

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

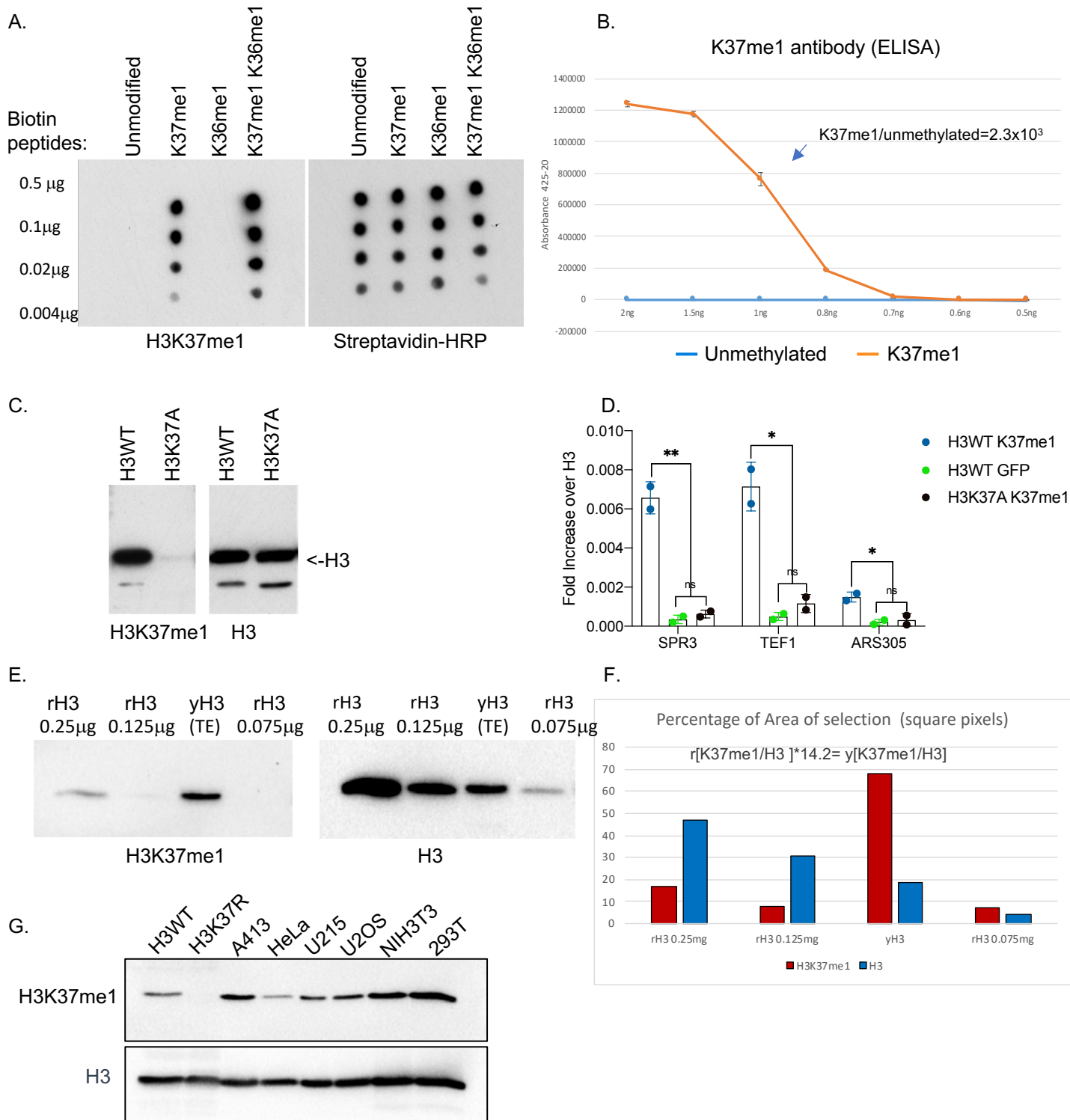


Figure S1. Characterization of H3K37me1 monoclonal antibody. Related to Figure 1. Dot-blot analysis using H3K37me1 specific antibody. 5-fold serial dilutions of different biotinylated H3 peptides (aa28-aa48), as indicated, were spotted onto a PVDF membrane. The membrane was subsequently blotted with 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.

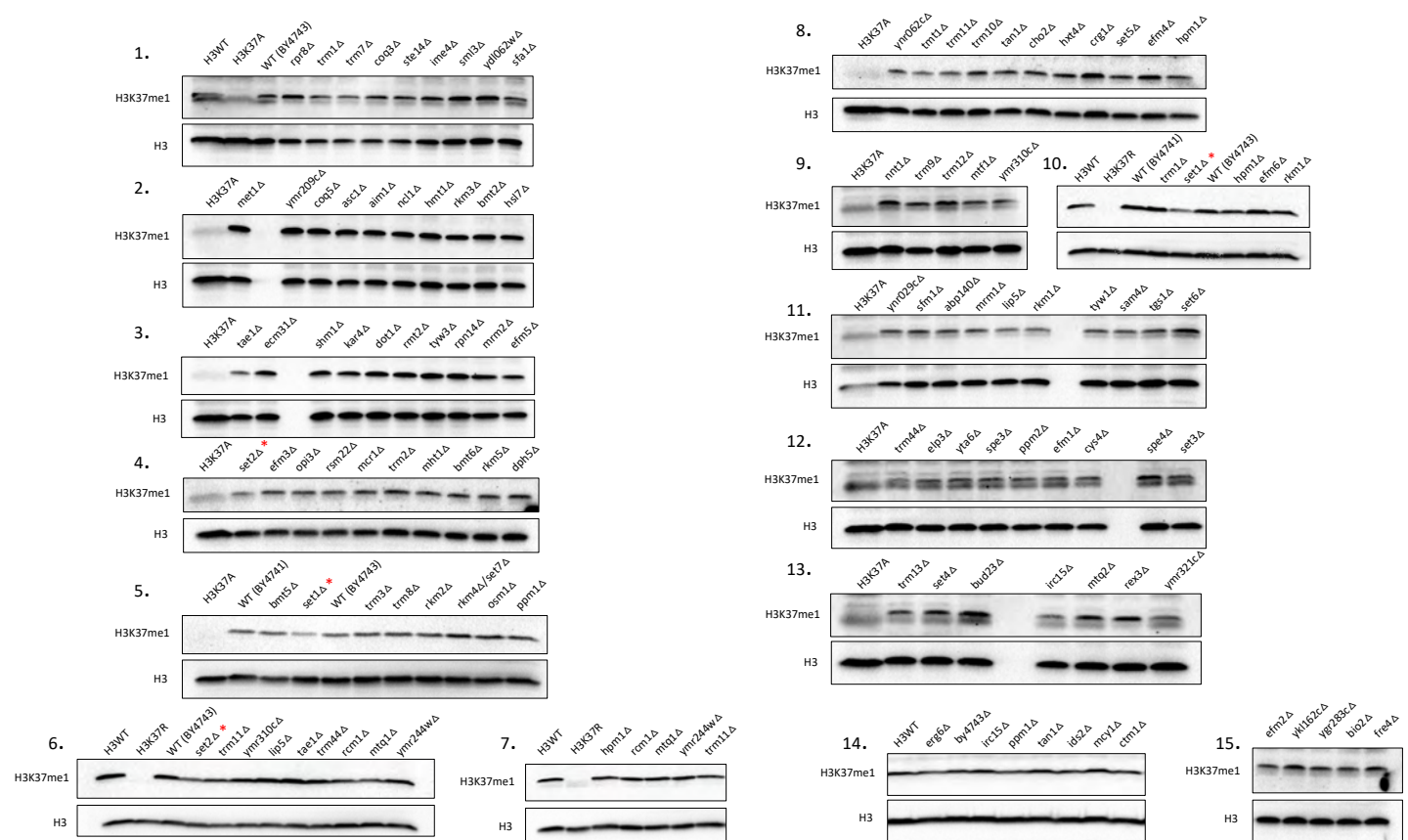


Figure S2. Set1 (COMPASS) and Set2 are responsible for H3 K37me1 *in vivo*. Related to Figure 1. (A) Immunoblot analysis of total protein extracts from wild-types (BY4741 and BY4743) and isogenic deletion mutants as specified. Proteins were separated by SDS-PAGE in 16% acrylamide gels. Blots were probed 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.

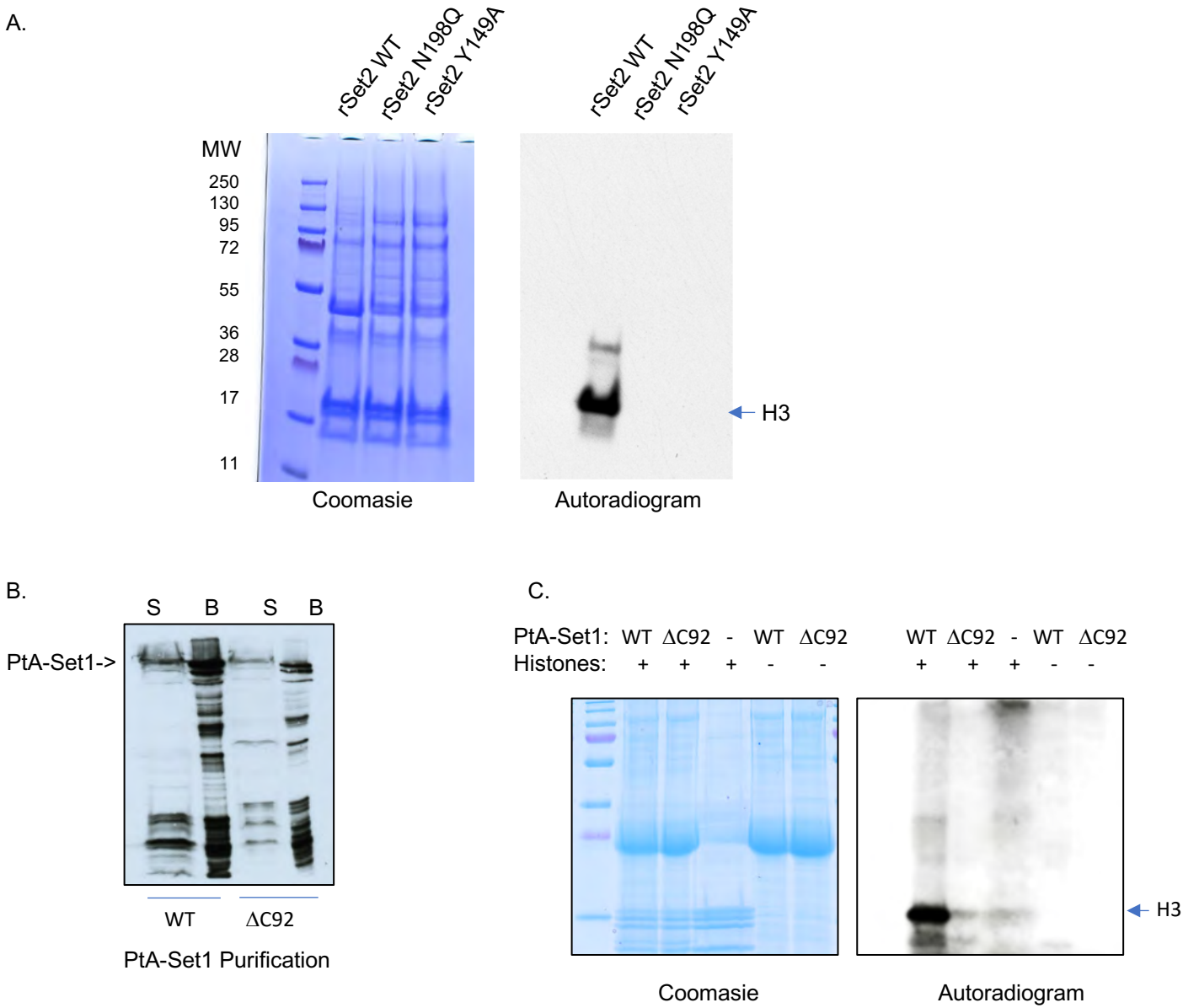


Figure S3. Set1 (COMPASS) and Set2 *in vitro* activity. Related to Figure 2. (A) ^3H -SAM dependent methyltransferase reactions on calf thymus histones, catalysed by recombinant wild-type (WT) Set2p, Set2N198Qp and Set2Y149Ap. Reactions were separated in a NuPAGE 4-12% gel. Left panel: Coomassie Blue 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) ^3H -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.

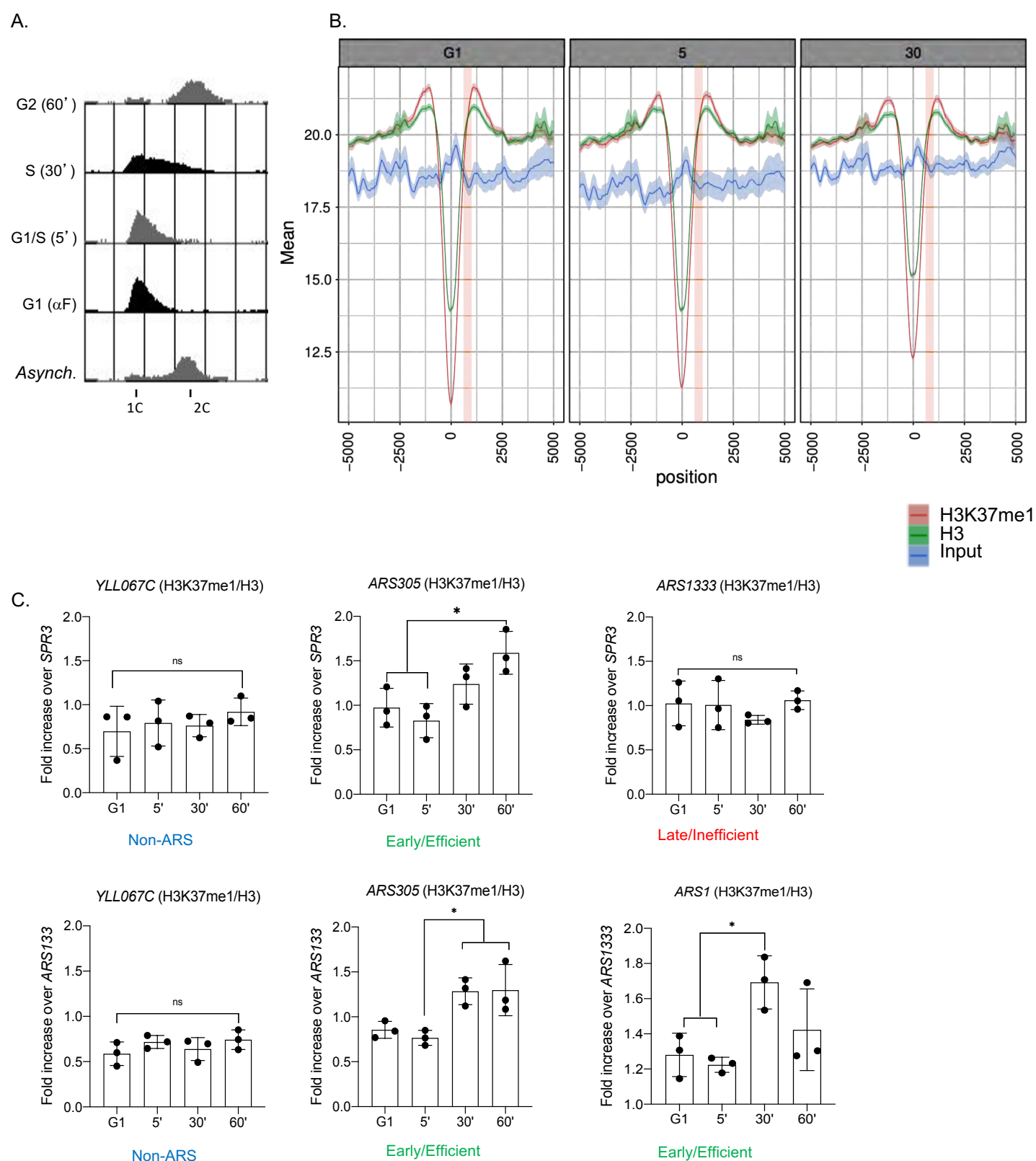
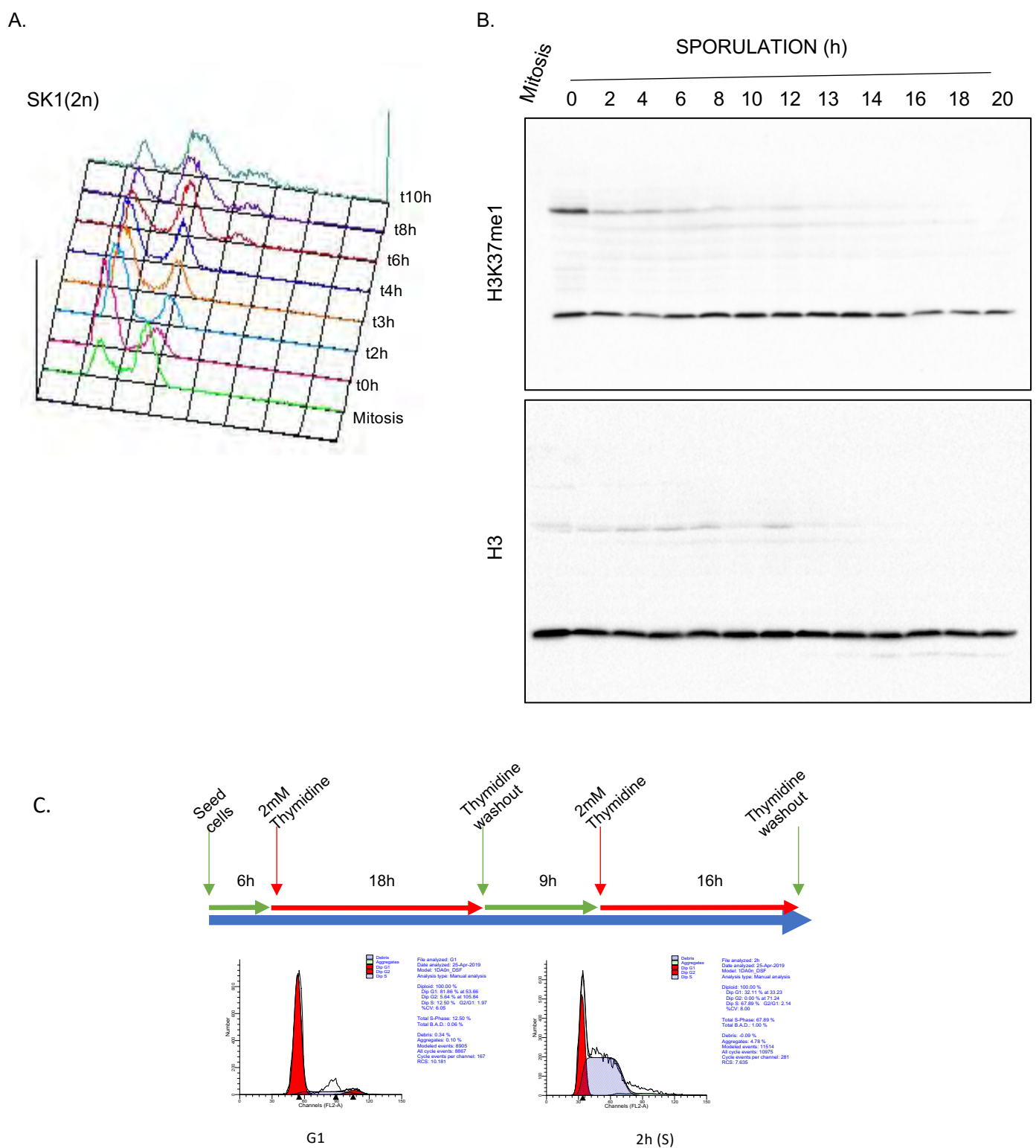


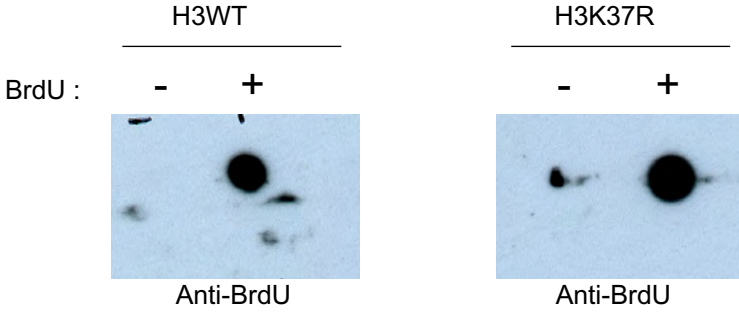
Figure S4. H3 K37me1 is cell cycle regulated at replication origins.

Related to Figure 3. (A) Flow cytometry analysis. Wild-type cells were arrested in G1 using α -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 α -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

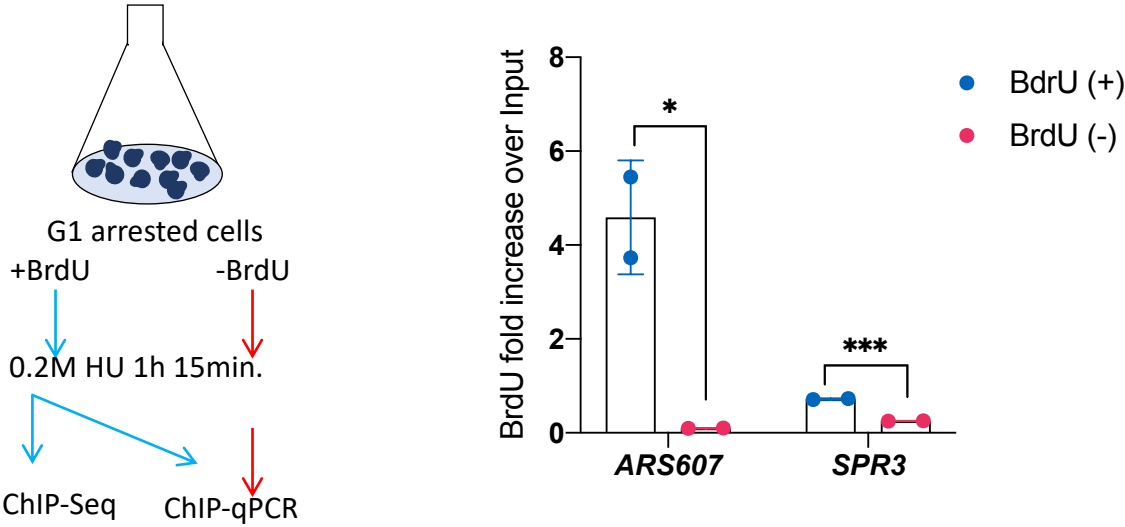


A.

Thymidine Kinase (TK) + Nucleoside transporter (hENT1)



B.



C.

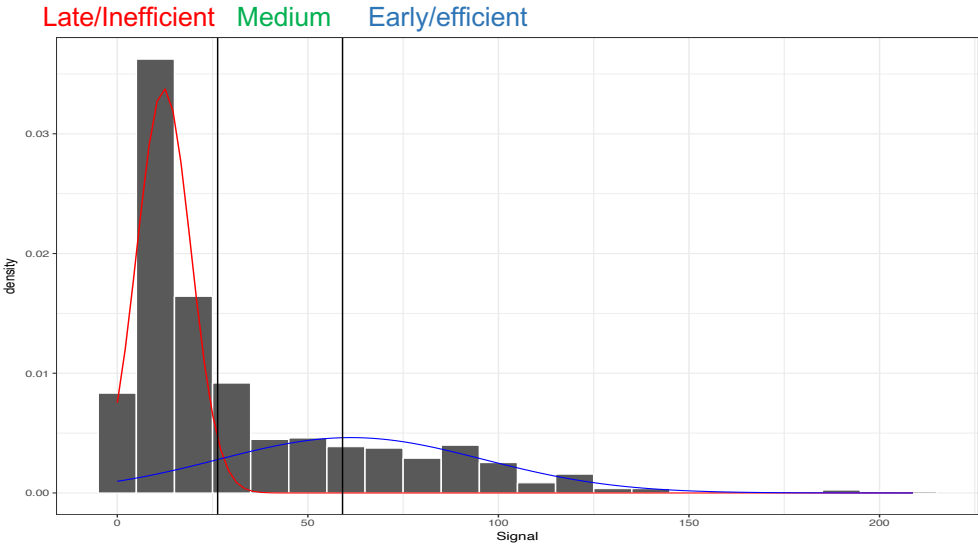


Figure S6. DNA replication profiling of H3WT and H3K37R mutant. Related to Figure 4. (A) BrdU incorporation of H3WT and isogenic H3K37R mutant strains. Equal amounts of DNA from H3WT and H3K37R cells incubated in the absence (-) or presence (+) of BrdU was immunoblotted with anti-BrdU antibody. (B) **Left:** 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).

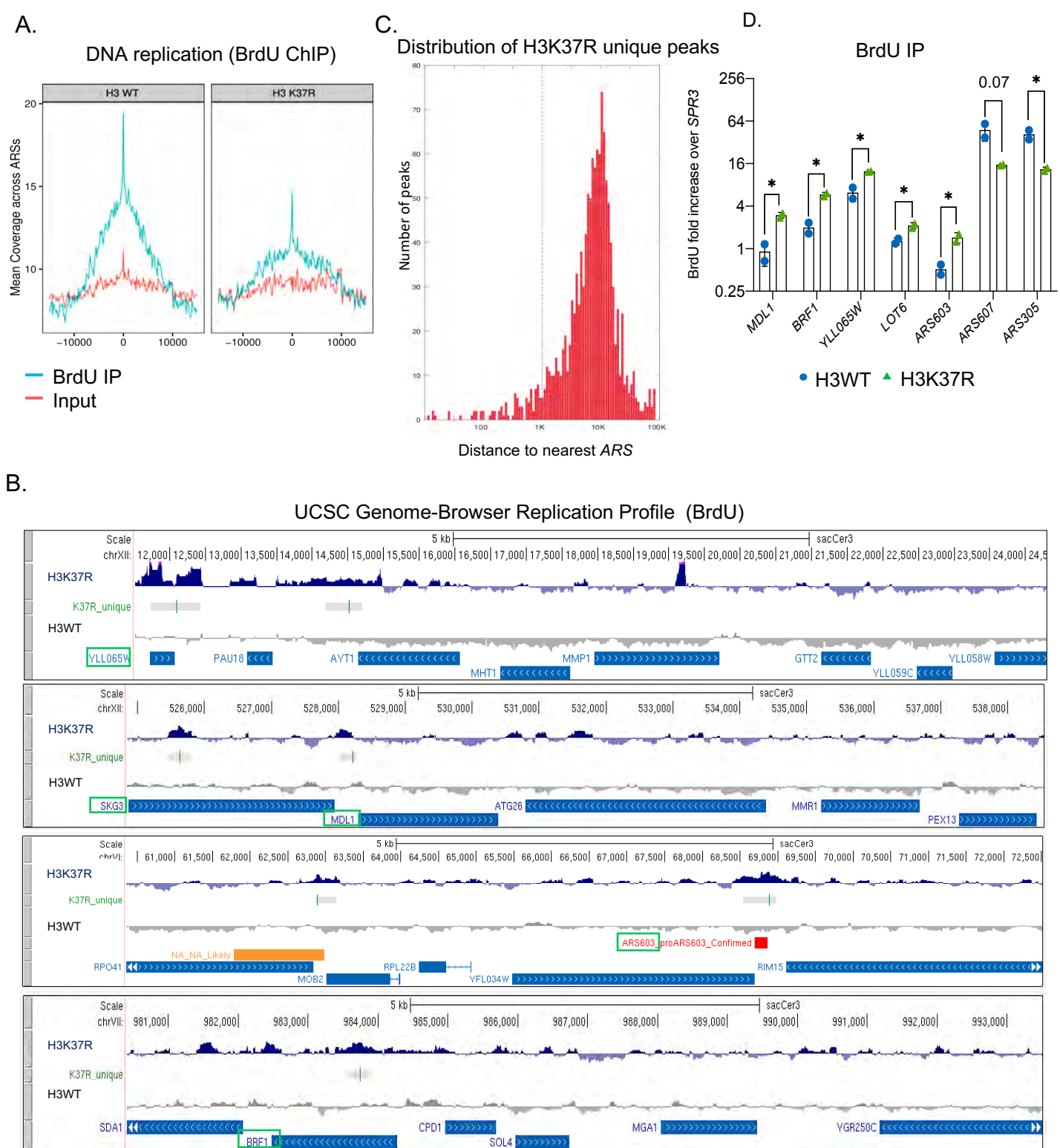


Figure S7. Lack of H3K37me1 results in genome-wide spurious replication events. Related to Figure 4. (A) Coverage plot of the mean signal across all *ARS* shown in (Fig.4A) centred at ACS (*ARS* Consensus Sequence). Input (red); IP (blue). **(B)** Representative genome browser snapshots of BrdU incorporation 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.

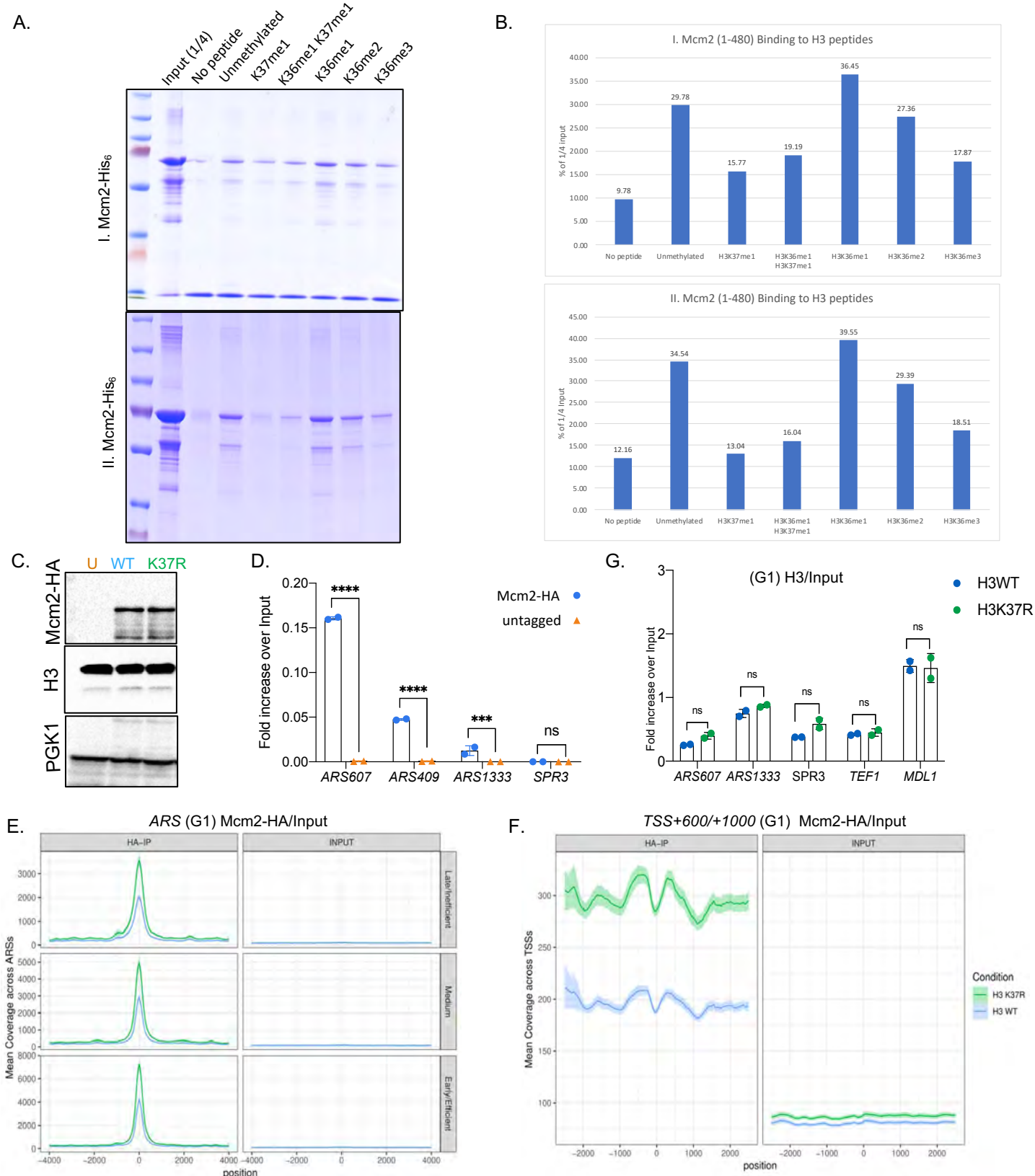


Figure S8. Lack of H3K37me1 leads to increased MCM recruitment to chromatin. Related to Figure 5. (A) Mcm2-His₆ *in vitro* binding to biotinylated H3 peptides. Input and peptide-bound Mcm2 were resolved by SDS-PAGE in 10% (assay I) or 8% (assay II) acrylamide gels and stained with Coomassie. I. (shown 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.

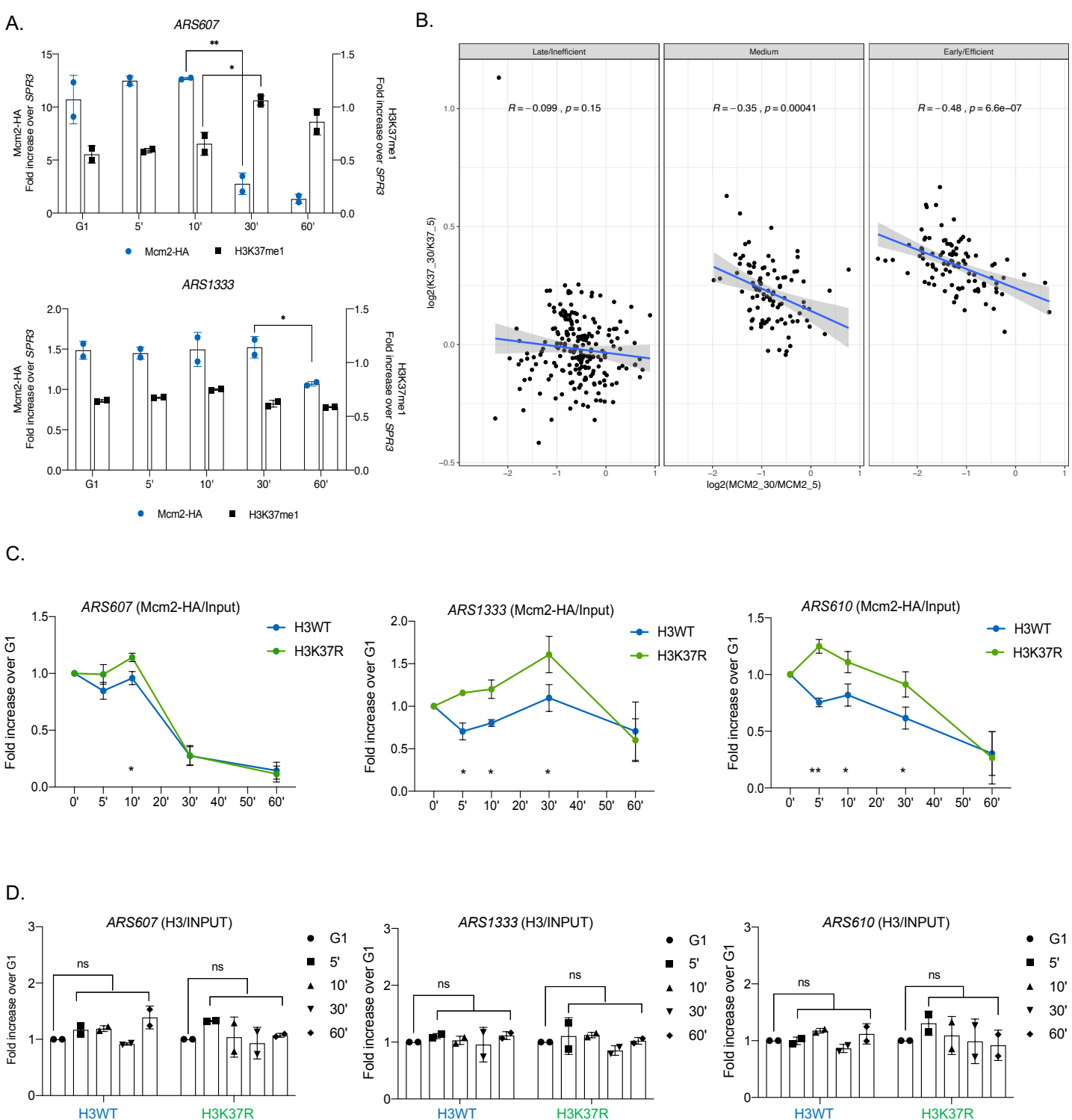
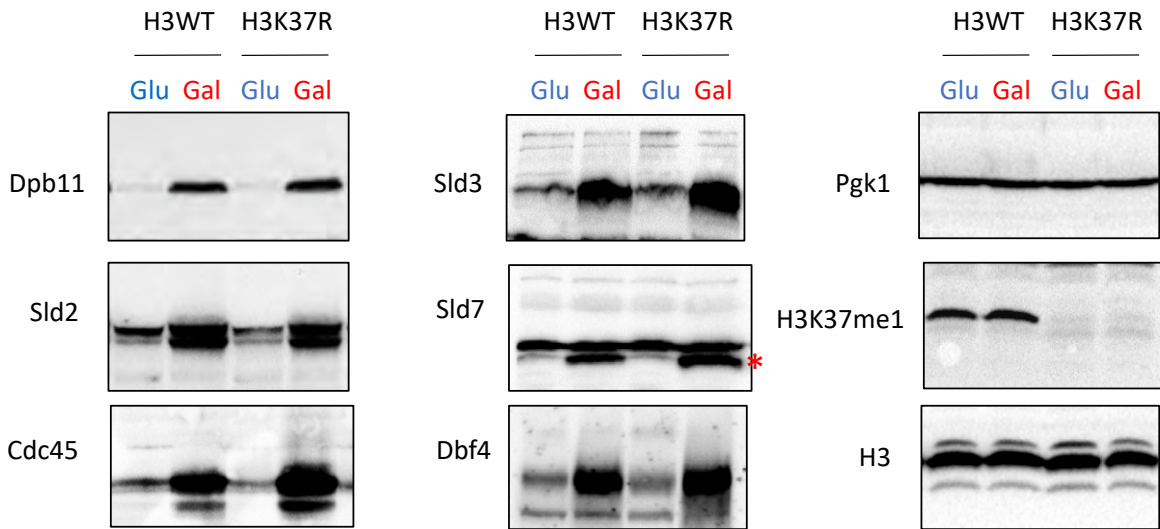
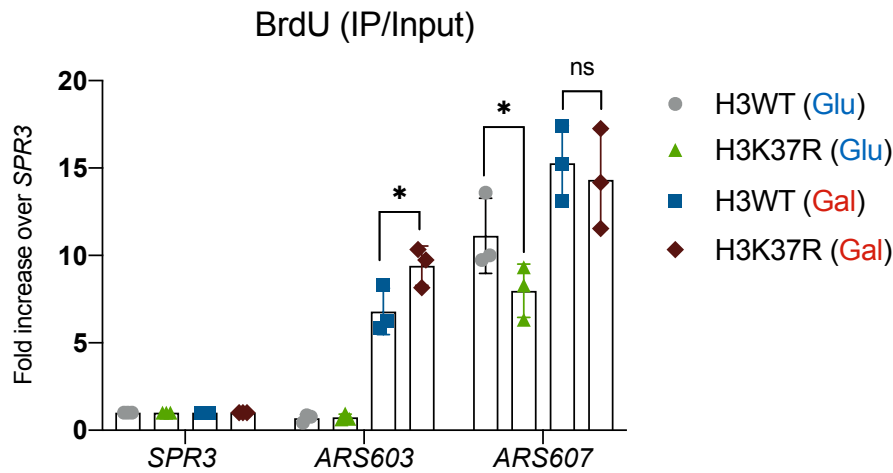


Figure S9. H3K37me1 and MCM anticorrelate on chromatin during the cell cycle. Related to Figure 5. (A) Time course ChIP qPCR experiments showing Mcm2-HA and H3K37me1 levels at *ARS*s. Wild-type cells were arrested in G1, released into the cycle and chromatin samples from indicated times were immunoprecipitated 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.

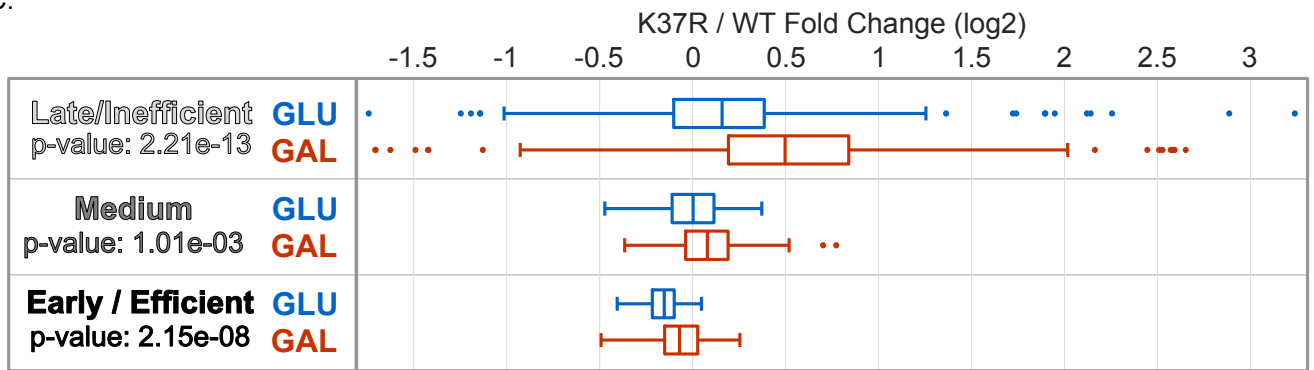


Figure S10. Overexpression of the MCM activators suppresses the H3H37R replication defect. Related to Figure 6. (A) Immunoblot analysis of total H3WT and isogenic H3K37R mutant extracts prepared under “non-induced” (Glucose) and “induced overexpression” (Galactose) of the MCM helicase activators as described 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-Pgk1 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.