Published in final edited form as:

Cell Rep. 2021 October 05; 37(1): 109799. doi:10.1016/j.celrep.2021.109799.

Collective regulation of chromatin modifications predicts replication timing during cell cycle

Capucine Van Rechem[1,2,8], **Fei Ji**[3,8], **Damayanti Chakraborty**[1,8], **Joshua C. Black**[1,7], **Ruslan I. Sadreyev**[3,4,*], **Johnathan R. Whetstine**[1,5,6,9,*]

[1]Massachusetts General Hospital Cancer Center and Harvard Medical School, Charlestown, MA 02129, USA

[2]Department of Pathology, Stanford Medicine, Stanford, CA 94305, USA

[3]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

[4]Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

[5]Cancer Signaling and Epigenetics Program, Fox Chase Cancer Center, Philadelphia, PA 19111, USA

[6]Cancer Epigenetics Institute, Fox Chase Cancer Center, Philadelphia, PA 19111, USA

[7]Present address: Department of Pharmacology, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA

[8]These authors contributed equally

[9]Lead contact

SUMMARY

Replication timing (RT) associates with genome architecture, while having a mixed relationship to histone marks. By profiling replication at high resolution and assessing broad histone marks across the cell cycle at the resolution of RT with and without genetic perturbation, we address the causal relationship between histone marks and RT. Four primary chromatin states, including an uncharacterized H3K36me2 state, emerge and define 97% of the mappable genome. RT and local replication patterns (e.g., initiation zones) quantitatively associate with chromatin states, histone mark dynamics, and spatial chromatin structure. Manipulation of broad histone marks and enhancer elements by overexpressing the histone H3 lysine 9/36 tri-demethylase KDM4A impacts

*Correspondence: sadreyev@molbio.mgh.harvard.edu (R.I.S.), johnathan.whetstine@fccc.edu (J.R.W.).

AUTHOR CONTRIBUTIONS

Conceptualization and design, C.V.R. and J.R.W.; conceptualization and design of quantitative analyses, F.J. and R.I.S.; computation analyses, F.J.; methodology, investigation, writing – original and revised draft, F.J., R.I.S., C.V.R., and J.R.W.; investigation and methodology, J.C.B. and D.C.

SUPPLEMENTAL INFORMATION

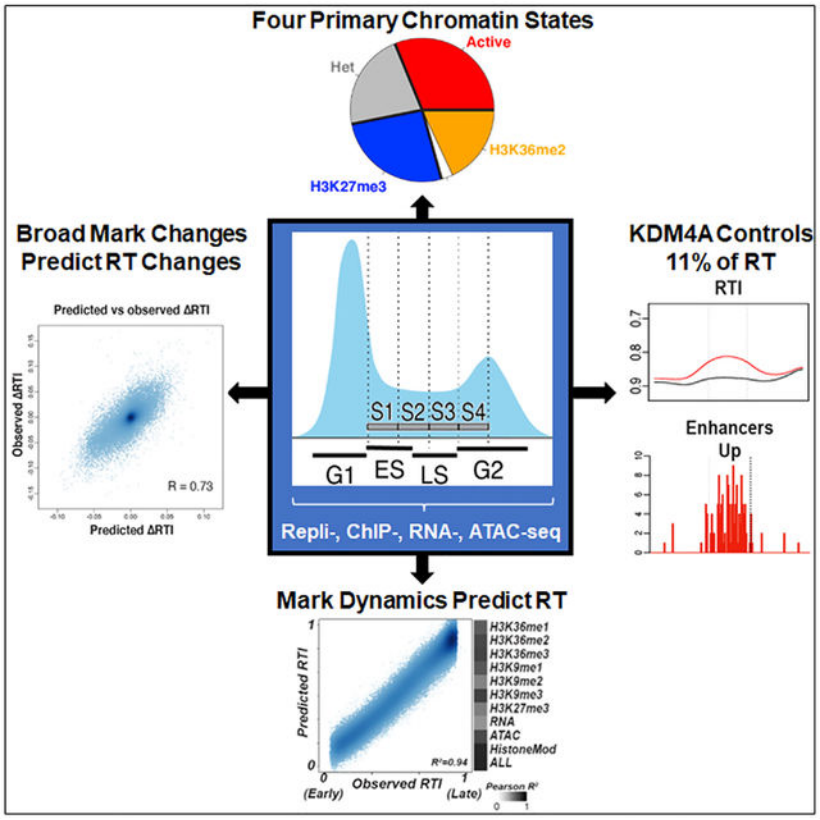
Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109799.

DECLARATION OF INTERESTS

In the past year, J.R.W. is or was serving as a consultant or advisor for Qsonica, Salarius Pharmaceuticals, and Daiichi Sankyo, Inc. The other authors declare no competing interests.

RT across 11% of the genome. Broad histone modification changes were strong predictors of the observed RT alterations. Lastly, replication within H3K36me2-enriched neighborhoods is sensitive to KDM4A overexpression and is controlled at a megabase scale. These studies establish a role for collective chromatin mark regulation in modulating RT.

Graphical Abstract



In brief

Van Rechem et al. establish that collective histone mark dynamics and their associated regulation modulate replication timing (RT). Four primary chromatin states define 97% of the mappable genome and associate with RT patterns. The K9/K36me3 demethylase KDM4A controls 11% of RT genome-wide through broad histone marks, especially H3K36me2, and enhancers.

INTRODUCTION

DNA replication occurs during S phase and is temporally regulated. The temporal order in which genomic regions replicate is referred to as replication timing (RT). This process is conserved, developmentally regulated, and dysregulated in diseases such as cancer (Nathanailidou et al., 2020; Vouzas and Gilbert, 2021). In recent years, studies have focused on establishing critical genomic and epigenomic features that regulate RT. Active and heterochromatic chromatin states, transcription, chromatin accessibility, individual chromatin marks and associated regulatory proteins, as well as elements of spatial chromatin

organization (A and B compartments, interaction domains, and loops), have varying degrees of influence on RT (Blumenfeld et al., 2021; Nathanailidou et al., 2020; Vouzas and Gilbert, 2021). However, the full extent to which their interplay regulates RT is not fully understood.

A study in mouse embryonic stem cells (mESCs) identified early replicating control elements (ERCEs). They contribute to topologically associated domain (TAD) structure, interact with transcription factors, are enriched for acetylated histone marks, and often overlap with super-enhancers in actively replicating regions (Sima et al., 2019). This study demonstrated that coordination of these *cis*-elements facilitates early replication. These elements have not been mapped in human cells, nor have the regulatory proteins controlling ERCEs been documented. Previous studies suggested a link between enhancers and replication; however, they focused on active expression and earlier replication, which could be the result of gene activation and more accessible regions in the active compartments (Iguchi-Ariga et al., 1988; Rao et al., 2014; Siefert et al., 2017; Yue et al., 2014). Collectively, these studies suggest that enhancers could be important RT determinants.

Replication domains have distinct RT patterns that are evolutionarily conserved (Ryba et al., 2010; Zhao et al., 2020). A recent high-resolution study that profiled replication at multiple time points during S phase revealed several types of local replication patterns: initiation zone (IZ), constant timing region (CTR), timing transition region (TTR), and termination site (TS) (Zhao et al., 2020). IZ corresponds to the genomic locus with earliest RT in a local neighborhood, from which replication progresses in both directions. Using a series of histone marks (H3K4me3, H3K9me3, H3K27me3, H3K27Ac), IZs were shown to correlate with active histone marks, while the repressive marks lacked clear correlations with IZs. CTRs are large regions that contain many initiation sites that fire at the same time, coordinating the regional replication that can occur early or late (Zhao et al., 2020). The CTRs that replicate in late S (LS) phase associate with heterochromatin enriched in H3K9me3 (Klein et al., 2021; Zhao et al., 2020); however, other CTRs have different RT and chromatin states. For example, loss of a key replication factor Rif-1 disrupts ~40% of RT, but not all H3K9me3-containing CTRs (Klein et al., 2021). TTRs correspond to a slope in the replication track where the replication fork moves away from initiation sites, and TSs are the points where two TTRs from opposite directions converge and thereby terminate replication (Zhao et al., 2020). Additional studies leveraging systematic temporal RT profiling, comprehensive combinations of histone marks, and chromatin accessibility across cell cycle should lead to a better understanding of the relationships between RT, local RT patterns, and chromatin states.

A prior report suggested that H3K9me2 could be more predictive of RT across species (Ryba et al., 2010); however, depletion of the key H3K9me2 methyltransferases had a nominal effect on global RT (Yokochi et al., 2009). In contrast, the H3K9/36 tri-demethylase KDM4A was shown to impact RT from *C. elegans* to human cells (Black et al., 2010). This study evaluated limited sites but established that H3K9me3 reduction was associated with earlier replication at heterochromatic loci. Although this study did not evaluate the global impact on RT or investigate how H3K36 methylation was involved, follow-up studies demonstrated significant effects on bromodeoxyuridine (BrdU) incorporation in cells overexpressing KDM4A (Black et al., 2013; Van Rechem et al., 2011). Along with

these observations, both KDM4A amplification, which occurs in ~20% of tumors, and overexpression were shown to promote site-specific DNA rereplication and amplification (Black et al., 2013, 2015, 2016; Clarke et al., 2020; Mishra et al., 2018). These data suggest that KDM4A could be a chromatin-modifying enzyme that shapes genome-wide replication and its timing. Therefore, KDM4A overexpression could serve as a unique model to interrogate the interplay between chromatin modulation and RT control genome-wide.

Using a FACS method to isolate chromatin and DNA across various time points of cell cycle and S phase, we classified the genome into four primary chromatin states that define 97% of the mappable genome at 50-kb resolution. Cell-cycle lysine methylation dynamics quantitatively predicts RT across the genome, and local replication patterns (IZs, TSs, and CTRs) associate with distinct chromatin states and spatial structure. Finally, we report that KDM4A overexpression impacts RT across 11% of the genome. The direction and magnitude of KDM4A effects on RT were predictively correlated with combined effects on broad histone marks, especially H3K36me2. Furthermore, these RT changes were associated with effects on enhancer elements and their clusters in the affected regions. Taken together, these results suggest that certain chromatin modifiers can function as important regulators of DNA replication during cell cycle through the modulation of broad histone marks and enhancer elements.

RESULTS

Profiling RT and histone modification dynamics during cell cycle

RT profiles and histone modifications were assessed genome-wide in a nearly diploid, immortalized, but not transformed, retinal pigment epithelial cells (RPEs) cell model (Black et al., 2013; Jiang et al., 1999). To avoid the impact that drug or metabolic arrest would have on replication and methylation-state dynamics, we adapted a FACS method to isolate multiple time points during S phase to evaluate DNA replication (S1–S4) and across the entire cell cycle for histone marks, gene expression, and chromatin accessibility (ATAC-seq) (G1, early S [ES], late S [LS], and G2/M; Figure 1A) in GFP-control and GFP-KDM4A-overexpressing cells. Biological replicates were consistent across the genome (Figures S1A–S1C). We also analyzed relationships with spatial chromatin structure based on previously published Hi-C data (Figure 1B) (Darrow et al., 2016).

We first defined the genome-wide patterns for the combinations of histone lysine methylation marks at 50-kb resolution in GFP-control RPE cells (control) at the level of cell-cycle averages without temporal resoultion (Figures 1B–1E and S2A). These analyses identified patterns previously noted (e.g., enriched H3K9me1 and H3K36me3 at active genes; Figure 1B) (Black et al., 2012). We also reveal that H3K36me2 occupies a large portion of the genome (29%, 904 Mb; Figure 1D) and is anti-correlated with H3K9me3 (Figure 1C). When both H3K36me2 and H3K9me3 are depleted at a genomic region, these regions are predominantly enriched in H3K27me3 (Figure 1C, red points), demonstrating that H3K36me2, H3K9me3, and H3K27me3 demarcate the majority of the genome into mutually exclusive neighborhoods. By analyzing publicly available data (Stafford et al., 2018), we confirmed these findings. The contraction of H3K27me3 regions upon a genetic perturbation was accompanied by a coordinated expansion of neighboring H3K36me2

regions to the new H3K27me3 boundary (Figure S2B). Collectively, these data are consistent with previous observations documenting the exclusivity between combinations of these histone marks at the local and more global level (Alabert et al., 2020; Streubel et al., 2018; Weinberg et al., 2019; Yuan et al., 2011).

To understand the degree of conservation for the H3K36me2 genomic distribution, we compared H3K36me2-enriched regions in our dataset with two other diverse cell types for which H3K36me2 chromatin immunoprecipitation sequencing (ChIP-seq) data were publicly available (García-Carpizo et al., 2016; Kuo et al., 2011; Stafford et al., 2018). H3K36me2 is enriched across a large fraction of the genome in each cell type with significant genomic overlap (~40%–50% at 50-kb resolution; Figure 1D). These data suggest significant conservation of these large H3K36me2 regions in the human epigenome.

Chromatin states defined at 50-kb resolution annotate 97% of the mappable genome

We applied the ChromHMM method (Ernst and Kellis, 2012) to define chromatin states at a 50-kb genomic resolution using our combination of broad histone marks across the genome. The genome was classified into four primary chromatin states (Figures 1E and S2C): active chromatin (red) (e.g., H3K36me1-3, H3K9me1), heterochromatic (gray; H3K9me3-enriched), H3K27me3-enriched (blue), and H3K36me2-enriched regions with an absence of other surveyed broad marks (orange). We compared these states with those defined by ENCODE or the Roadmap Epigenomics initiatives and observed a high degree of overlap (Figures 1F and S2D) (Kundaje et al., 2015). The orange H3K36me2 state was not previously identified by ENCODE or Roadmap Epigenomics, which was likely due to the H3K36me2 mark not being included in their analyses (Ernst and Kellis, 2012; Kundaje et al., 2015). The orange state corresponds to the previously annotated “quiescent” regions that lacked enriched marks and almost fully completes the assignment of genomic regions to primary chromatin states, increasing the annotated genomic fraction from 78% to 97% (Figures 1F and S2D). This state is mostly intergenic, multi-megabase in size, and includes multiple TADs (Figure 1B). Our data are consistent with a prior report noting wide H3K36me2-enriched regions; however, the breadth of genome coverage was unappreciated (Streubel et al., 2018; Weinberg et al., 2019).

These four chromatin states have preferential patterns for their adjacency (Figure S2E). The orange state corresponds to 17% of the genome and is separate but closest to the active red state in the PCA plot, which is likely explained by the presence of H3K36me2 in both states (Figure S2C). The red and orange states are strongly separated from the gray heterochromatic state, whereas the blue H3K27me3-enriched state has an intermediate positioning (Figure S2C). The four chromatin states are often associated with chromatin interaction domains defined by Hi-C (Figures 1B, S2F, and S2G). The average intensity of chromatin interactions is higher within a region with the same chromatin state than between adjacent regions with different states (Figure S2F), which is consistent with the average chromatin insulation scores being elevated at the chromatin-state boundaries (Figure S2G). This observation was especially true at the boundaries between active and H3K27me3-enriched states (Figures S2F, top middle, and S2G, red curve). The red active and gray heterochromatic states mostly localize to A and B chromatin compartments,

respectively (Hi-C heatmap in Figure 1E, compartments A and B marked as green and purple, respectively) (Lieberman-Aiden et al., 2009). However, regions in the blue H3K27me3 and orange H3K36me2 states were distributed between A and B compartments (Figure 1E). These two states have different self-insulation properties. The H3K27me3 blue state preferentially interacts with itself (Figure S2H), which was consistent with a distinct H3K27me3 spatial compartment in colorectal cancer cells (Johnstone et al., 2020). In contrast, the H3K36me2 orange state does not have a strong preference to interact with itself (Figure S2H), suggesting that this state does not form a separate spatial compartment but has more distributed interactions. Collectively, these data reveal an enrichment in the structural domain boundaries between these chromatin states and suggest that the blue and orange states lack strict adherence to the A or B compartments.

Histone lysine methylation dynamics associate with RT within individual chromatin states

Genomic regions retain their primary chromatin state throughout each phase of cell cycle (Figure S3A). However, many regions show quantitative changes of lysine methylation levels and chromatin accessibility during cell cycle (Figure 2). Therefore, we focused on each individual chromatin state and assessed the relationship of RT to the temporal progression of histone methylation, chromatin accessibility, and gene expression (Figure 2). In addition to the higher-resolution RT profiles (S1–S4), we also performed Repli-seq at a more traditional resolution of ES, LS, and G2/M phases (Figure 1A), allowing a direct comparison to our histone mark profiling throughout cell cycle (Figures 2 and S3B). With the exception of the gray heterochromatic state, genomic regions with a given chromatin state can have multiple RT patterns (Figure 1E, S1–S4; Figure 2, replication heatmaps). These patterns are often associated with distinct combinations of genomic features: gene density, gene activity, and lamina-associated domains (LADs) and repeats (specifically SINEs) (Figure 1E).

The gray heterochromatic state is H3K9me3 enriched (Figures 1E and 2A, left histone mark heatmaps in blue) and has a uniform homogeneous pattern of late RT and virtually no RNA expression (Figure 2A). H3K9me3 levels gradually increased between G1 and LS phase, followed by the abrupt drop between LS and G2 phases (Figures 2A–2C). Genomic tracks of H3K9me3 enrichment are shown for both the average levels in asynchronous cells (Figure 2B, top panel; Clarke et al., 2020) and the levels at each cell-cycle phase (Figure 2B, bottom panel, colored lines). Figure 2C graphically summarizes the genome-wide progression of the average level of replication signal (gray dotted line) and H3K9me3 enrichment (blue line) in heterochromatin across cell-cycle phases. This uniformly late replication at the heterochromatic regions and the concomitant reduction of H3K9me3 enrichment after LS phase (Figure 2C, gray shaded area) are consistent with a dilution of chromatin marks as the new DNA copy is synthesized during replication (Reverón-Gómez et al., 2018; Stewart-Morgan et al., 2020).

In contrast with the uniform temporal behavior of heterochromatin regions, other chromatin states contain multiple RT patterns that associate with distinct patterns of temporal progression of histone methylation and chromatin accessibility (Figures 2D–2I). For example, the red active chromatin state is enriched in H3K36me1-3 and H3K9me1 (Figure

2D, blue histone mark heatmap) and separated into three RT subgroups. The first subgroup (top cluster in the heatmaps of Figure 2D) has the expected pattern of very early replication, with the maximum BrdU signal in S1. This subgroup has increased chromatin accessibility at G1 and ES phase and enrichment of activating histone marks at G1, which are reduced at the time of replication in ES phase and restored at G2/M (Figures 2D and 2E, early on top). The second subgroup has a less expected behavior for active regions (Figure 2D, mid-cluster). These regions primarily replicate closer to the early-middle of S-phase, with the maximum of replication signal at time point S2 (Figures 2D and 2E). This distinct pattern of replication is associated with temporal patterns of chromatin accessibility and histone marks (Figure 2D). Finally, a smaller third subgroup of active chromatin-state regions replicate much later in S-phase, with the maximum of replication signal at mid-late time point S3 (Figure 2D, late cluster). In these regions, chromatin accessibility peaks in LS phase, followed by a drop in G2. The histone modifications are lowest at the LS cell-cycle phase (Figure 2E, late on bottom), consistent with chromatin mark dilution during replication.

Genomic regions in the blue H3K27me3-enriched state (Figure 2F, histone modification heatmap) have a wide range of temporal changes coordinated between histone methylation marks, chromatin accessibility, and RT. H3K27me3 regions replicate either in ES phase (early cluster: S1–S2 time points), middle-LS phase (middle cluster: S2–S3 time points), or LS (late cluster: S3–S4 time points) (Figure 2F, replication heatmaps). These subgroups have distinct temporal patterns of chromatin accessibility and H3K27me3 across cell cycle (Figure 2F). Consistent with the repressive role of H3K27me3, the ATAC-seq patterns of chromatin accessibility are often different from the dynamics of H3K27me3 and typically follow the increased H3K27me3 levels in the preceding cell-cycle phase.

The orange H3K36me2-enriched state (Figure 2H, histone modification heatmap) is another example of a chromatin state with a wide range of RT profiles. Genomic regions in this state comprise three distinct subgroups: the early cluster (BrdU signal achieves its maximum at the S1–S2 time points), the middle cluster (maximal BrdU signal at the S2–S3 time points), and the late cluster (maximal BrdU signal at the S3–S4 time points) (Figure 2H). These regions show a particularly strong association between the dynamics of a single histone modification (H3K36me2) and RT pattern at a given genomic region. Distinct patterns of chromatin accessibility demarcate the same RT subgroups but often differ from the H3K36me2 dynamics (Figure 2H). For example, the middle and late clusters have increased H3K36me2 preceding the increase in ATAC signal, whereas the early cluster has complete discordance between H3K36me2 maximum in G2/M and ATAC-seq maximum in ES (Figure 2H).

Studies have demonstrated a 2-fold dilution of chromatin marks at the time of synthesis of the new DNA copy during replication (Stewart-Morgan et al., 2020). These patterns are observed at many genomic regions in our data (especially in red active and gray heterochromatin states; Figures 2A–2E); however, many regions differ from these associations. For example, the orange state early subgroup (Figures 2H and 2I) reaches maximal BrdU signal in ES phase (S1–S2, Figures 2H and 2I) but maximal H3K36me2 in G2 (Figures 2H and 2I), suggesting that H3K36me2 accumulates at G2/M and then becomes reduced in G1-LS. Similarly, the late subgroup (S3–S4, Figures 2H and 2I) has an abrupt

increase in H3K36me2 between G1 and ES phase, inconsistent with gradual restoration after the previous cell-cycle dilution. Notably, H3K36me2 dynamics in the orange state differs from the active (red) state (Figures 2D and 2E), even when RT patterns are similar. For example, H3K36me2 progression in early and mid-S subgroups in the orange state (Figures 2H and 2I) differs from the early and mid-S subgroups in the red state (Figures 2D and 2E). Taken together, these data suggest another layer of influence by the primary chromatin state on the temporal relationship between RT and histone mark dynamics.

New measure of DNA RT: RT index

We generalized the traditional metric of RT, early-to-late (E/L) ratio, to a sensitive numerical measure, RT index (RTI), that allows RT quantification at a given genomic region as one continuous value between 0 and 1 for RT data of any experimental nature from multiple time points (see STAR Methods for details). This measure is conceptually similar to prior RT quantitation used for microarray data (Raghuraman et al., 2001). Figure 3A shows examples of RTI calculation at three selected loci with early, late, and intermediate RT. The locus with early replication (green) has the strongest read density at time point S1, which results in a small RTI close to 0 (RTI = 0.10), whereas the locus with late RT (purple) has the strongest read density at time point S4, which results in a large RTI close to 1 (RTI = 0.91). When compared with the genomic BrdU tracks at time points S1–S4, RTI as a single number recapitulates the original replication signal across the S phase (Figure 3A and S3C). RTI has a general genome-wide correlation with the traditional E/L ratio; however, many genomic regions showed differences at the four-point resolution (Figures S3C and S3D). These data are consistent with a recent report where using more S-phase time points revealed additional RT details (Zhao et al., 2020).

Cell-cycle histone lysine methylation dynamics predict RT across the genome

We analyzed the quantitative relationship of RTI to histone methylation and spatial chromatin compartmentalization on a continuous scale. We observed that the known binary relationships between RT and chromatin compartments (A, replicating early, green; B, replicating late, purple) hold on a continuous scale (Figure 3B, two top panels). Genomic tracks of the Hi-C eigenvalue as a continuous measure of the propensity for a region to belong to either A or B compartment (Figure 3B, Hi-C) are strikingly similar to the genomic tracks of RTI as a continuous measure of RT (Figure 3B, RT). The similarity suggests that the association between structural compartments and RTI holds at a higher level of quantitative detail, and even regions with intermediate A/B propensities replicate at the corresponding intermediate time points of the S phase. The densities of histone marks also show a notable correlation with both RTI pattern and chromatin compartments (Figure 3B, tracks below RT).

We then trained and evaluated computational models of RTI as a function of the combination of ChIP-seq, ATAC-seq, and RNA sequencing (RNA-seq) signals at a given genomic region. We observed that the combination of average signals across cell cycle was correlated with RT (Figures 3C and S3E; average, $R^2 = 0.71$); however, this correlation was substantially increased when considering the dynamics of these signals through cell cycle (Figures 3C and S3E; Sync, $R^2 = 0.94$). Among specific chromatin marks, the dynamics of H3K9me3

and H3K36me2/3 were the strongest predictors of RT ($R^2 \geq 0.80$), whereas gene expression (RNA-seq; $R^2 = 0.47$), H3K9me2 ($R^2 = 0.54$), and H3K27me3 ($R^2 = 0.56$) were the least correlated (Figures 3C and S3E). Taken together, these data illustrate that the collective dynamics of histone lysine methylation marks can predict RT across the genome.

Local chromatin states and spatial compartments associate with RT

We further assessed the preferences of chromatin states to replicate at specific time intervals within S phase. Based on RTI tracks, we split the whole genome into local replication regions between a local minimum and the nearest local maximum of RTI, which correspond to the local start and end of replication, respectively. We separated these local genomic regions into subgroups by the temporal interval of their replication from start to end (e.g., S1–S4, S1–S3, etc.; Figures 3D, 3E, and S4). Unexpectedly, the local regions whose replication spans the whole S phase (row S1–S4 in Figure 3E) do not comprise the most prominent fraction among all possible timing intervals. Instead, the largest number of local regions fully replicate within a part of cell cycle including the mid-S phase (Figures 3D and 3E). In each subgroup, we determined the propensity for the red, blue, orange, and gray chromatin states (Figure 3E). Consistent with our initial observations (Figure 2A), if replication ends at the latest time point S4, it is strongly associated with the gray heterochromatic state. However, when replication ends earlier than S4 (i.e., S2 or S3), this ending typically occurs in the blue H3K27me3 state and to a lesser extent in the orange H3K36me2 state (Figures 3D and 3E). By contrast, the red active chromatin state has a more general association with the start of replication, not just “early” RT (i.e., S1), even if replication starts later, for example, at time point S2 (Figure 3D). Taken together, these data suggest that the replication of a local genomic neighborhood is most often completed within only a fraction of S phase, and the timing of this fraction is associated with chromatin states in this neighborhood.

Local replication patterns associate with chromatin states and structure

Zhao et al. (2020) recently analyzed genomic replication profiles at a high temporal resolution and classified local patterns of replication into four primary types. Our replication data at four-point resolution identified the same local replication patterns (Figure 4A). We further subdivided CTRs (regions of RTI variance <0.006 over >500 kb) into (1) the termination CTRs (TCTRs) located at the end of the local replication process and preceded by the converging TTRs at their genomic flanks; and (2) the initiation CTRs (ICTRs) located at the start of the local replication process and followed by the diverging TTRs at their flanks (Figure 4A).

We then assessed the distribution of the four chromatin states among each type of RT pattern. Each RT pattern associates with a preferred chromatin state, except TTRs where replication simply passes through the region (Figures 4B and 4C). The majority of IZs correspond to the active red state (62.6 Mb; Figure 4C) regardless of whether the IZ occurs at the earliest time point or later in S phase (Figure 4B). TSs are primarily in the H3K27me3-enriched blue state regardless of whether the termination occurs earlier or later in S phase. However, most TSs replicate earlier than the end of S phase. Lastly, most TCTRs are in the gray heterochromatic state and replicate at the very end of S phase

(S4). In contrast, ICTRs are enriched in the red active state, replicating early in S phase (Figure 4B). A large fraction of IZs in active chromatin state correspond to individual interaction domains and have high Hi-C insulation scores at their boundaries that coincide with the genomic boundaries between the active state and a neighboring chromatin state, most often blue H3K27me3 or orange H3K36me2 states (Figures 4D, 4E, and S2G). These domains of ~100 kb or less are much smaller than the domains corresponding to the gray heterochromatic TCTRs or the blue H3K27me3-enriched TSs. TSs are typically located in the center of a large chromatin interaction domain, which corresponds to the surrounding H3K27me3-enriched region (Figures 4F and 4G). The boundaries of this domain are marked by high Hi-C insulation scores and co-localize with the boundaries of the H3K27me3 state.

The spatial organization of CTRs depends on the chromatin state. The TCTRs in the gray heterochromatic or orange H3K36me2 state are typically located within a single large chromatin interaction domain (Figures 4H and 4I), which is in contrast with ICTRs in the red active chromatin state with multiple interaction domains (Figure 4J). Chromatin interactions within the single TCTR domain formed regardless of TCTR size (Figure 4K) may facilitate the concerted replication initiation across the whole TCTR as a single unit. However, the number of interaction domains within active-state CTRs is roughly proportionate to its width, suggesting a different spatial organization that involves multiple smaller domains of similar size. Taken together, these results demonstrate that local replication patterns preferentially occur in specific chromatin states and associate with internal spatial organization.

KDM4A controls RT in association with broad histone modifications

KDM4A overexpression affected replication across 11% of the genome (Figure 5A) based on a cutoff of RTI change consistent with the cutoff stringency previously applied to the changes in E/L ratio (see STAR Methods for details). This large genomic fraction was observed because of the refined resolution of four S1–S4 time points, while the traditional resolution of ES versus LS phase provided the detection of replication changes among only 2% of the genome (Figure 5A). This difference was mainly due to the profiling of mid-S phase (S2, S3; Figures 5B and 5C). Figure 5C shows examples of genomic Repli-seq tracks highlighting the cases of RT differences that were apparent in the S1–S4 time points but were challenging to detect using the traditional ES and LS profiles, which was likely a result of the averaging of variable signals across these two larger temporal intervals. These RT shifts predominantly occurred toward an earlier RT (7.5% out of total 11% of the genome; Figures 5A and 5B).

We then analyzed the effects of KDM4A overexpression on histone marks with broad diffuse patterns. H3K9me3 and H3K36me3 levels were preferentially reduced, consistent with the enzymatic role of KDM4A, whereas H3K9me1, H3K36me1, and H3K27me3 were preferentially increased (Figures 5D, 5E, and S5A). The widest changes, over ~2.4% of the genome, were observed for H3K36me2, which occurred in both directions to approximately equal extent (Figure S5A). Effects of KDM4A overexpression were correlated between individual broad histone marks (Figure S5B). These changes, however, were mostly

quantitative, and the majority of these regions (72%) did not switch chromatin states (Figures S5C and S5D).

Although the changes of individual histone marks had mostly modest degrees of genome-wide correlation with the direction and/or magnitude of RT shift (Figure S5E), the combination of these changes showed a much stronger association with RT. The increase of H3K36me1-3 and of H3K9me1, as well as the coordinated reduction of H3K9me3 (Figure 5E, top part of the heatmap), was largely corresponding to a more active state and the shift to an earlier replication. The reduction of H3K36me1-3 and H3K9me1, as well as the coordinated increase of H3K9me3 (Figure 5E, bottom part of the heatmap), was largely corresponding to a more repressed state and the shift to a later replication. We observed that RT changes were centered on the strongest changes of histone marks and extended to a wider genomic vicinity (Figure 5F).

We then tested whether a combined quantitative analysis of all absolute levels of histone marks in control cells and their changes upon KDM4A overexpression could predict RT change at any given genomic region. We observed a correlation between predicted and experimental values of RT shift (Pearson R = 0.73; Figure 5G), suggesting that the behavior of the seven broad histone marks is sufficient to quantitatively explain the observed RT changes with high genome-wide accuracy. The actual RT shifts were observed in most predicted regions (Figure 5G). The directions of these changes were coordinated among these marks, regardless of whether the magnitude of these individual changes was strong (Figure 5H, top of the heatmaps) or modest and possibly below the threshold of confident detection (Figure 5H, lower parts of the heatmaps). These changes of histone marks were a stronger predictor of RT shifts than RNA- or ATAC-seq changes (Figures S5F and S5G). Taken together, these data suggest that combined changes of multiple broad histone marks can be an accurate predictor of RT change throughout the genome (Figures 5H and S5H).

KDM4A-driven RT changes accompany enhancer element modulation genome-wide

We then assessed the impact that KDM4A overexpression had on the genomic distribution of focused histone marks with narrow peaks (i.e., H3K27ac, H3K4me1-3, H3K9ac; Figure S6A) and their relationship to RT changes. Many genomic loci where these peaks were increased or decreased by KDM4A overexpression coincided with corresponding changes of broad histone marks (Figure 6A). These peaks were often located within wider regions of RT change and sometimes occurred as genomic clusters of multiple loci (Figure 6A).

Regions of KDM4A-associated RT shifts strongly overlapped with these changing peaks. For example, earlier RT shifts had much more significant overlap with the increased peaks of focused activating histone marks ($p = 1 \times 10^{-1531}$) than with the increased peaks of chromatin accessibility (ATAC, $p = 2 \times 10^{-5}$) or increased gene expression (upregulated genes, $p = 2 \times 10^{-3}$) (Figure 6B, upper Venn diagrams). A similar pattern was observed for the overlaps of regions shifting to a later replication with decreasing ATAC-seq peaks and transcriptionally downregulated genes (Figure 6B, lower Venn diagrams). Furthermore, very few peak changes coinciding with RT shifts were associated with gene expression changes in the genomic neighborhood, because less than 10% of these peaks were positioned within 1 Mb from a differentially expressed gene (DEG). Their distribution around the RT shift was

similar to the peaks without any positional adjacency to a DEG (Figure 6C). Taken together, these data suggest that these peaks are likely linked to RT, but not to gene expression.

Virtually all loci of observed KDM4A effects on focused histone marks had a similar histone mark signature of strong H3K27ac and H3K4me1-2, much weaker H3K4me3 levels, and the absence of H3K27me3 or H3K36me3 (Figure 6D). This signature is consistent with active enhancer elements and was similar to histone mark densities at chromHMM-annotated enhancers in wild-type RPE cells and different from the active promoter signature (Figure S6B). We further confirmed this prevalence of enhancers by applying the Roadmap Epigenomics HMM to our data, which confidently assigned 92% of these KDM4A-affected loci to the enhancer state (Figure 6E). KDM4A overexpression modulated both focused and broad histone marks in a coordinated fashion (Figures 6A and 6F). Specifically, increased histone marks at the enhancer elements (Figure 6F, peaks up) often corresponded to increased broad activating marks, reduced broad repressive marks, and a shift to earlier replication (ΔRTI heatmap; Figure 6F). The association between enhancer elements, broad histone marks, and RT was pronounced among 864 enhancers with strong, high-confidence earlier RT shifts (Figure 6F, peaks up, upper row of heatmaps). However, this relationship extended to ~1,800 enhancers with shifts below our increased RTI cutoff (Figure 6F, peaks up, second row of heatmaps). In contrast, loci with shifts to a later RT had reduced focused and broad activating histone marks and increased broad repressive marks (Figure 6F, peaks down).

We further demonstrate that the change of focused histone marks at multiple adjacent enhancer elements was associated with RT shift in this neighborhood (Figure 6F, differential peak density). Among the clusters of adjacent enhancer elements with a concerted increase of focused histone marks, 85% shifted to earlier RT, but none shifted to later RT (Figure 6G), while the clusters of enhancer elements with reduced focused histone marks were enriched in later RT shifts (Figure 6G). Taken together, these data suggest that concomitantly regulated clusters of enhancer-like elements are tightly linked to the regulation of RT.

KDM4A-associated RT changes depend on both chromatin state and local replication pattern

After assessing the genome-wide effect KDM4A overexpression had on histone marks and RT, we assessed its impact on individual types of chromatin states and local replication patterns. The red active chromatin state tends to replicate early (Figure 7A, red points with early control RTI, x axis) and often shifts to even earlier RT upon KDM4A overexpression (Figure 7A, red points above the cutoff line for ΔRTI, y axis). Both the mid-S phase replicating orange H3K36me2 state (Figure 7A, orange points above the cutoff line) and the late replicating gray heterochromatic state often shift to earlier replication (Figure 7A, gray points above the cutoff line). In contrast, the mid-S-phase replicating blue H3K27me3 state consistently shifts to a later RT (Figure 7A, blue points below the lower cutoff line). Taken together, these data suggest that KDM4A overexpression has different RT effects among the primary chromatin states.

We then assessed these different replication effects in connection with local replication patterns. The largest portion of the genome with a shift to earlier RT corresponded to the orange H3K36me2 state (Figure 7B). In this state, all types of local RT patterns shifted to earlier RT, with TCTRs covering the largest total length (22.7 Mb; Figure 7B). Other chromatin states had stronger preferences to be affected within a specific local RT pattern. Earlier RT shifts disproportionately affected active state within IZs and heterochromatin within TCTRs (Figure 7B). In contrast, later RT shifts disproportionately affected TSs in H3K27me3 state (Figure 7C). These data highlight the different modes of KDM4A regulation depending on both chromatin state and local replication pattern and emphasize the importance of the orange H3K36me2 state in KDM4A-driven RT changes.

Diverse KDM4A effects on local RT patterns associate with distinct modes of enhancer modulation

We then focused on the relationships between KDM4A effects on RT and histone marks among genomic neighborhoods grouped by local replication pattern and chromatin state. RT changes in gray heterochromatic TCTRs typically presented as a narrow region forming a new IZ within the wider TCTR (Figure 7D). These local RT shifts largely co-localized with reduced H3K9me3 and increased activating histone marks (H3K4me1-3, H3K27ac, H3K9ac, H3K36me2; Figures 7D and 7E). In fact, enhancer elements with increased activating histone marks specifically concentrated at the region of RT shift (Figure 7E). The spreading of local RT shift from these focally enriched enhancer elements had an interesting behavior. When replication entered the surrounding heterochromatin neighborhood, this RT shift did not abruptly drop to zero but was gradually reduced over hundreds of kilobases until replication reached the original timing (Figure 7E). This pattern suggests that the focal KDM4A effect on RT is gradually dampened in the adjacent gray heterochromatic state.

We next analyzed the association between earlier RT shifts and histone mark changes at IZs in the red active state (Figure 7F). The locus of strongest RT shift coincided with the earliest replicating site in the middle of IZ (Figures 7F and 7G). Similar to the effect in TCTRs, these RT shifts did not abruptly stop at the boundaries of active chromatin state but instead continued for at least a few hundred kilobases into the neighboring regions, suggesting that the propagation of replication is gradually slowed down along the genomic length (Figure 7G). The earlier RT shifts were associated with increased activating histone marks and decreased repressive histone marks (Figures 7F and 7G). These regions had an especially strong enrichment for enhancer elements with increase of focused activating histone marks (Figure 7G), which further emphasized the importance of enhancer elements at regions undergoing RT changes.

In contrast with the other local replication patterns that were typically enriched in earlier RT changes, KDM4A-associated RT changes at TSs occurred equally in either direction (Figures 7B and 7C). TSs in the blue H3K27me3 state shifted to later RT, with the strongest difference at the TS (Figures 7H and 7I). Although we detected only modest changes of repressive broad histone marks (Figure 7I), a large number of enhancer elements had reduced activating histone mark peaks across the wider H3K27me3-enriched region (Figure 7I).

Unlike TSs in the blue H3K27me3-enriched state, TSs in the orange H3K36me2 state shifted to earlier RT (Figure 7B) as a part of wider H3K36me2-enriched neighborhoods that often include smaller red active-state regions (Figure 7J). The earlier RT shift was maintained across the entire neighborhood, included all types of local RT patterns, and often occurred at a multi-megabase scale (Figures 7J and 7K). Within these neighborhoods, regions with the strongest RT shifts were associated with changes in broad histone marks that were accompanied by increased peaks of activating histone marks at annotated enhancer elements (Figures 7J and 7K). The degree of RT change in these regions was similar to a typical RT change at an IZ (Figure 7G). This change extended, albeit at a lower magnitude, into the wider H3K36me2 neighborhood for long genomic distances (Figures 7J–7L), which contrasts with the sharper decline in RT shifts upon crossing into blue H3K27me3-enriched or gray heterochromatic states (Figure 7L). These data suggest that the orange H3K36me2-enriched state is controlled as larger genomic units and is highly sensitive to KDM4A overexpression. Taken together, these results emphasize the important roles of both chromatin state and local replication pattern in determining KDM4A effect on replication, as well as the association of enhancer elements with RT changes in either direction.

DISCUSSION

Our study demonstrates that at 50-kb resolution there are three distinct histone marks that are exclusive of one another across the genome (H3K9me3, H3K27me3, and H3K36me2), which is consistent with prior studies (Alabert et al., 2020; Streubel et al., 2018; Weinberg et al., 2019; Yuan et al., 2011). The corresponding chromatin states, along with the active state, were insulated from one another and maintained during cell cycle, while having distinct roles in local initiation, propagation, and completion of replication. Regions within a state can have various temporal dynamics of chromatin marks, as well as different RT profiles. The orange H3K36me2 state is unique in that it is predominantly intergenic and contains no other tested modifications, often expands over megabase-scale regions, is enriched in the mid-S phase, and has unusual relationships of its chromatin dynamics with RT. This state like the blue H3K27me3-enriched state does not exclusively belong to the A or B compartment. The H3K27me3 state is preferentially self-insulated, which is consistent with a recent report in colorectal cancer cells (Johnstone et al., 2020). Thus, three of the four primary chromatin states correspond to their own spatial compartments, consistent with each of these states physically forming separate phase condensates (Laflamme and Mekhail, 2020). In contrast, the H3K36me2 state has more promiscuous interactions with other states and likely forms a spatially distributed part of the genome around these condensates. Future studies need to evaluate additional genomic and epigenomic characteristics in these states and their relationship to genome organization.

RT profiling at four S-phase time points revealed details of local RT patterns and improved the analysis of KDM4A-induced RT changes. Our temporal profiling of chromatin marks and accessibility established temporal dynamics of chromatin marks as a strong quantitative RT predictor ($R^2 = 0.94$). We also highlight that gene expression is not the strongest RT predictor. We further tested the causal relationship between chromatin marks and RT by manipulating KDM4A and analyzing the resulting RT alterations. Broad histone

mark changes were quite predictive of RT changes across the genome (Pearson R = 0.73). Furthermore, the relationship between chromatin marks and RT depended on the context of chromatin state and local RT pattern. For example, the orange state had an enriched sensitivity to KDM4A overexpression across all types of local RT patterns and had concerted RT regulation across single, often large, genomic units. The variety of local scenarios for different chromatin states and RT patterns may explain why our integrative analysis of multiple histone marks combined with a higher-resolution RT produced more accurate overall RT predictions across the genome. Future studies should use high-resolution RT to assess how other chromatin factors impact chromatin states, RT, and local RT patterns.

Although links have been reported between RT, active genes, and enhancers, the question remained as to whether there was a strong association between these elements genome-wide. We document a relationship between adjacent enhancer elements and RT. Modulation of focused active marks at enhancers was not the strongest predictor of RT change (data not shown), suggesting their functioning in concert with other events facilitating RT change. Consistent with our observation, another study noted that ERCEs control some early RT events in mouse ESCs, have partial overlap with enhancers, and could act as clusters (Sima et al., 2019). Our data suggest that a multi-factorial relationship exists in controlling RT and RT patterns throughout the genome that incorporates chromatin states, broad histone methylation marks, and enhancers. These enhancers were not associated with differential gene expression, raising the question as to whether these elements are true enhancers or possibly distinct enhancer-like elements. Future studies need to (1) explore whether dynamically controlling this methylation mark serves as a switch from replication to transcription, (2) investigate additional features or factors that coordinate RT and local patterns through these enhancers, and (3) assess whether altering the transcriptional potential of enhancers that are associated with replication promotes a change in RT and/or local instability.

Our study emphasized the importance of chromatin regulation in understanding how RT is controlled. KDM4A regulates RT genome-wide with profound effects on mid-S-phase replication and the orange H3K36me2 state. This impact is a critical insight into KDM4A function because it is amplified in ~20% of tumors analyzed in The Cancer Genome Atlas (TCGA) (Black et al., 2013) and is associated with minimal gene expression changes unless evaluated during cell cycle (Van Rechem et al., 2020). Future studies should (1) evaluate the impact that KDM4A-amplified and -overexpressed tumors have on RT dynamics and the orange H3K36me2 state, and (2) address whether the effects on RT are a direct or indirect role for the KDM4A enzyme. Because many H3K36 KMTs and KDMs are mutated or altered in cancer (Van Rechem and Whetstine, 2014), future studies should consider how they impact RT, not just gene expression, because their control of RT could be a determinant in cancer.

Limitations of the study

In this manuscript, we apply standard input normalization techniques to ChIP-seq data. However, we cannot completely exclude the possibility that this normalization does not fully eliminate the confounding direct effect of DNA copy number on ChIP and input

DNA signals, which may lead to additional contribution of DNA copy number to the observed correlation of the normalized ChIP-seq intensities with RT (Figure S7). However, the arguments in favor of the biological role of the collective histone marks versus DNA copy number as the main driver are as follows: (1) there were different and opposite patterns of correlation with RT among individual marks despite the same input data being used for normalization (Figure 2), (2) there were different levels of correlation with RT among genomic regions that are enriched in a specific mark versus regions that are not enriched in the same mark when compared with the input (data not shown), and (3) the data reproduced correlations with both RT and structural compartments that were previously noted in other independent studies that focused on specific marks (Dileep et al., 2015; Gorkin et al., 2014; Johnstone et al., 2020; Lieberman-Aiden et al., 2009). Therefore, the data presented suggest that the combinatorial control of these histone marks has an impact on RT. Additional studies altering factors controlling these marks, like KDM4A, will provide additional testing of this impact.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Johnathan R. Whetstine (Johnathan.Whetstine@fccc.edu).

Materials availability—RPE cell lines generated in this study are available upon request to Johnathan Whetstine.

Data and code availability

- Sequencing data have been deposited at GEO and are publicly available as of the date of publication. The accession number is listed in the key resources table.
- All original code has been deposited at GitHub (https://github.com/MolBioBioinformatics/RTI_Epigenome) and is publicly available as of the date of publication. The DOI is listed in the Key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request, Ruslan I. Sadreyev (sadreyev@molbio.mgh.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines—Stable cell lines have been generated as in (Black et al., 2013). Stable RPE cell lines (hTERT-RPE-1 cell line; referred to as RPE) were generated by retroviral transduction of MSCV-GFP or MSCV-GFP-KDM4A. Cells were selected for 96 hours with puromycin. Cells were maintained in DMEM with 10% fetal bovine serum, 1% penicillin/streptomycin, and L-glutamine, and cultivated at 37°C under 5% CO2. RPE cell lines are female. Original RPE cell lines were obtained from Nicholas Dyson's laboratory (originally from ATCC hTERT RPE-1) and have not been further authenticated. Biological duplicates were used for all analyses in Figures 1, 2, 3, 4, 5, 6, and 7.

METHOD DETAILS

RNA-sequencing

FACS Sorting: The approach was previously described in (Van Rechem et al., 2020). Specifically, four 15 cm^2 plates of 2 million RPE cells were seeded for 48 hours. Cells were incubated with 1/1000 Hoechst 33342 directly into the media for 1h at 37°C degrees. Cells were trypsinized and resuspended in Hoechst-containing media. Cells were sorted with a BD FACS Fusion BV421-A laser into 1.5 mL cold Qiazol, based on DNA content. The gates for the sorting were set such that the G1 phase was collected from the bottom left of the first peak to the middle of the right slope, the G2 phase was collected from the bottom left of the second peak to the bottom right of that same peak. The ES and LS phase were equally divided in between G1 and G2 phases. 250,000 cells were collected per phase.

Library Preparation and Sequencing: RNAs were purified using the QIAGEN miRNeasy kit including a DNase treatment. Total RNA sequencing libraries were prepped using the TruSeq Stranded Total RNA Sample Preparation with Ribo-Zero kit (Illumina). 250 ng of RNA were used for library preparation following supplier's instructions, including 10 cycles of PCR amplification. Library length distribution were assessed by TapeStation (Agilent) and quantified with Qubit (ThermoFisher). Libraries were paired-end sequenced (101 cycles each way) using a NextSeq500 (Illumina).

ChIP-sequencing

FACS Sorting: 15 cm^2 plates of 2 million RPE cells were seeded for 48 hours. Cells were incubated with Hoechst 33342 at 1/1000 directly into the media for 1h at 37°C degrees. Cells were then trypsinized and resuspended in Hoechst-containing media before crosslinking with 1% formaldehyde for 13min at 37 degrees and quenching with 0.125M Glycine. Cells were washed with 1X PBS and resuspended in Hoechst-containing media. Cells were sorted with a BD FACS Fusion BV421-A laser, based on DNA content. The gates for the sorting were set such that the G1 phase was collected from the bottom left of the first peak to the middle of the right slop, the G2 phase was collected from the bottom left of the second peak to the bottom right of that same peak. The ES and LS phase were equally divided in between G1 and G2 phases.

Chromatin Preparation: Sorted cells were pelleted by centrifugation prior to resuspension in 200 µL lysis buffer (5 mM PIPES pH8 – 85 mM KCl – 0.5% NP40) and incubated 5 minutes on ice. Nuclei were pelleted by centrifugation and lysed in 10–100 µL nuclear lysis buffer (50 mM Tris pH8 – 10 mM EDTA pH8 – 1% SDS), using an estimation of 1 million cells for 25 µl. Chromatin was sonicated in 0.5 mL thin wall PCR tubes using a Q800R (QSonica) at 70% amplitude for 20 minutes of total sonication time, to get < 250bp fragments. Debris were cleared by centrifugation.

Immunoprecipitation and DNA Purification: *Pre-binding of antibodies to beads*: using 200 µL PCR strip tubes, 0.1–0.2 µg of antibody was incubated with 100 µL dilution IP buffer high SDS (16.7 mM Tris pH 8 – 1.2 mM EDTA pH 8 – 167 mM NaCl – 0.2% SDS – 0.24% or 1.84% Triton X-100), 0.24% or 1.84% Triton X-100 depending on the antibody (see antibodies section) and 2.5 µL Dynabeads (ThermoFisher) protein A or G magnetic beads

(A for rabbit polyclonal Ab, G for mouse monoclonal) for 6 hours at 4°C under constant rotation.

Pre-clear of chromatin: 0.5 μg of chromatin was used per IP (except if indicated otherwise in antibodies section). Using 200 μL PCR strip tubes, chromatin was brought to 10 μL with nuclear lysis buffer, then to 100 μL with dilution IP buffer low SDS (16.7 mM Tris pH 8 – 1.2 mM EDTA pH8 – 167 mM NaCl – 0.1% SDS – 0.24% or 1.84% Triton X-100). 2.5 μL protein A agarose was added and incubated for 2 hours at 4°C under constant rotation. Beads were spun down and discarded. 2.5 μL protein A or G magnetic beads was added and incubated for 2 hours at 4°C under constant rotation.

Immunoprecipitation: the precleared chromatin was incubated with antibodies pre-binded to beads (after discarding the dilution IP buffer from the pre-binded beads) overnight at 4°C under constant rotation. IPs were washed extensively (vortex and quick spin for buffer trapped in the caps between each wash): twice with 100 μL dilution IP buffer high SDS, once with 100 μL TSE buffer (20 mM Tris pH8 – 2 mM EDTA pH 8 – 500 mM NaCl – 1% Triton X-100 – 0.1% SDS), once with 100 μL LiCl buffer (100 mM Tris pH8 – 500 mM LiCl – 1% deoxycholic acid – 1% NP40), and twice with 100 μL TE buffer (10 mM Tris pH 8 – 1 mM EDTA pH 8). Beads were transferred in 1.5 mL tubes and DNA was eluted with 50 μL elution buffer (50 mM $NaHCO_3$ – 140 mM NaCl – 1% SDS) – 1 μL 200 μg/ml RNase A 30 minutes at 37°C followed by the addition of 1 μL Proteinase K 10 mg/ml and incubation 1 hour at 55 degrees 1000 rpm in a thermomixer (if several IPs were pooled this was done at this step, using 50 μL of elution buffer for all IPs, see antibodies section). The eluates were removed from the beads and de-crosslinked 4 hours at 65°C 1000 rpm in a thermomixer.

DNA purification: 125 μL AMPure XP beads (Agencourt) was added to the samples and incubated 10 minutes at RT. On the magnet: liquid was removed before performing two washes with 450 μL fresh 70% EtOH, according to the supplier's instructions. DNA was eluted with 40 μL H_2O.

Input: 0.5 μg of chromatin was brought to 10 μL with nuclear lysis buffer. 40 μL of elution buffer was added and protocol was followed according to the IP samples from that step forward.

Antibodies—H3K4me1 Abcam ab8895 lot GR193882–1 (1.84% Triton, 0.2 mg antibody per IP); H3K4me2 Abcam ab32356 lot GR209821–1 (1.84% Triton, 0.1 μg antibody per IP); H3K4me3 Millipore 07–473 lot 2648189 (0.24% Triton, 0.2 μg antibody per IP); H3K27Ac Active Motif 39133 lot 31814008 (1.84% Triton, 0.2 μg antibody per IP); H3K9Ac Abcam ab4441 lot GR224698–1 (1.84% Triton, 0.15 μg antibody per IP); H3K27me3 Millipore 07–449 lot 26532203 (1.84% Triton, 0.15 μg antibody per IP); H3K9me3 Abcam ab8898 lot GR30928–1 (1.84% Triton, 0.2 μg antibody per IP); H3K36me3 Abcam ab9050 lot GR10860–1 (1.84% Triton, 0.2 μg antibody per IP); H3K9me1 Abcam ab8896–100 lot 815309 (1.84% Triton, 0.2 μg antibody per IP, two IPs were pooled); H3K9me2 Abcam ab1220 lot GR32351–2 (1.84% Triton, 0.2 μg antibody per IP, two 1 μg IPs were pooled); H3K36me1 Cell Signaling 14111S ref 03/2017 lot 1 (1.84% Triton, 0.2 μg antibody per IP);

H3K36me2 Abcam ab9049 lot GR316128–1 (1.84% Triton, 0.2 μg antibody per IP, two IPs were pooled).

Library Preparation and Sequencing—ChIP sequencing libraries were prepped using the TruSeq ChIP Sample Preparation kit (Illumina) following the supplier's protocol with slight modifications. 1 ng or 35 μL of DNA was used for library preparation. After ligating the adapters, fragments were amplified with 13 cycles of PCR ("Purify Ligation Products" section was skipped). The PCR products were purified using a Double-Sided SPRI approach: AMPure was added to 0.6X SPRI concentration (30 μL of AMPure is added to 50 μL of PCR reaction) and incubated for 10 minutes at RT before being placed on the magnet. Beads were discarded and supernatant was incubated with 0.85X SPRI concentration (12 μL of AMPure is added) and incubated for 10 minutes at RT before being placed on the magnet. Liquid was removed before performing two washes with 450 μL fresh 70% EtOH. DNA was eluted with 15 μL H_2O. Library length distribution were assessed by TapeStation (Agilent) and quantified with Qubit (ThermoFisher). Libraries were single-end sequenced (75 cycles) using a NextSeq500 (Illumina).

ATAC-sequencing

FACS Sort: Four 15 cm^2 plates of 2 million RPE cells were seeded for 48 hours. Cells were incubated with 1/1000 Hoechst 33342 directly into the media for 1h at 37°C degrees. Cells were trypsinized and resuspended in Hoechst-containing media. Cells were sorted with a BD FACS Fusion BV421-H laser based on DNA content. The gates for the sorting were set as described in Replication Timing Sequencing section. 50,000 or 100,000 cells were collected per phase.

Library Preparation and Sequencing: Protocol was adapted from (Buenrostro et al., 2013). Sorted cells were pelleted and resuspended in 250 μL cold RBS buffer (10 mM Tris-HCl pH 7.4 – 10 mM NaCl – 3 mM $MgCl_2$) before being pelleted by centrifugation. Pellet was resuspended in 250 μL cold RBS buffer – 0.1% NP-40, incubated 5 minutes on ice and pelleted by centrifugation. Pellet was resuspended in 50 μL Tagmentation mix (22.5 μL H_2O – 25 μL 2X Tagment DNA buffer – 2.5 μL Tagment DNA enzyme) and incubated 30 minutes at 37°C before being placed on ice. DNA was purified with MinElute PCR Purification kit (QIAGEN) following supplier's instructions and eluted in 10 μl. Purified DNA was amplified by PCR by adding 2.5 μL 25 μM Primer Ad2 (barcode, use different primers for each samples) and 37.5 μL PCR mix (10 μL H_2O – 2.5 μL 25 μM Primer Ad1 – 25 μL 2X NEBnext master mix) and incubated as follows: 72°C 5 minutes, 98°C 30 s, [98°C 10 s, 63°C 30 s, 72°C 1 minute] × 12 cycles, 4°C. Amplified DNA was purified using a Promega Gel Purification Kit following supplier's instructions and eluted in 25 μL H_2O. Library length distribution were assessed by TapeStation (Agilent) and quantified with Qubit (ThermoFisher). Libraries were paired-end sequenced (43 cycles each way) using a NextSeq500 (Illumina).

Primers sequences (Buenrostro et al., 2013): see Table S1.

Replication Timing Sequencing

FACS Sort: Four 15 cm^2 plates of 2 million RPE cells were seeded for 48 hours. Cells were incubated with 100 μM BrdU for 2 hours at 37°C prior to be trypsinized, spun and washed with cold 1X PBS – 1% FBS, spun and resuspended in 2.5 mL cold 1X PBS – 1% FBS. 7.5 mL cold 70% EtOH was added drop by drop while vortexing the samples. Fixed cells were stored at −20°C. Cells were pelleted at 4°C, resuspended in 3 mL 1X PBS – 1% FBS – 1/100 Propidium Iodide – 250 μg/ml RNase A, and incubated 1 hour RT in the dark. Cells were sorted into 500 μL cold 1X PBS with a BD FACS Fusion Tex Red A laser, based on DNA content. The gates for the sorting were set as for the RNA Sequencing method, or dividing the S phase in four fractions: S1 starting at the middle of the right slope of the G1 peak, S4 finishing at the midpoint of the G2/M peak, with the four phases being equally divided across. 300,000 to 500,000 cells were sorted per fraction.

DNA Extraction: After sort 0.8% SDS and 20 μg/ml proteinase K were added to the cells and incubated for two hours at 55°C. DNA was extracted by adding one volume of phenol-chloroform and collecting the upper phase after vortex and spin. DNA was precipitated overnight at −80°C after addition of 1/9 volume 3M NaAc pH 5.2 – 1 μL glycogen – 2.5 volume 100% EtOH and vortex. Precipitated DNA was pelleted by centrifugation, washed with 70% EtOH and dried pellets were resuspended in 100 μL 10 mM Tris pH8. DNA was sonicated using thin wall tubes using a QSonica Q700 at 70% amplitude for 12 minutes of total sonication time in order to get 300 bp fragments.

Immunoprecipitation: Sonicated DNA were denaturated by heat for 2 minutes at 95°C followed by cooling on ice. 75 μL of DNA was brought to 450 μL by adding 45 μL 10X IP buffer (10X PBS – 0.5% Triton X-100) and RNase free H_2O. Immunoprecipitation was performed by addition of 50 μL of Dynabeads protein G magnetic beads (pre-washed and resuspended in 1X IP buffer) and 2 μL of anti-BrdU antibody (BD PharMingen 555627) and incubation at RT 2 hours under constant rotation, in the dark. Immunoprecipitates were washed five times with 1X IP buffer and vortex, and eluted with 200 μL 1X IP buffer – 250 μg/ml proteinase K 2 hours in the dark. Eluates were taken off the beads and 200 μL phenol – chloroform were added prior to vortex and spin. 1 μL glycogen – 1/9 volume NaAc pH 5.2 – 2.5 volume cold 100% EtOH were added to the upper phase and DNA was precipitated overnight at −80°C. Precipitated DNA was pelleted by centrifugation, washed with 70% EtOH and dried pellets were resuspended in 75 μL 10 mM Tris pH8.

Library Preparation and Sequencing: Libraries were prepared using the Accel-NGS Methyl-Seq DNA Library Kit (Swift Biosciences) following supplier's protocol, starting with 15 μL of immunoprecipitated DNA and following the recommendation for < 10 ng, using 7 cycles of PCR amplification. Library length distribution were assessed by TapeStation (Agilent) and quantified with Qubit (ThermoFisher). Libraries were single-end sequenced (75 cycles) using a NextSeq500 (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq analysis—STAR aligner was used to map sequencing reads to transcripts in the hg19 reference genome (Dobin et al., 2013). Read counts for individual transcripts were

produced with HTSeq-count (Anders et al., 2015), followed by the estimation of expression values using EdgeR (Robinson et al., 2010).

ChIP-seq analysis—Sequencing reads were aligned against the hg19 reference genome using BWA (Li and Durbin, 2010). Alignments were filtered for uniquely mapped reads and duplicates were removed. To determine the regions of ChIP-Seq tag enrichment for broad histone marks (H3K9me1, H3K9me2, H3K9me3, H3K36me1, H3K36me2, H3K36me3, H3K27me3), we analyzed tag counts in 50 Kb bins across the chromosome length. All counts in both replicates were quantile normalized across both control and KDM4A overexpression datasets in all phases (G1/ES/LS/G2). These normalized counts were then used to calculate ChIP enrichment over input as the $\log_2$ ratio of ChIP to input tag density. Peaks of enrichment for focused histone marks (H3K4me1, H3K4me2, H3K4me3, H3K27ac, H3K9ac) were called using HOMER (Heinz et al., 2010). To identify differential ChIP-seq enrichment between control and KDM4A overexpressing cells, we analyzed input-normalized ChIP-seq tag densities at either all 50 Kb genomic bins for broad marks, or all regions corresponding to the union of peaks called in individual samples for focused marks, and used edgeR (Robinson et al., 2010) with the cutoff of at least 1.5-fold difference between replicate averages. Public ChIP-seq data were downloaded from GEO (GSE73696, GSE118954). To compare our RPE data on H3K36me2 enrichment to other cell types, H3K36me2-enriched regions were defined as 50 Kb bins with at least 1.5-fold ChIP enrichment over input, followed by determining the overlaps between these regions in different datasets and estimating their statistical significance as Z-scores and P values based on the distribution of overlap lengths between randomly shuffled regions.

ATAC-seq analysis—ATAC-seq sequencing reads were aligned to hg19 reference genome using BWA (Li and Durbin, 2010), followed by filtering for uniquely mapped non-mitochondrial reads and removal of duplicates. Peak calling was performed using HOMER (Heinz et al., 2010) and the union of the peaks called in individual samples was used to calculate the ATAC-seq tag density over each peak region across all samples. Differentially accessible regions were identified using edgeR (Robinson et al., 2010) with the cutoffs of at least 1.5-fold difference and FDR < 0.01. To calculate genomic ATAC-seq tag density over 50 Kb intervals, reads in all peak regions were excluded and remaining reads were counted over 50 Kb genomic bins. This lower-resolution ATAC-seq density, unlike the density at ATAC-seq peaks, showed a high correlation with replication throughout cell cycle.

Repli-seq analysis—Repli-seq 75 bp reads in each of the surveyed cell cycle phases (S1, S2, S3, S4, ES, LS, G2/M) were mapped to hg19 reference genome, followed by the removal of duplicates and counting reads over a 50 Kb bins across the genome. These counts were then quantile normalized and LOESS smoothed as described in (Marchal et al., 2018).

Replication Timing Index (RTI) and differential RT—The traditional metric of RT, Early-to-Late (E/L) ratio, is based on two time points (early and late S-phase). We generalized this metric to a more sensitive numerical value in order to quantify RT based on four time points in our experiments, or any other number of profiled time points $N \geq 2$. The

RT index (RTI) is based on a weighted sum of normalized replication signals (Repli-seq read densities D_n) from each time point n:

$$RTI = \frac{\sum_{n=1}^{N} n\, D_n}{\sum_{n=1}^{N} D_n}$$

where n is the time point of the cell cycle (1 to 4, corresponding to time points S1 to S4) and D_n is the density of BrdU reads (per bp) within the given region at this time point. This formula is schematically explained in Figure 3A, which shows an example of a genomic region with BrdU intensities at time points S1-S4 indicated by four rows of heatmaps with a trace of the pattern within each S phase time point. RTI corresponds to the "center of mass" of the replication signal at a continuous scale between 0 (earliest replication) and 1 (latest replication). RTI faithfully recapitulates the original pattern of four replication signals at time points S1-S4 (Figure 3A and S3C). Regions of differential RT between control and KDM4A overexpressing cells were identified using the difference of RTI in a given genomic bin, with RTI difference cutoff of 0.05. This cutoff of RTI difference was largely consistent with a stringent cutoff of difference in E/L ratio (change of log2 E/L ratio > 0.3) used for calling RT change in previous publications (Rivera-Mulia et al., 2015). RTI difference cutoff of 0.05 provided the detection of 99% of all regions with > 0.3 $\log_2$ scale difference of Repli-seq density at any time point S1-S4. Compared to the previously reported approach to calling RT changes based on the cutoff of two standard deviations in pairwise comparisons of biological replicates (Rivera-Mulia et al., 2017; Sarni et al., 2020) (Figure S3F), our RTI cutoff was similar but somewhat more stringent since in the genome-wide distribution of RTI differences (Figure S3F), two standard deviations corresponded to the RTI difference of 0.044.

Classification of local replication patterns—To identify types of local replication patterns, we analyzed RTI values at 50 Kb genomic bins and surveyed RTI patterns over a window of 10 adjacent bins sliding across the genome. To represent the local shape of RT pattern in each window, RTI value in each of the 10 bins was normalized by the average RTI across window. The resulting local RT patterns across all windows were clustered using hierarchical clustering. Four most distinct clusters corresponded to the patterns of constant RT, local RTI minimum, local RTI maximum and RTI slope, similar to previously reported RT patterns (Zhao et al., 2020). These patterns corresponded to constant timing regions (CTR, defined as a window with RTI variance < 0.006 among 10 bins), initiation zones (IZ, a local minimum of RTI), termination sites (TS, a local maximum of RTI) and the remaining genomic windows classified as timing transitioning regions (TTR). We further subdivided the constant timing regions (CTRs) into initiation constant replication regions (ICTR) and termination constant replication regions (TCTR) based on RTI profiles in the flanking 500 Kb windows. Local replication domains were defined as genomic regions between a local minimum and the nearest local maximum of RTI in the genomic RTI track. Each of these domain boundary points of local start (local RTI minimum) and local end of replication (local RTI maximum) was assigned to one of four time points (S1, S2, S3, or

S4) corresponding to one of four bins of RTI ([0–0.25], [0.25–0.5], [0.5–0.75], or [0.75–1], respectively).

Hi-C analysis—Public Hi-C data in RPE cells (Darrow et al., 2016) were downloaded and normalized using Juicer (Durand et al., 2016). Compartment eigenvalues were calculated at 50 Kb resolution using *eigenvector* command in Juicer. Insulation scores were calculated as previously described (Bonev et al., 2017). Hi-C interaction frequencies were transformed into distance-specific Z-score and chromatin interaction domains were called using local normalized Hi-C contrast as previously described (Kundu et al., 2017).

Classification of primary chromatin states—To identify chromatin states based on the combination of chromatin marks, we applied systematic classification of genomic regions using the chromHMM method (Ernst and Kellis, 2012) implementing a multivariate Hidden Markov Model (HMM). We first applied a standard Roadmap Epigenomics HMM model that uses the levels of six histone marks (H3K4me3, H3K4me1, H3K27ac, H3K27me3, H3K36me3, H3K9me3) to classify genomic regions into 18 epigenomic states. Average ChIP-seq tag density of histone marks and input over cell cycle were calculated over 200 bp bins in both replicates and then converted into the binary presence or absence calls using *BinarizeBam* function of chromHMM. These binary calls for six histone marks at 200 bp resolution were applied to the 18-state HMM model publicly available from Roadmap Epigenomics [https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/core_K27ac/jointModel/final/] to call chromatin states. To call chromatin states at a lower 50 Kb resolution similar to the resolution of Repli-seq, we used input-normalized normalized ChIP-seq densities of seven broad histone marks (H3K36me1, H3K36me2, H3K36me3, H3K9me1, H3K9me2, H3K9me3, H3K27me3) to make binary presence or absence calls at 50 Kb bins across genome and then apply chromHMM classification to establish primary chromatin states. Classifying the genome into four chromatin states with distinct combinations of marks and a quiescent state with no marks present (a small minority of mostly unmappable genomic regions) produced biologically meaningful annotation that was functionally consistent with subgroups of the detailed 18 states at 200 bp resolution. To assess the genome-wide enrichment of overlaps between each of the four primary states and each of the 18 detailed states (Figure S2D), we calculated $\log_2$ enrichment $E_{ij} = log_2(f_{ij}/f_i f_j)$, where $f_i (i = 1…18)$ and $f_j (j = 1…4)$ are genomic fractions corresponding to one 18 Roadmap (i) and one of 4 primary states (j), respectively, while f_{ij} is the genomic fraction of overlap state i and state j. To visualize the association between RT patterns and other genomic functions for each 50 Kb region in each primary chromatin state (Figure 1E), we calculated RNA-seq counts per million reads CPM, the density of SINE elements, Hi-C eigenvalue based on public Hi-C data (Darrow et al., 2016), and the presence of previously annotated LADs (Guelen et al., 2008). We also calculated gene density in a 500 Kb vicinity of each bin. To evaluate the preferences for adjacency between each pair of different primary states (Figure S2E), the observed genome-wide frequency of adjacency was compared to the average random frequency based on random shuffling of these regions across the genome and presented as a heatmap of $\log_2$ fold enrichment.

Analyses of cell cycle dynamics of broad histone marks and chromatin accessibility—All 50 Kb genomic bins were separated by their primary chromatin state and then sub-clustered by the patterns of Repli-seq signal across four time points of S-phase (S1-S4) using hierarchical clustering. To investigate the corresponding temporal dynamics of chromatin and gene expression throughout cell cycle, input-normalized ChIP-seq tag density of broad histone marks, lower-resolution ATAC-seq tag density, and RNA-seq signal were calculated in each bin at each profiled time point (G1, ES, LS, G2/M) and then normalized as a log2 ratio to the average across all time points for the given bin.

Quantitative RT modeling—To analyze correlative relationships of RT in control RPE cells with histone marks, chromatin accessibility, and RNA expression, we used linear regression models implemented as function *lm* in the R *stat* package. Log_2 enrichment of broad histone marks, RNA-seq, and lower resolution ATAC-seq signal were calculated over 50 Kb genomic bins at each of the four cell cycle time points in both replicates. Regression models based on various combinations of these values were assessed for the correlation between predicted and observed RTI. In addition, we compared the models based on temporal averages of signals across cell cycle to the models based on the combination of separate values at four individual time points.

To analyze the associations of KDM4A effects on RT with the effects on histone marks, chromatin accessibility, and RNA expression, we used random forest algorithm to develop (a) regression models for predicting RT change on a continuous scale and (b) classifier models for discriminating between RT changing and non-changing regions. Both model types were implemented using *randomForest* R package with default parameters. For training and testing of these models we used leave-one-out cross-validation strategy where a randomly sampled quarter of all 50 Kb genomic bins was used as testing set after the other three quarters were used as a training set. The regression model shown in Figure 5 G used the levels of seven broad histone marks (H3K36me1-3, H3K9me1-3, H3K27me3) and their difference between control and KDM4A overexpressing cells, averaged across four surveyed time points and among two biological replicates in each 50 Kb genomic bin as input to predict quantitative changes of RTI between KDM4A-overexpressing and control cells in this bin. Adding focused histone mark densities (H3K27ac, H3K4me1-3, H3K9ac) at the loci of differential narrow peaks, ATAC-seq, and RNA-seq signals to the input or using separate values for individual cell cycle time points rather than temporal averages did not lead to substantial improvement of model's performance.

We also trained and evaluated classifier models for the binary prediction of genomic bins with RT changes, defined by the 0.05 cutoff of RTI difference between control and KDM4A overexpressing cells. As input, we used various combinations of broad histone marks, chromatin accessibility, and RNA-seq signal, as well as their changes between control and KDM4A overexpressing cells. In addition to using all 50 Kb genomic bins, we separately evaluated the models that were specifically focused on (a) gene bodies with differential RT, which was modeled using differential expression of each gene as input, and (b) ATAC-seq peaks with differential RT, which was modeled using differential ATAC-seq tag density at each peak. Performance of each classifier model was evaluated using Receiver Operating Characteristic (ROC) curve, which compares the rates of true positives and false positives

among top N predictions, with N sliding from one top prediction to the whole list of evaluated genomic regions. Model's accuracy was assessed by the corresponding area under curve (AUC).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Drs. Alfonso Bellacosa and Cihangir Duy, as well as Gulnaz Alekbaeva, Reuben Duttweiler, and Zachary Gray, for comments on the manuscript. This study was supported by NIH NIGMS Grant R01GM097360 (to J.R.W.), NIH/NCI Cancer Center Support Grant P30 CA006927 (to J.R.W.), and NIH NIDDK Grant P30DK040561 (to R.I.S.). J.R.W. is a recipient of an American Lung Association Lung Cancer Discovery Award.

REFERENCES

Alabert C, Loos C, Voelker-Albert M, Graziano S, Forné I, Reveron-Gomez N, Schuh L, Hasenauer J, Marr C, Imhof A, and Groth A (2020). Domain Model Explains Propagation Dynamics and Stability of Histone H3K27 and H3K36 Methylation Landscapes. Cell Rep. 30, 1223–1234.e8. [PubMed: 31995760]

Anders S, Pyl PT, and Huber W (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169. [PubMed: 25260700]

Black JC, Allen A, Van Rechem C, Forbes E, Longworth M, Tschöp K, Rinehart C, Quiton J, Walsh R, Smallwood A, et al. (2010). Conserved antagonism between JMJD2A/KDM4A and HP1g during cell cycle progression. Mol. Cell 40, 736–748. [PubMed: 21145482]

Black JC, Van Rechem C, and Whetstine JR (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol. Cell 48, 491–507. [PubMed: 23200123]

Black JC, Manning AL, Van Rechem C, Kim J, Ladd B, Cho J, Pineda CM, Murphy N, Daniels DL, Montagna C, et al. (2013). KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. Cell 154, 541–555. [PubMed: 23871696]

Black JC, Atabakhsh E, Kim J, Biette KM, Van Rechem C, Ladd B, Burrowes PD, Donado C, Mattoo H, Kleinstiver BP, et al. (2015). Hypoxia drives transient site-specific copy gain and drug-resistant gene expression. Genes Dev. 29, 1018–1031. [PubMed: 25995187]

Black JC, Zhang H, Kim J, Getz G, and Whetstine JR (2016). Regulation of Transient Site-specific Copy Gain by MicroRNA. J. Biol. Chem 291, 4862–4871. [PubMed: 26755726]

Blumenfeld B, Masika H, Farago M, Yehuda Y, Halaseh L, Vardi O, Rapoport R, Levin-Klein R, Cedar H, Bergman Y, and Simon I (2021). Chromosomal coordination and differential structure of asynchronous replicating regions. Nat. Commun 12, 1035. [PubMed: 33589603]

Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, and Cavalli G (2017). Multi-scale 3D Genome Rewiring during Mouse Neural Development. Cell 171, 557–572.e24. [PubMed: 29053968]

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, and Greenleaf WJ (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 10, 1213–1218. [PubMed: 24097267]

Clarke TL, Tang R, Chakraborty D, Van Rechem C, Ji F, Mishra S, Ma A, Kaniskan HU, Jin J, Lawrence MS, et al. (2020). Histone Lysine Methylation Dynamics Control *EGFR* DNA Copy-Number Amplification. Cancer Discov. 10, 306–325. [PubMed: 31776131]

Darrow EM, Huntley MH, Dudchenko O, Stamenova EK, Durand NC, Sun Z, Huang SC, Sanborn AL, Machol I, Shamim M, et al. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. Proc. Natl. Acad. Sci. USA 113, E4504–E4512. [PubMed: 27432957]

Dileep V, Rivera-Mulia JC, Sima J, and Gilbert DM (2015). Large-Scale Chromatin Structure-Function Relationships during the Cell Cycle and Development: Insights from Replication Timing. Cold Spring Harb Symp Quant Biol. 80, 53–63. [PubMed: 26590169]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, and Aiden EL (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst. 3, 95–98. [PubMed: 27467249]

Ernst J, and Kellis M (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods 9, 215–216. [PubMed: 22373907]

García-Carpizo V, Sarmentero J, Han B, Graña O, Ruiz-Llorente S, Pisano DG, Serrano M, Brooks HB, Campbell RM, and Barrero MJ (2016). NSD2 contributes to oncogenic RAS-driven transcription in lung cancer cells through long-range epigenetic activation. Sci. Rep 6, 32952. [PubMed: 27604143]

Gorkin DU, Leung D, and Ren B (2014). The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell 14, 762–775. [PubMed: 24905166]

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, and van Steensel B (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453, 948–951. [PubMed: 18463634]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. [PubMed: 20513432]

Iguchi-Ariga SM, Okazaki T, Itani T, Ogata M, Sato Y, and Ariga H (1988). An initiation site of DNA replication with transcriptional enhancer activity present upstream of the c-myc gene. EMBO J. 7, 3135–3142. [PubMed: 3053161]

Jiang XR, Jimenez G, Chang E, Frolkis M, Kusler B, Sage M, Beeche M, Bodnar AG, Wahl GM, Tlsty TD, and Chiu CP (1999). Telomerase expression in human somatic cells does not induce changes associated with a transformed phenotype. Nat. Genet 21, 111–114. [PubMed: 9916802]

Johnstone SE, Reyes A, Qi Y, Adriaens C, Hegazi E, Pelka K, Chen JH, Zou LS, Drier Y, Hecht V, et al. (2020). Large-Scale Topological Changes Restrain Malignant Progression in Colorectal Cancer. Cell 182, 1474–1489.e23. [PubMed: 32841603]

Klein KN, Zhao PA, Lyu X, Sasaki T, Bartlett DA, Singh AM, Tasan I, Zhang M, Watts LP, Hiraga SI, et al. (2021). Replication timing maintains the global epigenetic state in human cells. Science 372, 371–378. [PubMed: 33888635]

Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, et al. ; Roadmap Epigenomics Consortium (2015). Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330. [PubMed: 25693563]

Kundu S, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, Dekker J, and Kingston RE (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. Mol. Cell 65, 432–446.e5. [PubMed: 28157505]

Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, Xi Y, Park BH, Shi X, Garcia BA, et al. (2011). NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. Mol. Cell 44, 609–620. [PubMed: 22099308]

Laflamme G, and Mekhail K (2020). Biomolecular condensates as arbiters of biochemical reactions inside the nucleus. Commun. Biol 3, 773. [PubMed: 33319830]

Li H, and Durbin R (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595. [PubMed: 20080505]

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293. [PubMed: 19815776]

Marchal C, Sasaki T, Vera D, Wilson K, Sima J, Rivera-Mulia JC, Tre-villa-García C, Nogues C, Nafie E, and Gilbert DM (2018). Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. Nat. Protoc 13, 819–839. [PubMed: 29599440]

Mishra S, Van Rechem C, Pal S, Clarke TL, Chakraborty D, Mahan SD, Black JC, Murphy SE, Lawrence MS, Daniels DL, and Whetstine JR (2018). Cross-talk between Lysine-Modifying Enzymes Controls Site-Specific DNA Amplifications. Cell 174, 803–817.e16. [PubMed: 30057114]

Nathanailidou P, Taraviras S, and Lygerou Z (2020). Chromatin and Nuclear Architecture: Shaping DNA Replication in 3D. Trends Genet. 36, 967–980. [PubMed: 32713597]

Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, Lockhart DJ, Davis RW, Brewer BJ, and Fangman WL (2001). Replication dynamics of the yeast genome. Science 294, 115–121. [PubMed: 11588253]

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, and Aiden EL (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680. [PubMed: 25497547]

Reverón-Gómez N, González-Aguilera C, Stewart-Morgan KR, Petryk N, Flury V, Graziano S, Johansen JV, Jakobsen JS, Alabert C, and Groth A (2018). Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. Mol. Cell 72, 239–249.e5. [PubMed: 30146316]

Rivera-Mulia JC, Buckley Q, Sasaki T, Zimmerman J, Didier RA, Nazor K, Loring JF, Lian Z, Weissman S, Robins AJ, et al. (2015). Dynamic changes in replication timing and gene expression during lineage specification of human pluripotent stem cells. Genome Res. 25, 1091–1103. [PubMed: 26055160]

Rivera-Mulia JC, Desprat R, Trevilla-Garcia C, Cornacchia D, Schwerer H, Sasaki T, Sima J, Fells T, Studer L, Lemaitre JM, and Gilbert DM (2017). DNA replication timing alterations identify common markers between distinct progeroid diseases. Proc. Natl. Acad. Sci. USA 114, E10972–E10980. [PubMed: 29196523]

Robinson MD, McCarthy DJ, and Smyth GK (2010). edgeR: a Bio-conductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. [PubMed: 19910308]

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, and Gilbert DM (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res. 20, 761–770. [PubMed: 20430782]

Sarni D, Sasaki T, Irony Tur-Sinai M, Miron K, Rivera-Mulia JC, Magnuson B, Ljungman M, Gilbert DM, and Kerem B (2020). 3D genome organization contributes to genome instability at fragile sites. Nat. Commun 11, 3613. [PubMed: 32680994]

Siefert JC, Georgescu C, Wren JD, Koren A, and Sansam CL (2017). DNA replication timing during development anticipates transcriptional programs and parallels enhancer activation. Genome Res. 27, 1406–1416. [PubMed: 28512193]

Sima J, Chakraborty A, Dileep V, Michalski M, Klein KN, Holcomb NP, Turner JL, Paulsen MT, Rivera-Mulia JC, Trevilla-Garcia C, et al. (2019). Identifying cis Elements for Spatiotemporal Control of Mammalian DNA Replication. Cell 176, 816–830.e18. [PubMed: 30595451]

Stafford JM, Lee CH, Voigt P, Descostes N, Saldaña-Meyer R, Yu JR, Leroy G, Oksuz O, Chapman JR, Suarez F, et al. (2018). Multiple modes of PRC2 inhibition elicit global chromatin alterations in H3K27M pediatric glioma. Sci. Adv 4, eaau5935. [PubMed: 30402543]

Stewart-Morgan KR, Petryk N, and Groth A (2020). Chromatin replication and epigenetic cell memory. Nat. Cell Biol 22, 361–371. [PubMed: 32231312]

Streubel G, Watson A, Jammula SG, Scelfo A, Fitzpatrick DJ, Oliviero G, McCole R, Conway E, Glancy E, Negri GL, et al. (2018). The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. Mol. Cell 70, 371–379.e5. [PubMed: 29606589]

Van Rechem C, and Whetstine JR (2014). Examining the impact of gene variants on histone lysine methylation. Biochim. Biophys. Acta 1839, 1463–1476. [PubMed: 24859469]

Van Rechem C, Black JC, Abbas T, Allen A, Rinehart CA, Yuan GC, Dutta A, and Whetstine JR (2011). The SKP1-Cul1-F-box and leucine-rich repeat protein 4 (SCF-FbxL4) ubiquitin ligase regulates lysine demethylase 4A (KDM4A)/Jumonji domain-containing 2A (JMJD2A) protein. J. Biol. Chem 286, 30462–30470. [PubMed: 21757720]

Van Rechem C, Ji F, Mishra S, Chakraborty D, Murphy SE, Dillingham ME, Sadreyev RI, and Whetstine JR (2020). The lysine demethylase KDM4A controls the cell-cycle expression of replicative canonical histone genes. Biochim. Biophys. Acta. Gene Regul. Mech 1863, 194624. [PubMed: 32798738]

Vouzas AE, and Gilbert DM (2021). Mammalian DNA Replication Timing. Cold Spring Harb. Perspect. Biol 13, a040162. [PubMed: 33558366]

Weinberg DN, Papillon-Cavanagh S, Chen H, Yue Y, Chen X, Rajagopalan KN, Horth C, McGuire JT, Xu X, Nikbakht H, et al. (2019). The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. Nature 573, 281–286. [PubMed: 31485078]

Yokochi T, Poduch K, Ryba T, Lu J, Hiratani I, Tachibana M, Shinkai Y, and Gilbert DM (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. Proc. Natl. Acad. Sci. USA 106, 19363–19368. [PubMed: 19889976]

Yuan W, Xu M, Huang C, Liu N, Chen S, and Zhu B (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem 286, 7983–7989. [PubMed: 21239496]

Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, et al. ; Mouse ENCODE Consortium (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature 515, 355–364. [PubMed: 25409824]

Zhao PA, Sasaki T, and Gilbert DM (2020). High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. Genome Biol. 21, 76. [PubMed: 32209126]

Highlights

- Four chromatin states define 97% of genome and associate with different RT patterns
- H3K36me2 state defines 17% of the genome and largely replicates in mid-S phase
- KDM4A controls RT in 11% of the genome, especially within the H3K36me2 state
- Regulation of broad histone modifications and enhancer elements predicts RT

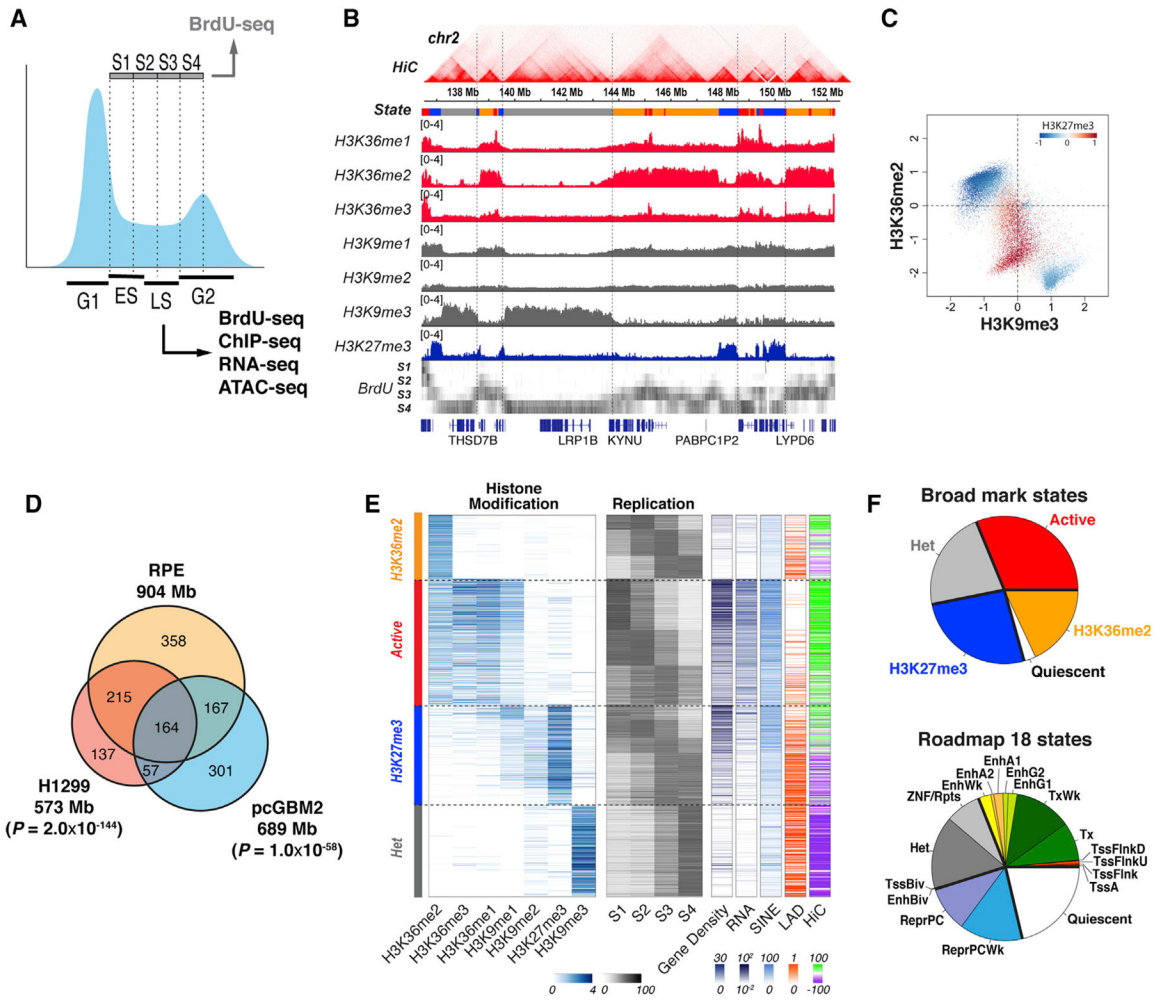


Figure 1. Four chromatin states define 97% of the mappable genome

(A) Experimental design.

(B) Representative genomic tracks of Repli-seq, ChIP-seq, Hi-C, and chromatin states.

(C) Scatterplot of $\log_2$ levels of H3K9me3 versus H3K36me2 at 50-kb bins genome-wide. H3K27me3 levels: a scale of blue and red.

(D) Statistical significance of overlap with RPE H3K36me2-enriched regions.

(E) Primary chromatin states associate with different replication timing patterns (gray: BrdU cpm), gene density and expression, repeat density, LAD overlap (Guelen et al., 2008) and spatial A/B compartments (Hi-C eigenvalue, green and purple, respectively).

(F) Primary chromatin states (top pie chart) correspond to groups of similar chromatin states in the Roadmap Epigenomics annotation (bottom pie chart, state groups separated by bold lines). The H3K36me2 state (top, orange) replaces the Roadmap quiescent state (bottom, white).

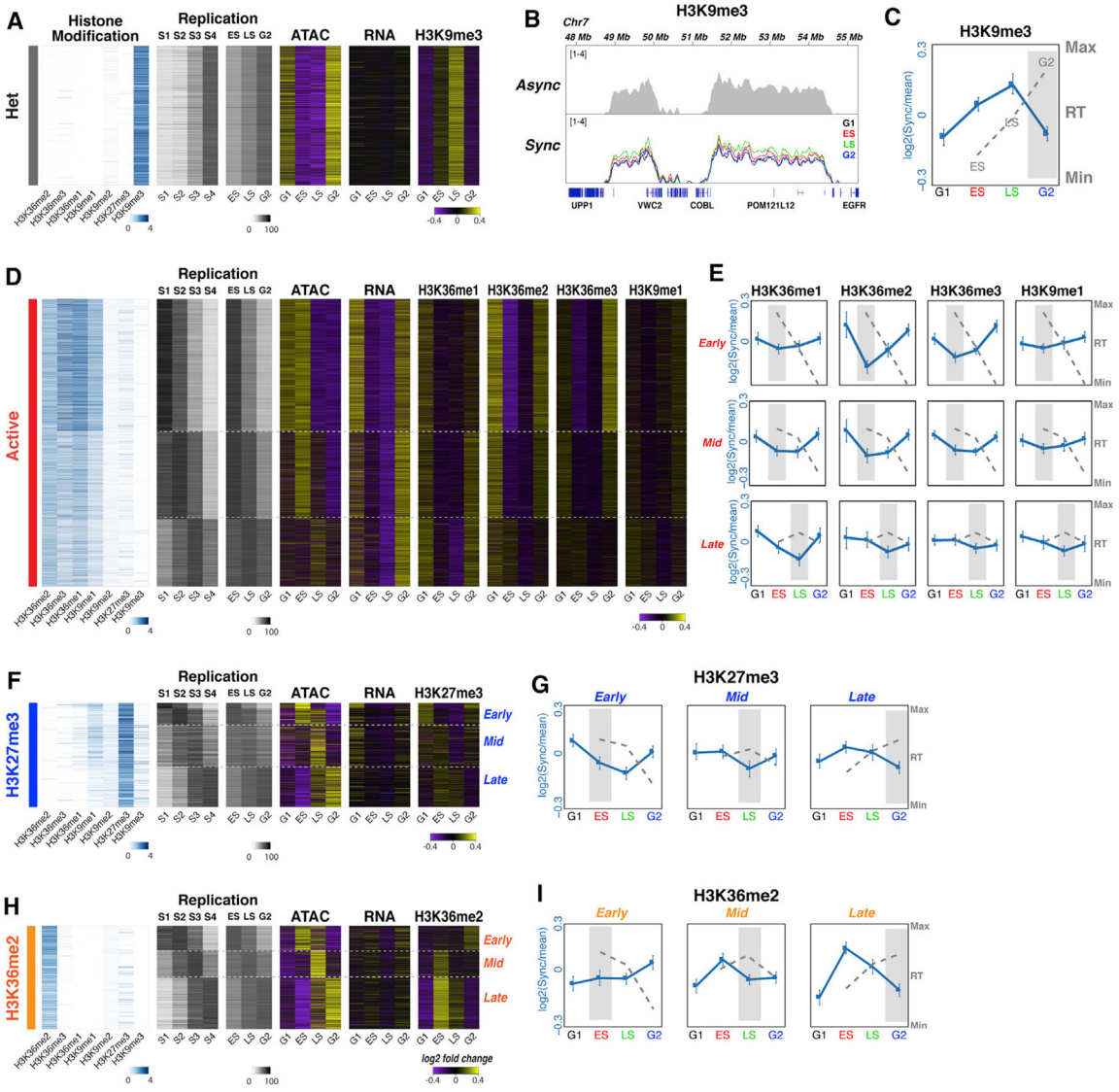


Figure 2. Methylation dynamics associate with RT within individual chromatin states

(A) Heterochromatic regions replicate late and correlate with temporal dynamics of H3K9me3 and chromatin accessibility (ATAC).

(B) H3K9me3 genomic tracks in asynchronous cells (top) and at four time points of cell cycle (bottom).

(C) The maximum of replication signal in G2 (shaded) coincides with the minimum of H3K9me3 (blue line). Average time course of H3K9me3 density relative to cell-cycle mean among regions of the heterochromatic state (blue line, $\log_2$ scale) compared with average relative Repli-seq levels (gray line).

(D) Active chromatin states present multiple replication patterns (gray: BrdU cpm) correlating with temporal dynamics of chromatin marks and accessibility (ATAC).

(E) Average levels of histone mark density relative to cell-cycle mean (blue line, $\log_2$ scale) compared with average relative Repli-seq levels (gray line). Active chromatin state.

(F and G) H3K27me3 state (similar to D and E, respectively).

(H and I) H3K36me2 state (similar to D and E, respectively).
Purple-yellow heatmap colors indicate signal densities at 50-kb genomic bins relative to cell-cycle average (log_2 scale). Error bars represent SD.

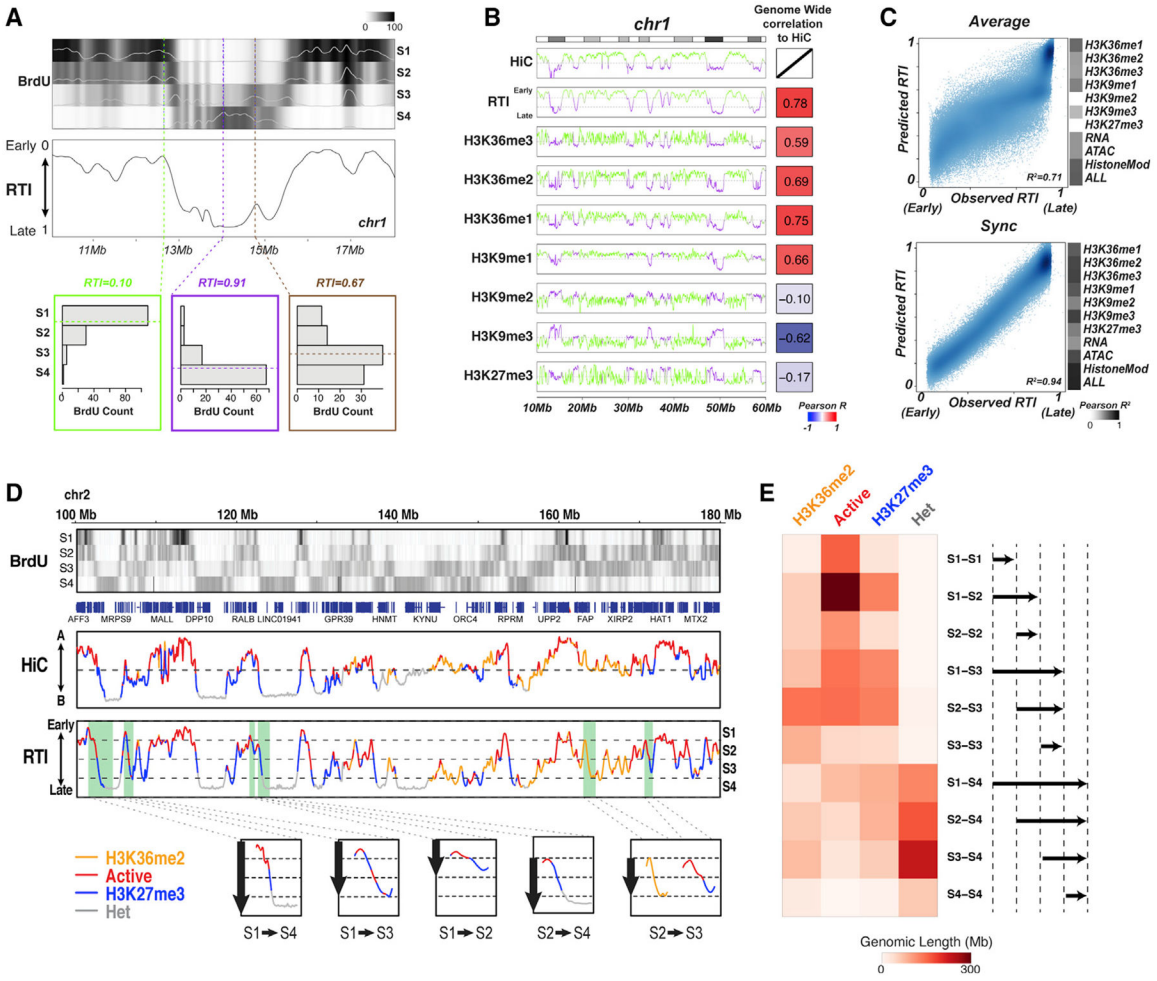


Figure 3. Methylation dynamics predict RT across the genome, and chromatin states have preferential RT

(A) Representative genomic tracks with examples of S1–S4 Repli-seq signals (bar plots) and RTI for early (green), late (purple), and mid-S-phase (brown) replicating loci.

(B) Quantitative tracks of Hi-C eigenvalues, RTI, and histone marks across a chromosomal arm. Green: compartment A; purple: compartment B.

(C) Scatterplots of RTI at all genomic bins (x axis) versus RTI predicted by two multivariate models (y axis) based on the histone marks, ATAC-seq, and RNA-seq signals averaged across cell cycle (top, $R^2 = 0.71$) and surveyed at four time points (bottom, $R^2 = 0.94$). Shading on the right indicates correlations.

(D) Upper: representative genomic tracks of Repli-seq (BrdU), Hi-C compartment eigenvalues (Hi-C), and RTI colored by chromatin states of loci. Lower: typical patterns of traversed chromatin states in local neighborhoods with different replication time spans.

(E) Chromatin-state propensities (total lengths) in neighborhoods grouped by the time span of their complete replication (right).

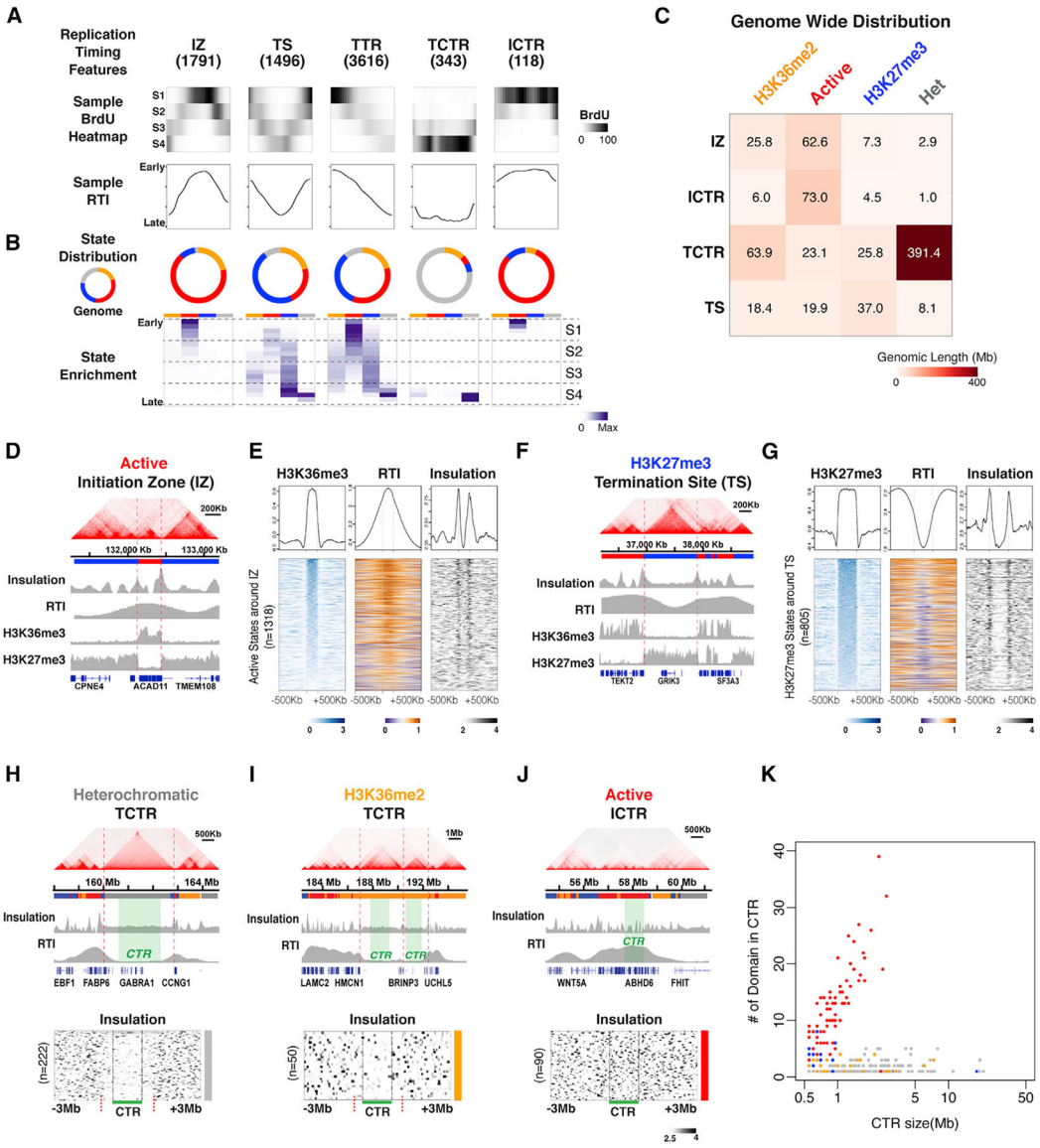


Figure 4. Local replication patterns associate with chromatin states and structure

(A) Local replication patterns: initiation zones (IZs), termination sites (TSs), timing transition regions (TTRs), termination constant timing regions (TCTRs), and initiation constant timing regions (ICTRs), with representative BrdU and RTI tracks.

(B) Doughnut plots of genomic lengths among four chromatin states (top) and enrichment of replication across S phase for each state (bottom). Overall genomic distribution of states is shown on the left.

(C) Heatmap of genomic lengths occupied by each chromatin state within each replication pattern.

(D–G) Example of genomic tracks of Hi-C, chromatin insulation score, RTI, and histone mark densities around an IZ in active state (D) and around a TS in H3K27me3 state (F). Dotted lines indicate interaction domain. Heatmaps and average profiles of H3K36me3, RTI,

and insulation score in a 500-kb vicinity of all IZs in active chromatin state (E) and all TSs in H3K27me3 state (G).

(H–J) Top: examples of genomic tracks of Hi-C, insulation score, and RTI around constant timing regions (CTR, shaded green): heterochromatic TCTR (H), H3K36me2 TCTR (I), and active-state ICTR (J). Bottom: heatmaps of insulation scores in a 3-Mb vicinity of all individual CTRs.

(K) Scatterplot of CTR size versus the number of overlapping Hi-C domains, colored by CTR's main chromatin state.

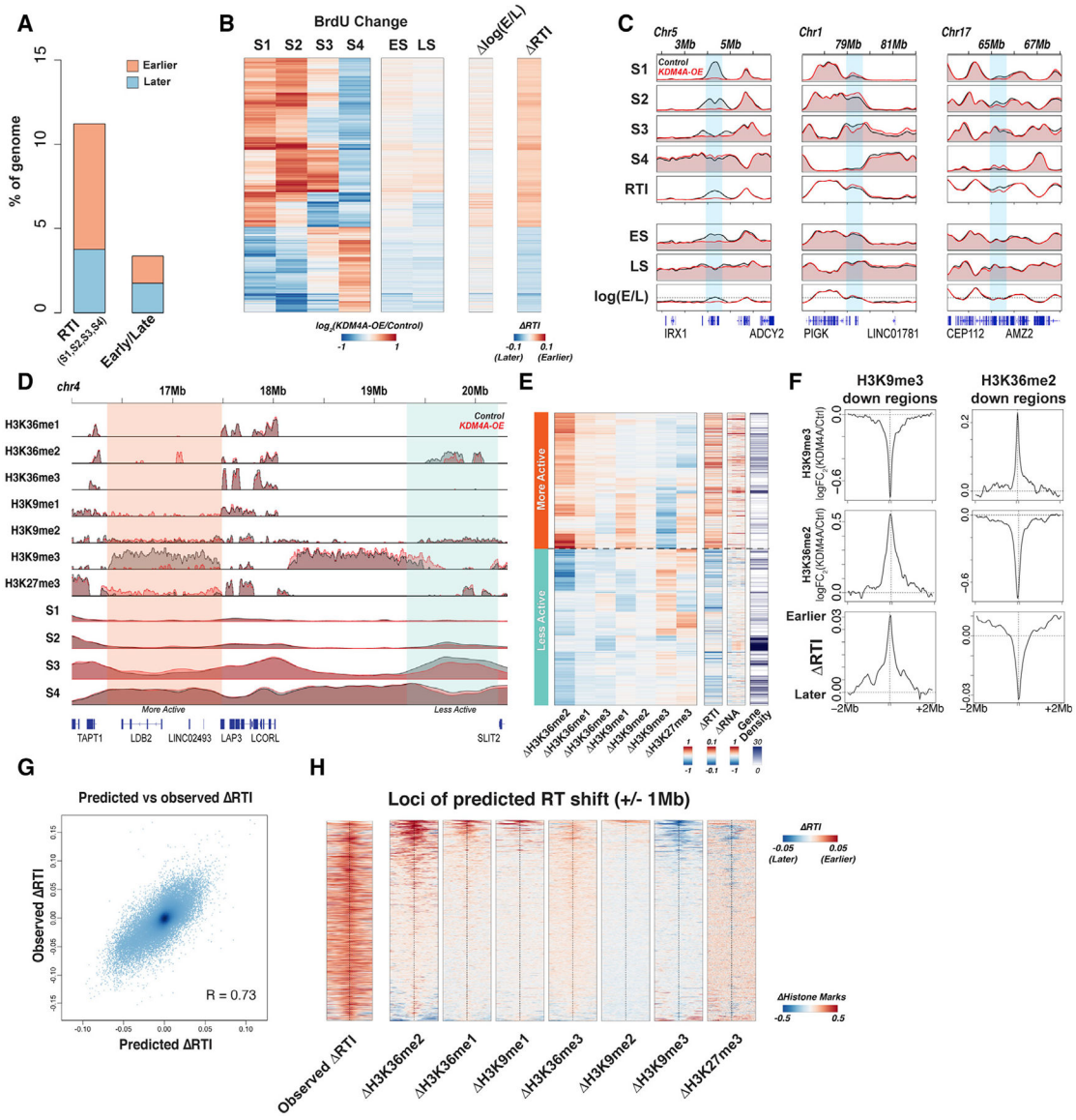


Figure 5. KDM4A controls RT in association with broad histone modifications

(A and B) Fractions of the genome shifting to earlier (orange) and later RT (blue) detected from four (RTI) or two (early/late) time points (A) and heatmaps of $\log_2$ fold change in BrdU signal (B) between control and KDM4A-overexpressing cells.

(C) Examples of genomic tracks where differences in replication signal were detected at four-point but not early/late resolution (middle, right) and where differences were detected by both (left) (black: control; red: KDM4A OE).

(D) Examples of correlated changes in RT and broad histone marks H3K36me1-3, H3K9me1-3, and H3K27me3 upon KDM4A overexpression (black: control; red: KDM4A OE).

(E) Heatmap of $\log_2$ changes of histone marks compared with changes in RTI and RNA expression over 50-kb genomic intervals. Gene density is represented on the right.

(F) Average positional profiles of histone mark and RTI changes in the vicinity of regions with H3K9me3 (left) or H3K36me2 decrease (right).
(G) Scatterplot of observed versus predicted RTI based on levels and changes of broad histone marks.
(H) Heatmaps of concomitant RT and broad histone marks changes in the vicinity of all regions predicted as changing to earlier RT by the computational model (G).

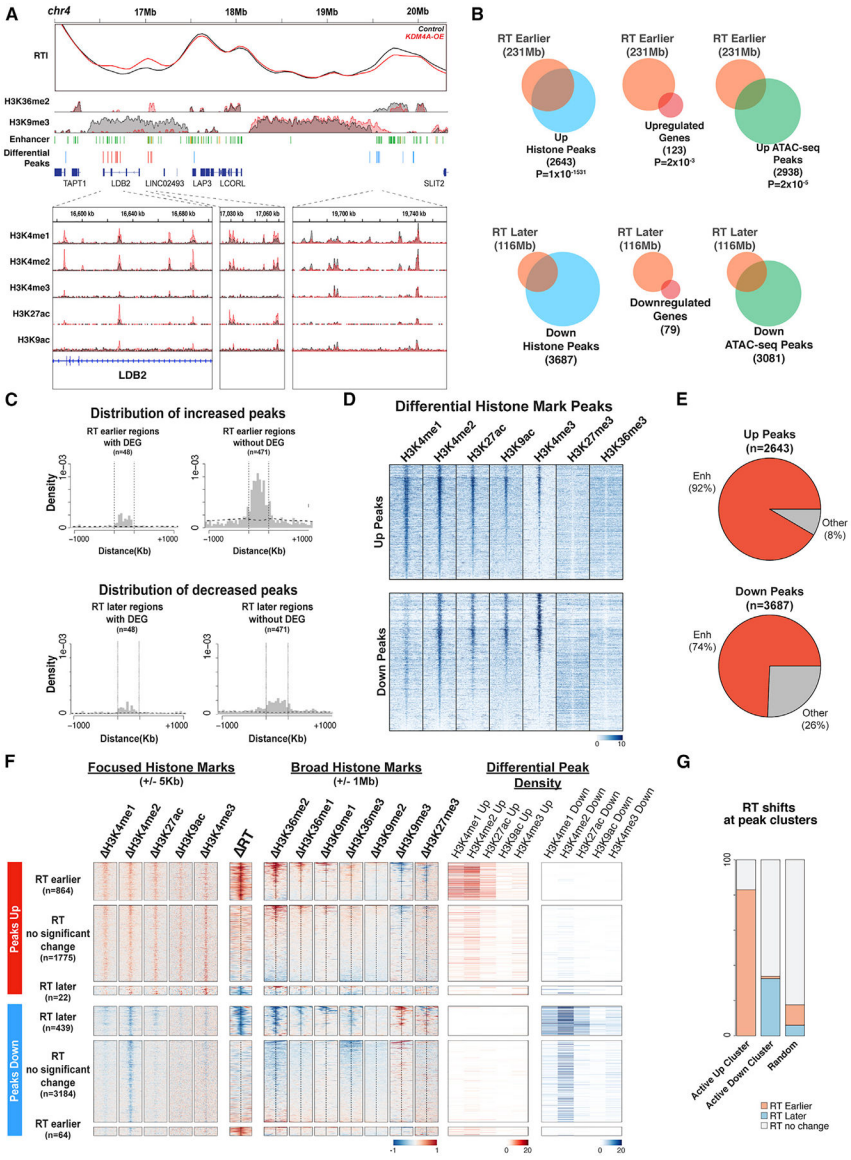


Figure 6. KDM4A-driven RT changes accompany enhancer element modification dynamics

(A) Changes of peaks of focused histone marks H3K4me1-3, H3K27ac, and H3K9ac correlate with changes in RT and broad histone marks upon KDM4A overexpression (black: control; red: KDM4A OE).

(B) Venn diagrams of regions with earlier (top) and later RT shifts (bottom) versus differential histone marks (blue), ATAC-seq peaks (green), and differentially expressed genes (red). p values were calculated based on the distribution of overlap lengths between randomly shuffled regions (see STAR Methods).

(C) Distributions of increased and decreased peaks near regions with shifts to earlier (top) and later RT (bottom), respectively, that have (left) and do not have (right) a differentially expressed gene in a 1-Mb vicinity.

(D) Heatmaps of histone mark densities in a vicinity of all peaks that were increased (top) or decreased (bottom) upon KDM4A overexpression.

(E) Loci with KDM4A-induced changes of focused histone marks are enhancers according to chromHMM classifier: up (top Venn diagram) and down (bottom Venn diagram) regulated peaks.

(F) Increase (top) and decrease of focused histone marks (bottom) accompanying changes of broad histone marks associated with earlier and later RT shifts, respectively. Higher density of adjacent differential peaks corresponds to stronger RT changes (right).

(G) The bar plots represent the distributions of earlier, later, and no RT shifts among all clusters of increasing and decreasing peaks, compared with random distribution.

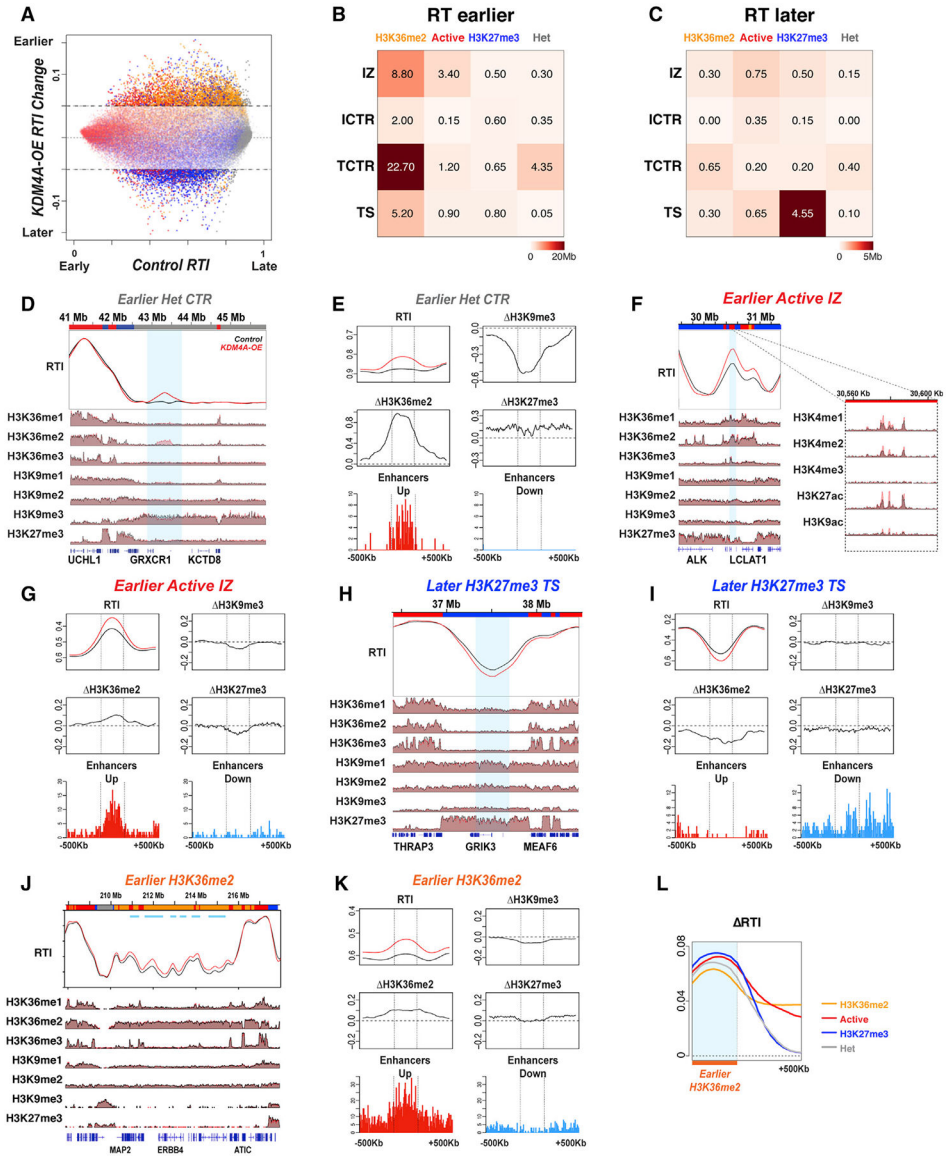


Figure 7. KDM4A-driven RT changes depend on both chromatin state and local replication pattern with enhancer modification modulation

(A) Scatterplot of RTI in control (x axis) versus KDM4A-induced RTI change (y axis) at all 50-kb genomic regions as points colored by chromatin states. Horizontal dashed lines: RTI change cutoffs.

(B and C) Heatmaps of genomic lengths of replication patterns in each chromatin state that shift to an earlier (B) and later RT (C).

(D) Genomic tracks of concomitant changes of RTI and broad histone marks around a region shifting RT (shaded blue).

(E) Average profiles of RT and changes of broad histone marks H3K9me3, H3K36me2, and H3K27me3 in a 500-kb vicinity of regions that shifted RT. Bottom panels: distributions of enhancer elements in this vicinity that increased (red) or decreased (blue) focused active histone marks. All regions that shifted to earlier RT within a broader heterochromatic CTR.

(F) Same as (D): RT shifting earlier around an IZ in active chromatin state.

(G) Same as (E): all IZs in active chromatin state that shifted to earlier RT.
(H) Same as (D): RT shifting later around a TS in H3K27me3 state.
(I) Same as (E): all TSs in H3K27me3 state that shifted to later replication.
(J) Same as (D): RT shifting earlier over an extremely wide H3K36me2-enriched neighborhood.
(K) Same as (E): all loci in H3K36me2 state that shifted to earlier RT.
(L) Average profiles of RT shift (DRTI) across all boundaries between regions in H3K36me2 state with significant RT shift (earlier H3K36me2, shaded blue) and adjacent regions, grouped by their chromatin state (lines colored by state). Black lines are control, and red lines are KDM4A OE.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti bodies		
Anti-H3K4me1	Abcam	Cat# ab8895; RRID: AB_306847
Anti-H3K4me2	Abcam	Cat# ab32356; RRID: AB_732924
Anti-H3K4me3	Millipore	Cat# 07–473; RRID: AB_1977252
Anti-H3K27Ac	Active Motif	Cat# 39133; RRID: AB_2561016
Anti-H3K9Ac	Abcam	Cat# ab4441; RRID: AB_2118292
Anti-H3K27me3	Millipore	Cat# 07–449; RRID: AB_310624
Anti-H3K9me3	Abcam	Cat# ab8898; RRID: AB_306848
Anti-H3K36me3	Abcam	Cat# ab9050; RRID: AB_306966
Anti-H3K9me1	Abcam	Cat# ab8896; RRID: AB_732929
Anti-H3K9me2	Abcam	Cat# ab1220; RRID: AB_449854
Anti-H3K36me1	Cell Signaling Technology	Cat# 14111; RRID: AB_2798395
Anti-H3K36me2	Abcam	Cat# ab9049; RRID: AB_1280939
Anti-BrdU	BD PharMingen	Cat# 555627
Chemicals, peptides, and recombinant proteins		
Dulbecco's Modified Eagle's Medium - High Glucose	Sigma Aldrich	Cat# D5648
FBS	GIBCO	Cat# 26140–079
Penicillin-Streptomycin	Life Technologies	Cat# 15140122
L-Glutamine	Life Technologies	Cat# 25030–081
Hoechst 33342	ThermoFisher Scientific	Cat# H3570
Trypsin-0.25% EDTA	Life Technologies	Cat# 2520056
Formaldehyde	Electron Microscopy Sciences	Cat# 15686
Ribonuclease A	Sigma	Cat# R4875
Proteinase K	Sigma	Cat# P6556
5-Bromo-2-deoxyuridine	Sigma	Cat# B5002
QIAzol	QIAGEN	Cat# 79306
AMPure XP	Beckman Coulter	Cat# A63881
Dynabeads Protein A	ThermoFisher Scientific	*Cat# 10001D*
Dynabeads Protein G	ThermoFisher Scientific	*Cat# 10003D*
Critical commercial assays		
miRNeasy Mini Kit	QIAGEN	Cat# 217004
TruSeq Stranded Total RNA Library Prep Gold	Illumina	Cat# 20020598
TruSeq ChIP Library Preparation Kit Set A	Illumina	Cat# IP-202–1012
TruSeq ChIP Library Preparation Kit Set B	Illumina	Cat# IP-202–1024
Nextera DNA	Illumina	Cat# FC-121–1030
MinElute PCR Purification Kit	QIAGEN	Cat# 28004
NEBNext DNA Library Prep	NEB	E6040S
Wizard SV Gel and PCR Clean-Up System	Promega	Cat# A9282
Accel-NGS Methyl-Seq DNA Library Kit	Swift Biosciences	Cat# 30024

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw sequencing data: Repli-seq, RNA-seq, ATAC-seq, ChIP-seq	This paper	GEO: GSE175752
Processed data and code for the calculation and analysis of Replication Timing Index (RTI) and the implementation of machine learning models for RTI prediction from the combined levels of histone marks	This paper	https://doi.org/10.5281/zenodo.5173191
Experimental models: Cell lines		
hTERT RPE-1	Nicholas Dyson's lab, Massachusetts General Hospital	N/A
RPE-GFP-CTRL	This paper	N/A
RPE-GFP-KDM4A	This paper	N/A
Oligonucleotides		
ATAC-seq primers: Table S1	Buenrostro et al., 2013	N/A
Recombinant DNA		
MSCV-GFP-CTRL	Black et al., 2010	N/A
MSCV-GFP-KDM4A	Black et al., 2010	N/A
Software and algorithms		
BWA	Li and Durbin, 2010	http://bio-bwa.sourceforge.net
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
HTseq	Anders et al., 2015	https://htseq.readthedocs.io/en/release_0.11.1/index.html#
EdgeR	Robinson et al., 2010	https://www.bioconductor.org/packages/release/bioc/html/edgeR.html
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/
chromHMM	Ernst and Kellis, 2012	http://compbio.mit.edu/ChromHMM/