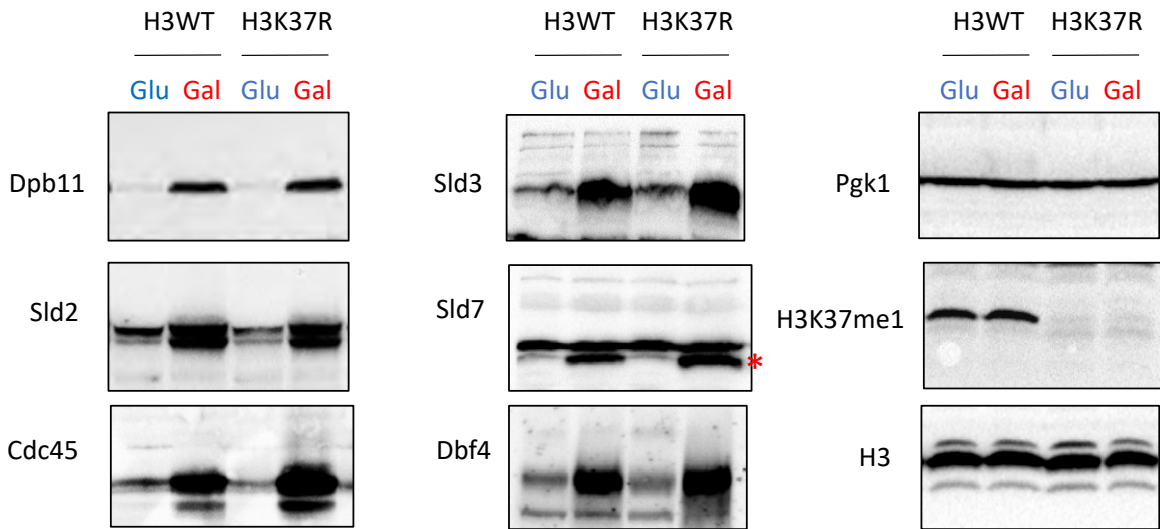
**Figure S8. Lack of H3K37me1 leads to increased MCM recruitment to chromatin. Related to Figure 5. (A)** Mcm2-His<sub>6</sub> *in vitro* binding to biotinylated H3 peptides. Input and peptide-bound Mcm2 were resolved by SDS-PAGE in 10% (assay I) or 8% (assay II) acrylamide gels and stained with Coomassie. I. (shown in Fig.5A) and II are independent binding assays. **(B)** Quantification of assays I and II by ImageJ. Binding is represented as % of the signal corresponding to 1/4 of the input. **(C)** Immunoblot of protein extract from G1-arrested untagged (U) and Mcm2-HA tagged (WT and H3K37R) cells. Blots were probed with anti-HA or anti-H3 antibodies and re-probed with an anti-PGK1 antibody as control. **(D)** Mcm2-HA ChIP in Mcm2-HA (blue) or untagged Mcm2 (orange) cells. IPs were analysed by qPCRs at different genomic locations with specific primers. Statistical analysis was performed using Two-way ANOVA corrected for the comparisons using the Holm-Sidak method (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ , \*\*\* -  $P \leq 0.001$ , \*\*\*\* -  $P \leq 0.0001$ . Error bars represent the mean  $\pm$  SD of 2 independent experiments **(E)** Coverage plots of Mcm2-HA occupancy at ARSs centred at ACS **(F)** Coverage plots of Mcm2-HA occupancy at non-ARS regions centred at TSS. Data correspond to the average of biological triplicates. Note the difference in the scales. **(G)** H3 occupancy in G1 arrested H3WT and H3K37R cells. IPs were analyzed by qPCR with primers specific for each location. Statistical analysis, using Two-way ANOVA, was corrected for the comparisons using the Holm-Sidak method. “ns” refers to non-statistical significance. Error bars represent the mean  $\pm$  SD of 2 biological replicates.



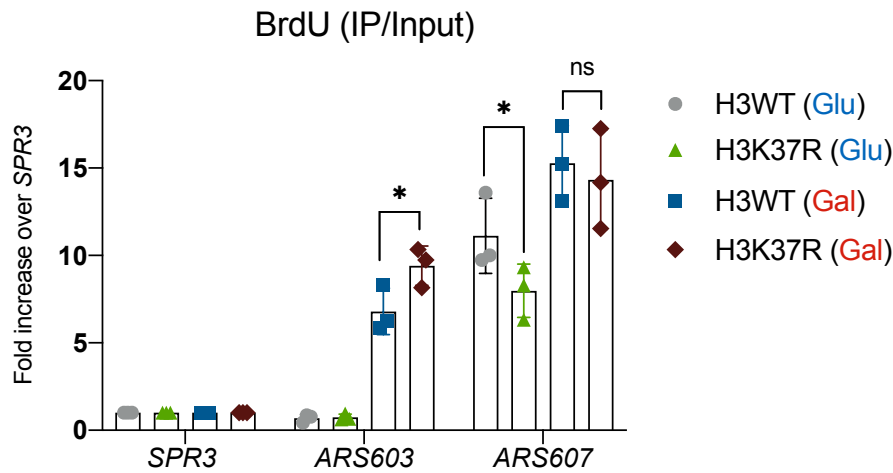
**Figure S9. H3K37me1 and MCM anticorrelate on chromatin during the cell cycle. Related to Figure 5. (A)** Time course ChIP qPCR experiments showing Mcm2-HA and H3K37me1 levels at *ARS*s. Wild-type cells were arrested in G1, released into the cycle and chromatin samples from indicated times were immunoprecipitated using anti-HA, anti-H3K37me1 and anti-H3 antibodies. The IP material was analyzed by qPCR with primers towards efficient *ARS607* or inefficient *ARS1333* and normalized to a non-*ARS* region (*SPR3*). Statistical analysis was performed using multiple t test without correction for multiple comparisons (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 2 independent experiments. **(B)** Scatter plot ( $\log_2$  30'/5' released from G1) representing Mcm2 (*x*-axis) versus H3K37me1 (*y*-axis) at early/efficient, medium and late/inefficient *ARS*s. **(C)** ChIP qPCR experiments showing Mcm2-HA during a time course in wild-type and H3K37R mutant. Cells were arrested in G1, released into the cycle and chromatin samples from indicated times were immunoprecipitated using anti-HA antibody. The IP material was analyzed by qPCR with primers specific towards late/inefficient (*ARS1333* and *ARS610*) and early/efficient (*ARS607*) replication origins. Time points are normalized to the G1 signal. Statistical analysis was performed using multiple t test without correction for multiple comparisons (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 2 independent experiments. **(D)** ChIP qPCR experiments showing H3 levels during the time course shown in (C). The IP material was analyzed by qPCR with primers specific towards late/inefficient (*ARS1333* and *ARS610*) and early/efficient (*ARS607*) replication origins. Time points are normalized to the G1 signal. Statistical analysis was performed using multiple t test without correction for multiple comparisons. "ns" refers to non-statistical significance ( $P > 0.05$ ). Error bars represent the mean  $\pm$  SD of 2 independent experiments.



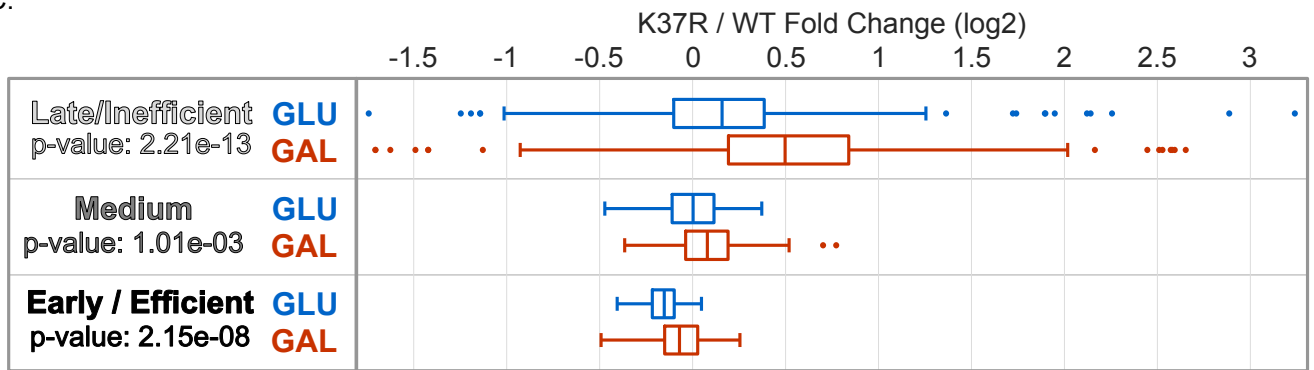
A.



B.



C.



**Figure S10. Overexpression of the MCM activators suppresses the H3H37R replication defect. Related to Figure 6. (A)** Immunoblot analysis of total H3WT and isogenic H3K37R mutant extracts prepared under “non-induced” (Glucose) and “induced overexpression” (Galactose) of the MCM helicase activators as described in “Method details”. Blots were probed with antibodies specific to each of the activators (generous gift from H. Araki) and anti-H3K37me1 antibodies as indicated. Anti-Pgl1 and anti-H3 were used as controls. Sld7 specific signal is highlighted \*. **(B)** H3WT and isogenic H3K37R DNA replication profiles. Equal amounts of DNA from H3WT and H3K37R cells grown in the presence of BrdU under “non-induced” (Glucose) and “induced overexpression” (Galactose) of the MCM helicase activators, were immunoprecipitated using anti-BrdU antibody. IP material was quantified by qPCR using primers specific towards late/inefficient *ARS603* or early/efficient *ARS607*, then normalized to a non-*ARS* region (*SPR3*). Statistical analysis was performed using multiple t test without correction for multiple comparisons (Alpha: 0.05); \* -  $P \leq 0.05$ , \*\* -  $P \leq 0.01$ . Error bars represent the mean  $\pm$  SD of 3 independent colonies. **(C)** Box-plot showing the distribution of mean BrdU signal (H3K37R/H3WT ratio) incorporated at the replication origins shown in B. The p-values were calculated with the Mann-Whitney-Wilcoxon test.