

Parallel analysis of replication timing, gene expression, and copy number with PARTAGE

Lakshana Sruthi Sadu Murari,^{1,5} Quinn Dickinson,^{1,5} Silvia Meyer-Nava,¹ and Juan Carlos Rivera-Mulia^{1,2,3,4}

¹Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA; ²Masonic Cancer Center, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA; ³Stem Cell Institute, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA; ⁴Masonic Institute on the Biology of Aging and Metabolism, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA

The human genome is partitioned into functional compartments that replicate at specific times during S-phase. This temporal program, referred to as replication timing (RT), is coregulated with the 3D genome organization, is cell type-specific, and changes during development in coordination with gene expression. Moreover, RT alterations are linked to abnormal gene expression, genome instability, and structural variation in multiple diseases, including cancer. However, mechanistic links between RT, large-scale 3D genome architecture, and transcriptional regulation remain poorly understood. A major limitation is that current approaches require the separate profiling of RT and transcriptomes from independent batches of samples, which obscures the complex coregulation between the epigenome and transcriptome. Here, we present *PARTAGE*, a multiomics approach that enables joint profiling of copy number variation (CNV), RT, and gene expression from the same sample, providing a more accurate integrated view of the complex relationships between RT and gene regulation.

[Supplemental material is available for this article.]

Human cells duplicate their genome by the firing of thousands of origins that are activated in clusters following a precise temporal order (Rivera-Mulia and Gilbert 2016a; Rhind 2022; Vouzas and Gilbert 2023; Hyrien et al. 2025). This replication timing (RT) program partitions the genome into functional units that segregate to distinct nuclear compartments (Rivera-Mulia and Gilbert 2016b; Liu et al. 2025). In fact, RT aligns with the 3D genome organization measured by chromosome conformation methods (Dekker et al. 2002; Lieberman-Aiden et al. 2009), with active and inactive compartments replicating early and late, respectively (Ryba et al. 2010; Dixon et al. 2012; Rivera-Mulia and Gilbert 2016b). Establishment and maintenance of chromatin epigenetic states also require proper RT regulation (Lande-Diner et al. 2009; Klein et al. 2021). RT also changes dynamically during development and is cell type-specific (Hiratani et al. 2008, 2010; Rivera-Mulia et al. 2015, 2019a; Nakatani 2025). Indeed, around half of the human genome changes RT during differentiation into distinct cell fate specification pathways, and these changes occur in coordination with the establishment of cell type-specific transcriptional programs (Rivera-Mulia et al. 2015). Moreover, abnormal RT is linked to altered gene regulation in distinct disease states (Rivera-Mulia et al. 2017; Sasaki et al. 2017; Dixon et al. 2018; Rivera-Mulia et al. 2019b, 2022), highlighting the importance of RT control. The correlation between RT and transcriptional activity has been reported for decades, but the mechanistic bases remain elusive (Gilbert 2002; Hiratani et al. 2009; Rivera-Mulia and Gilbert 2016b; Vouzas and Gilbert 2023). One limitation is the lack of methods enabling the simultaneous measurement of RT and gene expression from the same samples. Although multiple RT profiling meth-

ods exist at the ensemble population level (Marchal et al. 2018; Hulke et al. 2020; Rivera-Mulia et al. 2022; Wheeler et al. 2025), as well as at the single-cell level (Dileep and Gilbert 2018; Miura et al. 2020; Bartlett et al. 2022; Gnan et al. 2022; Massey and Koren 2022), links to gene regulation are still being explored by transcriptome analyses on separate sample batches, hindering the complex coregulation of RT and gene expression.

The rapid development of multiomic methods provided tools for the assessment of cell identity and cell-to-cell heterogeneity. Multiomic, single-cell methods also inform on distinct aspects of the epigenome in parallel with gene expression (Cao et al. 2018; Zhu et al. 2019; Flynn et al. 2023). Approaches that map features of 3D genome architecture and transcriptomes have been developed (Cao et al. 2018; Chen et al. 2019; Liu et al. 2023). However, to date, there is no robust method for the joint mapping of RT and gene expression. In addition, distinct challenges remain to align 3D genome features to gene regulation due to data sparsity, and compartment identification is feasible only with data aggregation (Zhang et al. 2022). Here, we developed *PARTAGE*, a multiomic approach for the simultaneous profiling of RT and gene expression. This approach is based on sample preparation under conditions that enable the preservation of both nucleic acids, enabling joint profiling of RT and transcriptome programs. Because this method is performed on asynchronous cultures of cells maintained in regular media conditions, it eliminates the need for cell cycle synchronization, allowing a more accurate measurement of gene regulation. Moreover, the precise collection of cell populations across distinct stages of the cell cycle allows *PARTAGE* to also provide accurate measurements of CNV. Overall, *PARTAGE* is a robust approach for the simultaneous analysis of CNV, RT, and gene expression.

⁵These authors contributed equally to this work.

Corresponding author: riveramj@umn.edu

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.281532.125>. Freely available online through the *Genome Research* Open Access option.

© 2026 Sadu Murari et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Results

Overview of PARTAGE experimental design

PARTAGE is a multiomics method designed to accurately map CNV, RT, and transcriptome programs from the same samples. It is based on the sample collection for cell cycle analysis under conditions that preserve both nucleic acids. First, cells are labeled with the nucleotide analog BrdU to label nascent DNA and enable downstream purification by immunoprecipitation; then, single-cell suspensions are obtained and intact nuclei purified (Fig. 1). To enable accurate cell cycle analyses and cell sorting into specific populations while preserving both nucleic acids, intact nuclei are stained with cell-permeant DNA dyes. Specifically, we exploited cell-permeant DyeCycle dyes (Invitrogen), which we previously validated for Repli-seq sample preparation (Meyer-Nava et al. 2023). Then, cell cycle analyses are performed based on DNA content, and fractions of nuclei at distinct stages of the cell cycle are collected by fluorescence-activated cell sorting (FACS). Next, genomic DNA and nuclear RNA are copurified using silica-based strategies. Genomic DNA is processed following the optimized Repli-seq method, consisting in enzymatic fragmentation, immunoprecipitation with antibodies against BrdU, and library preparation (Rivera-Mulia et al. 2022). Nuclear RNA is also processed by standard RNA-seq. Because PARTAGE enables the isolation of a pure population of nuclei in G1, an accurate CNV mapping can be performed. DNA isolated from S-phase fractions is exploited to measure RT, and the RNA enables accurate detection of transcriptome control during the cell cycle (Fig. 1).

PARTAGE sample preparation preserves DNA and RNA integrity

To test PARTAGE, we exploited the HepG2 human cell line, a reference cell line used at the ENCODE (The ENCODE Project Consortium et al. 2020) and Roadmap Epigenomics (Roadmap Epigenomics Consortium et al. 2015) consortia with extensive

omics data publicly available. Cells were grown up to 60% confluence to avoid contact inhibition, which halts normal cell cycle proliferation, then, dissociated into a single-cell suspension, counted, and assessed for viability (Fig. 2A; Supplemental Fig. S1). Then, intact nuclei were isolated by detergent lysis (Fig. 2B; Supplemental Fig. S1) and DNA stained for cell cycle analysis (Fig. 2C) as previously described (Meyer-Nava et al. 2023). Purified nuclei were analyzed by flow cytometry, and the cell cycle was estimated based on DNA content (Vybrant Violet DyeCycle intensity). Flow cytometry gating strategies, and identification of cell cycle fractions are shown in Figure 2D. Populations of 20 K nuclei for each fraction derived from G1, S-phase (divided into four fractions), and G2/M were sorted into multiwell plates containing a nucleic acid stabilization solution to preserve integrity. Then, high-quality total genomic DNA and RNA were copurified using silica-based column methods. PARTAGE sample preparation generated nucleic acid samples of high quality and integrity adequate for downstream processing by optimized Repli-seq and total nuclear RNA-seq (Fig. 2E–J; Supplemental Figs. S2, S3).

PARTAGE simultaneously maps CNV, RT, and transcriptomes

To evaluate PARTAGE's capacity to capture the multimodal genomic signals from the same samples, we analyzed HepG2 cells and compared the results against those generated using standard Repli-seq and RNA-seq methods (Fig. 3). Standard Repli-seq was performed on HepG2 nuclei as previously described (Rivera-Mulia et al. 2022). For standard RNA-seq, RNA was isolated from 0.5 million cells, and libraries were prepared for stranded mRNA sequencing. These standard analyses confirmed the close alignment of RT programs with genome compartmentalization as measured by Hi-C, as well as enrichment of active gene expression in active-early replicating genomic compartments (Fig. 3A). However, these data sets were generated independently from separate batches of samples processed by the standard methods of Repli-seq, Hi-C, and RNA-seq (Fig. 3A). Thus, we performed

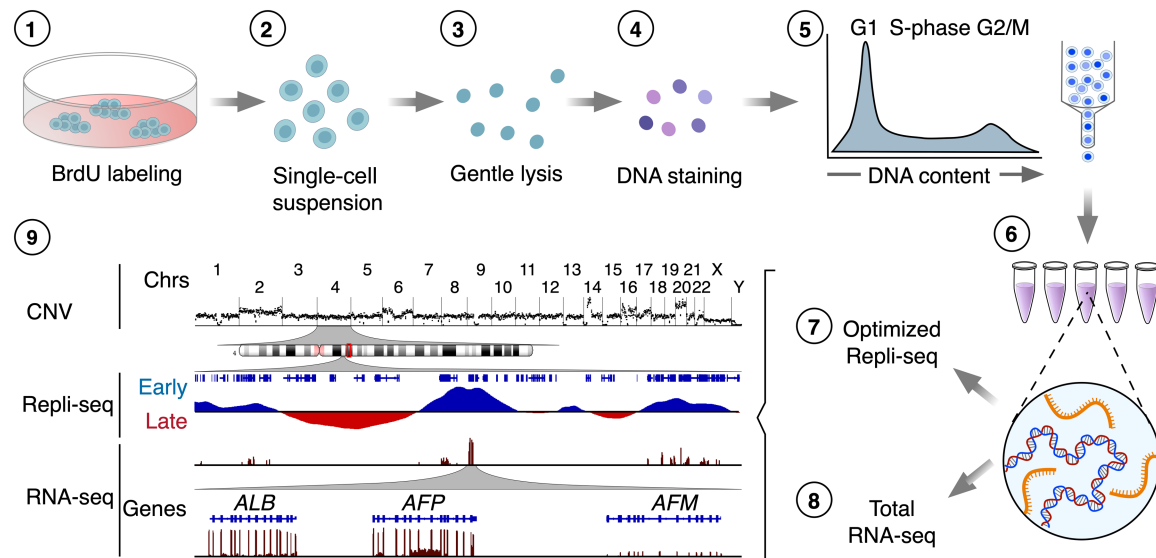


Figure 1. PARTAGE: Parallel analysis of RT and gene expression. Multiomics approach for the simultaneous profiling of CNV, RT, and transcriptome programs. Target cells are BrdU-labeled (1), dissociated (2), and intact nuclei isolated (3). Then, DNA was stained with cell-permeant dyes (4) and cell cycle analyzed by flow cytometry (5). Nuclei at specific cell cycle stages were collected for nucleic acid copurification (6), and optimized Repli-seq (7) and total RNA-seq (8) were performed. Finally, data analysis and visualization were performed to generate CNV, RT, and gene expression profiles (9).

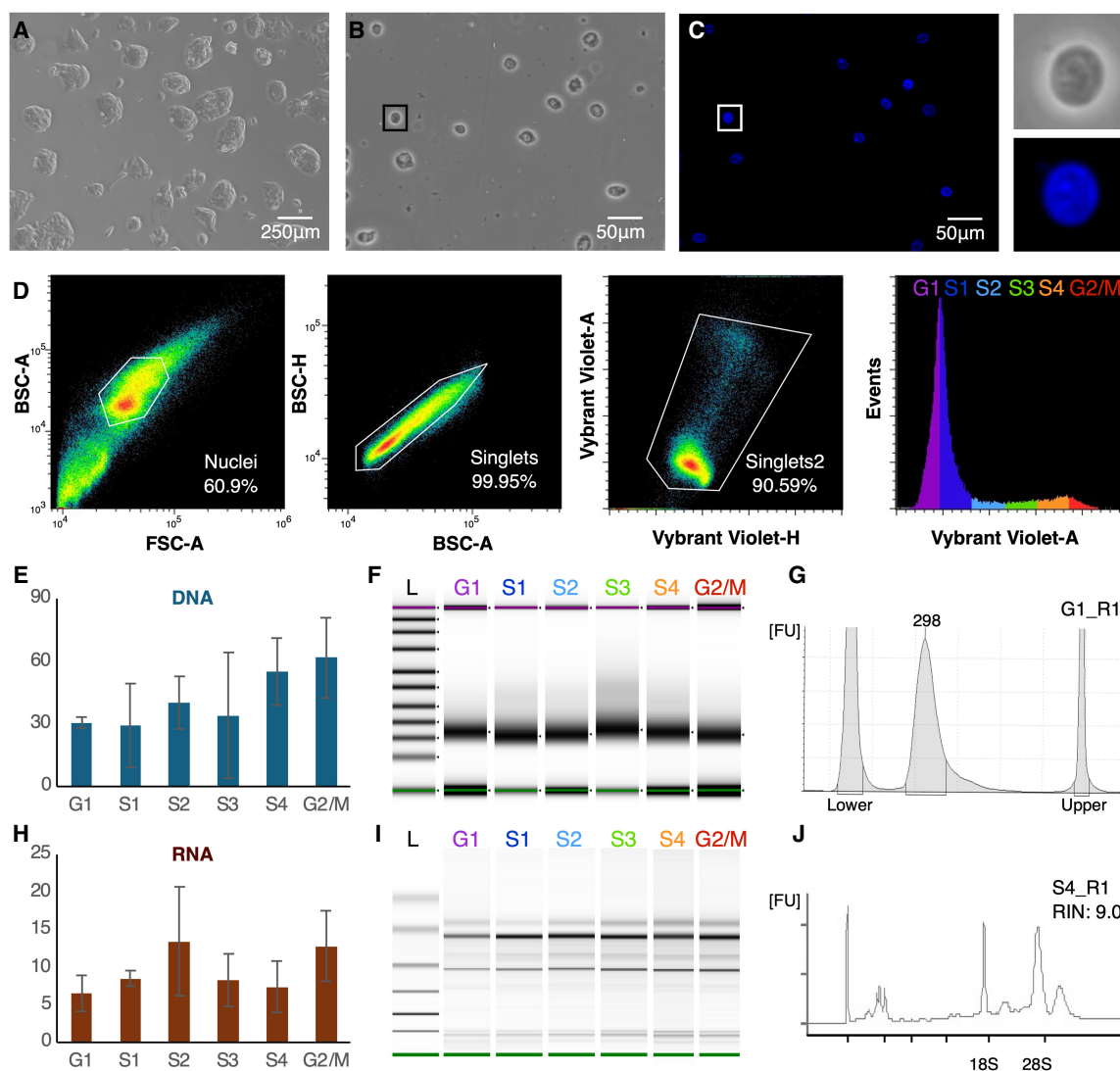


Figure 2. PARTAGE sample preparation. (A) HepG2 cells in culture. (B) Purified intact nuclei. (C) Nuclei staining with Vybrant DyeCycle Violet for accurate DNA content estimation. Magnification of a single nucleus is shown. (D) Flow cytometry and gating strategy. Single nuclei were identified, and gates were established to purify cell populations across the cell cycle: G1, S1, S2, S3, S4, G2/M. Twenty thousand nuclei were collected per fraction, and three independent biological replicates were processed. (E) DNA yields (total ng) obtained from each sorted population. (F) Integrity analysis and size distribution of purified and fragmented genomic DNA. (G) Electropherogram of purified and fragmented genomic DNA. (H) RNA yields copurified with the DNA (total ng) from the sorted cell populations. (I) RNA integrity analysis. (J) Exemplary electropherogram of purified RNA demonstrating RNA integrity (RIN = 0.90).

PARTAGE by processing the purified genomic DNA using the optimized Repli-seq method and the RNA using an ultralow input library preparation pipeline. PARTAGE allowed us to simultaneously map CNV, RT, and gene expression from the same sample (Fig. 3B). First, whole-genome sequencing of the cell populations isolated from G1 enabled the accurate mapping of CNV and the detection of multiple amplifications in the HepG2 genome (Fig. 3B). Next, RT was mapped using the multifraction Repli-seq method (Hansen et al. 2010; Zhao et al. 2020). Standard Repli-seq involves the sorting of early and late fractions of S-phase cells/nuclei, and the generation of binary profiles of early versus late replicating values across the genome. In contrast, multifraction Repli-seq involves cell sorting of multiple fractions across the entire S-phase to increase resolution and fine-scale detection of genomic features of the replication domains (Zhao

et al. 2020). Thus, PARTAGE multifraction Repli-seq was performed and data visualized as heat maps of normalized sequencing signals per genomic bin. We found that the patterns of early/late replication reproduced the RT timing data obtained with standard Repli-seq (Fig. 3B). Finally, PARTAGE total-nuclear RNA-seq reproduced the transcriptome of the HepG2 cells obtained from separate RNA-seq analysis (Fig. 3). To display global trends in gene expression, we displayed exemplary patterns in three distinct chromosomal regions. To confirm cross-protocol consistency in transcriptomes, we displayed zoomed-in tracks focused on hepatic-specific and housekeeping genes (Fig. 3B).

To validate reproducibility and accuracy, we performed in-depth comparisons of PARTAGE multiomic signals against those generated by the separate standard methods. First, because the HepG2 genome contains well-characterized amplifications and

deletions, and public CNV data is available; we benchmarked PARTAGE-derived CNVs against ENCODE whole-genome sequencing (WGS) CNV data (The ENCODE Project Consortium 2012). We applied CNVpytor (Suvakov et al. 2021) to the high-coverage ENCODE WGS dataset (60×) to define CNV segments, identifying 25 CNV events with resolution of 1-Mb bins. We then applied the same CNVpytor pipeline to three independent PARTAGE libraries (~0.5× coverage each) and compared the CNVs detected in PARTAGE to those identified in the ENCODE reference (Fig. 4A). Across the three PARTAGE replicates, we recovered between 92% and 100% of the ENCODE CNV events (Fig. 4B). The statistical significance of the overlap between ENCODE and PARTAGE CNVs was assessed using ProOvErlap (Gualandi et al. 2025), confirming that the concordance is substantially greater than expected by chance (Fig. 4B). Analysis at higher resolution (200-kb bins) resulted in 75 CNVs in the ENCODE reference data, with PARTAGE detecting >85% of these events (Supplemental Fig. S4). Thus, all major genomic amplifications observed in the ENCODE data can be detected using PARTAGE.

Next, we compared PARTAGE RT programs versus standard Repli-seq, and visual inspection confirmed that the multifraction PARTAGE Repli-seq signals faithfully recapitulated the patterns of the standard Repli-seq (Fig. 4C). To enable direct comparison of these different Repli-seq methods, we collapsed the PARTAGE multifraction Repli-seq into $\log_2(\text{Early/Late})$ ratios. Visual inspec-

tion overlaying the RT profiles from the distinct methods and replicates demonstrated the consistency between PARTAGE and standard Repli-seq (Fig. 4C). Moreover, cross-protocol scatterplot comparisons confirmed close concordance of RT values across the genome (Fig. 4D). Furthermore, pairwise/genome-wide correlation analysis confirmed high consistency between PARTAGE and standard Repli-seq RT profiles with $r \geq 0.95$ (Fig. 4D,E). Finally, we also validated the concordance between PARTAGE and standard RNA-seq pipelines to ensure that accurate transcriptome profiling was achieved in this method. Visualization of RNA-seq raw coverage tracks (normalized to sequencing depth) demonstrated that PARTAGE recapitulates the transcriptome data obtained from standard methods at specific genes (Figs. 3B, 4F). Because transcriptomes were obtained from distinct cell numbers (0.5 million cells for standard RNA-seq vs. 20,000 nuclei in PARTAGE RNA-seq) and using distinct library preparation pipelines (TruSeq stranded RNA-seq vs. Takara ultra-low input SMARTer Stranded Total RNA-seq), we scaled and normalized the data for proper comparison (Supplemental Fig. S5). Then, $\log_2\text{TPM}$ RNA-seq values were median-centered and batch-corrected (Johnson et al. 2007) to enable proper comparisons. We found strong gene-wise concordance between PARTAGE and standard RNA-seq, and replicate-level scatterplots of gene expression values showed near-isometric agreement, with housekeeping and hepatic-specific marker genes aligning closely (Fig. 4G; Supplemental

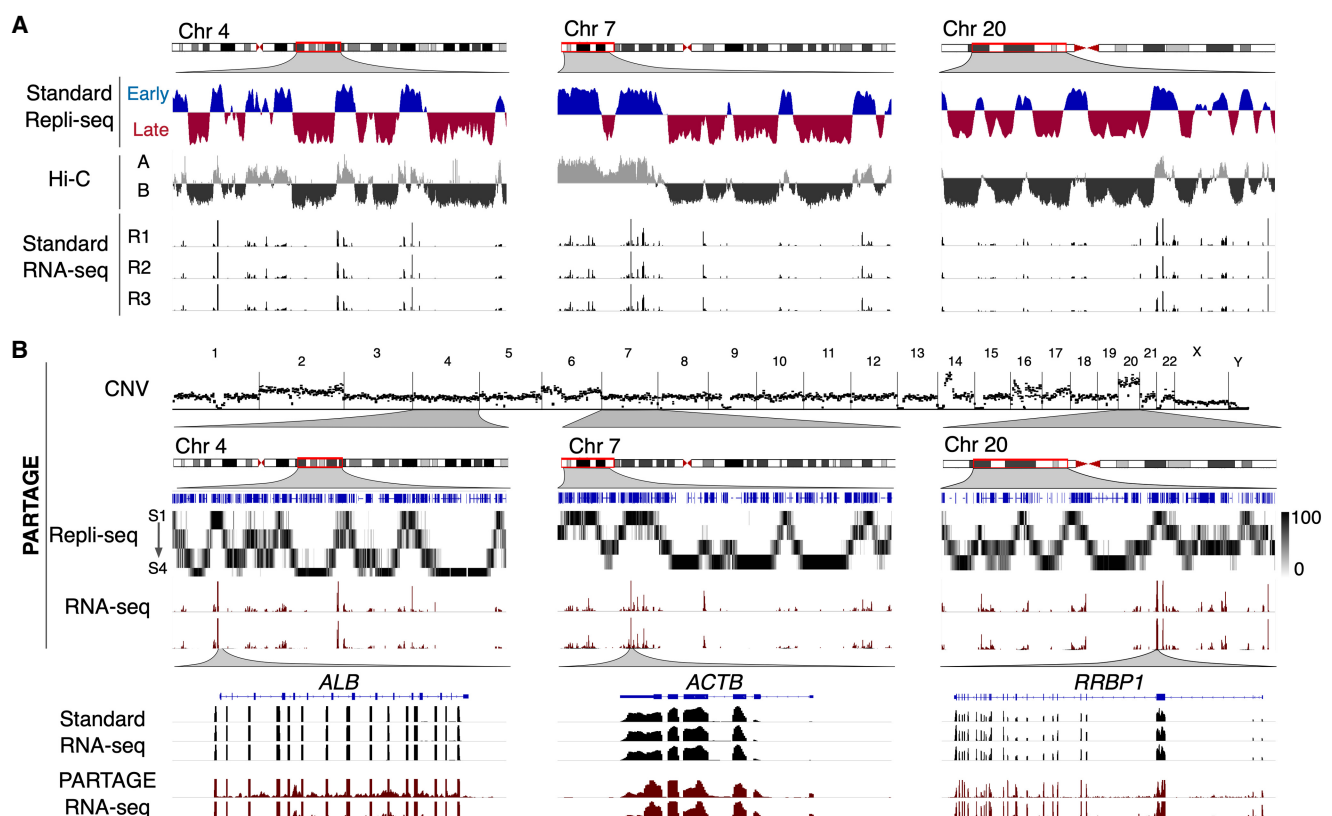


Figure 3. PARTAGE captures CNV, RT, and transcriptomes from the same samples. (A) Standard omics analysis generated from HepG2 cells. Three exemplary chromosomal regions are shown. Standard Repli-seq, chromosome confirmation capture (Hi-C), and RNA-seq were performed from separate sample batches of HepG2 cells. (B) PARTAGE simultaneous mapping of CNV, RT, and transcriptome from HepG2 nuclei. CNV was calculated from the sequencing reads obtained from G1 cell populations. Sequencing reads normalized for depth and per 1-Mb bin are shown. PARTAGE Repli-seq was performed using multifraction RT analysis. Heat map signals of G1-normalized reads are shown per fraction (S1–S4). RNA-seq coverage tracks are shown for the same chromosomal regions. Bottom tracks show a comparison between standard and PARTAGE RNA-seq signals zoomed in on hepatic-specific (*ALB*) and housekeeping genes (*ACTB* and *RRBP1*).

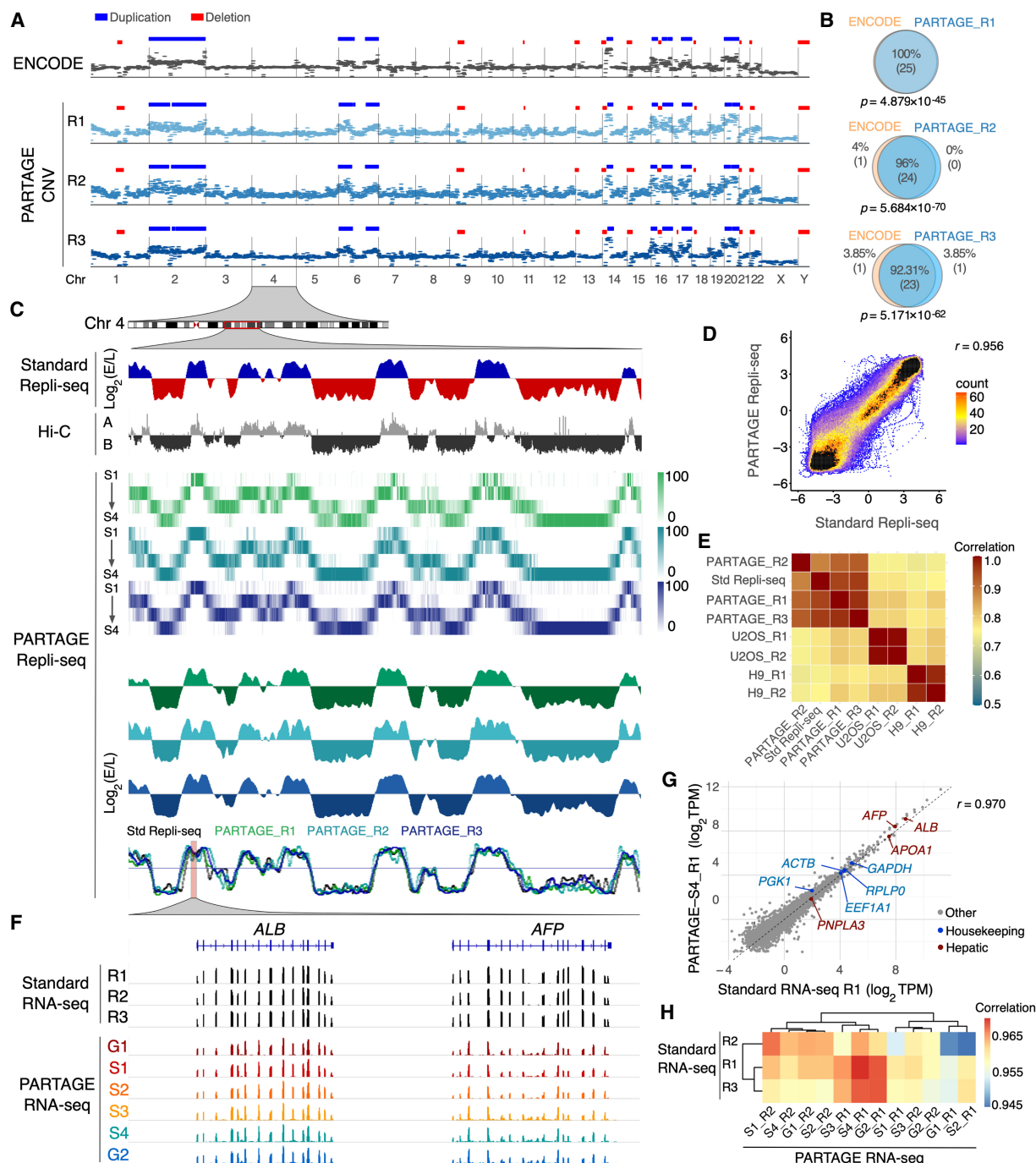


Figure 4. PARTAGE accurately profiles CNV, RT, and transcriptomes. (A) Comparative CNV analysis. CNV profiles derived from ENCODE high-coverage WGS (60× coverage; top track) and from PARTAGE (~0.5× coverage; three independent replicates) using CNVpytor (see Methods). CNV segments identified in the ENCODE data set ($n = 25$ events) are shown as duplications (blue) and deletions (red) and were used as a reference. PARTAGE CNV calls are also shown. (B) Overlap of CNV calls between ENCODE and PARTAGE replicates was quantified. PARTAGE correctly detected 92%–100% of the ENCODE CNV events across replicates. The statistical significance of the overlap between ENCODE and PARTAGE CNVs was assessed using ProOvOverlap (see Methods). (C) Comparative analysis of standard and PARTAGE Repli-seq data. Standard Repli-seq is shown as log₂(E/L) ratio signals. Multifraction Repli-seq is shown in distinct heat map colors per replicate. Signals are normalized counts against G1 per 20-kb bins. Collapsed E/L log₂ ratios were calculated from PARTAGE multifraction Repli-seq RT to obtain RT profiles directly comparable to standard E/L Repli-seq. (D) Scatterplot of PARTAGE (y-axis) versus standard (x-axis) Repli-seq data. (E) Genome-wide correlation between a PARTAGE and standard Repli-seq. Genome-wide RT data in 20-kb bins were used. RT from stem cells (H9) and cancer cells (U2OS) were included as a comparison. (F) PARTAGE and standard RNA-seq coverage tracks at the albumin locus showing expression patterns of the hepatic genes *ALB* and *AFP*. Raw coverage normalized for sequencing depth was used. (G) Gene-by-gene scatterplots on batch-corrected expression values (log₂TPM). Standard RNA-seq (x-axis) versus PARTAGE RNA-seq (y-axis) is shown. Housekeeping genes (blue) and hepatic markers (red) are highlighted; all other genes are gray. (H) Genome-wide cross-protocol concordance between a PARTAGE and standard RNA-seq transcriptomes. Pearson's correlation among all samples, computed on batch-corrected log₂TPM values, and clustered by Euclidean distance, is shown. Note that the correlation map is confined to a narrow range near 1.0 reflecting strong similarity across transcriptomes.

Fig. S6). Moreover, pairwise Pearson's correlations resulted in a very narrow, high range close to 1.0, indicating a strong agreement between standard RNA-seq and PARTAGE libraries (Fig. 4H). Overall, these results validate the simultaneous and precise mapping of CNV, RT, and gene expression patterns using PARTAGE.

Cell cycle dynamics in RT and gene expression

A long-standing correlation has been observed between early replication and active gene expression (Hiratani et al. 2010; Rivera-Mulia et al. 2015, 2019a; Rivera-Mulia and Gilbert 2016b; Vouzas and Gilbert 2023). However, previous studies were based on RT and gene expression data obtained independently from distinct batches of samples generated in parallel. Thus, we exploited PARTAGE to explore how closely early replication and gene expression are linked, taking advantage of the RT and transcriptome measurements from the same sample that reduces technical noise and batch effects. Genomic tracks of PARTAGE RT profiles and transcriptomic coverage patterns confirmed that active transcription is enriched at early replication chromosomal regions (Fig. 5A). To measure the strength of this correlation, we calculated the percentage of expressed genes in relation to their RT. To do so, we identified all expressed genes and plotted them according to their RT (collapsed E/L ratios). Our results recapitulate the well-established correlation between early replication and active gene expression, which further validates this multiomic method (Fig. 5B). Moreover, a regression analysis was performed and showed a strong correlation between RT and transcriptional activity (Fig. 5B). These results confirm the coregulation between RT and active transcription with the advantage of PARTAGE measuring these modalities within the same sample and cell-cycle-defined populations, minimizing batch and compositional differences.

PARTAGE also enables the profiling of cell cycle phase-resolved transcriptome profiling, providing a direct measurement of dynamic regulation of gene expression patterns across the cell cycle. We first determined the degree of similarity between the transcriptomes obtained from the distinct stages of the cell cycle. After removing nonexpressed genes, applying batch correction, and performing pairwise correlation analysis, we observed strong correlations ($r \geq 0.95$) within a narrow, high range between all cell cycle stages (Fig. 5C), indicating technical reproducibility of our methods and global stability of the HepG2 transcriptome across the cell cycle in unperturbed conditions. To identify transcriptional changes during the cell cycle, we identified genes expressed in a minimum of two samples and filtered those with $\geq 4\times$ change across cell cycle stages. Under these criteria, we detected 83 protein-coding genes changing expression patterns during the cell cycle. Clustering analysis revealed a group of coregulated genes that increase expression during cell cycle progression, containing cell cycle regulators (Fig. 5D). As expected, ontology analysis confirmed that these variant genes are associated with processes regulating the cell cycle progression (Fig. 5E; Supplemental Fig. S7). Notably, the number of cell cycle regulated genes is substantially smaller than previous reports based on synchronized cells. Exemplary PARTAGE tracks of cell cycle-regulated genes are shown in Figure 5F. Overall, these analyses illustrate how PARTAGE provides an integrated framework with the potential to detect CNV, RT, and gene expression regulation during the cell cycle from the same experimental samples.

Discussion

Analyses of genome organization and function have unveiled the coregulation between RT and gene expression (Rivera-Mulia et al. 2015, 2019a; Dileep et al. 2019; Sima et al. 2019). Recent studies of early mammalian embryogenesis have begun to elucidate when RT emerges during development and how this relates to chromatin states and transcriptional activity (Halliwell et al. 2024; Nakatani et al. 2024, 2025; Takahashi et al. 2024; Xu et al. 2024), but they reach differing conclusions regarding the presence and consolidation of RT programs in the zygote and earliest cell divisions. These discrepancies highlight the complexity of RT regulation and the need for technological advances to jointly interrogate RT and gene expression. Notably, previous studies rely on separate profiling of these genomic properties, which might obscure the complex coregulation between the epigenome and transcriptome. Multiomic approaches can help to address these questions; however, rigorous testing is necessary to avoid artifacts in RT mapping (Rivera-Mulia 2025). In this work, we present PARTAGE, a multiomics approach that enables the parallel analysis of RT and gene expression, as well as CNV, from the same samples.

The major limitation for the joint mapping of transcriptomes and RT was that methods for cell cycle analysis by FACS relied on methods that do not preserve the RNA. Standard cell cycle analysis by flow cytometry depends on the precise DNA content measurement using propidium iodide (PI) or diamidino-2-phenylindole (DAPI) staining. PI staining requires RNase treatment, whereas DAPI staining requires permeabilization that leads to RNA loss. To overcome this challenge, we exploited nontoxic, cell-permeant dyes with narrow emission spectra (DyeCycle dyes, Invitrogen). These dyes enable the stoichiometric DNA staining for accurate cell cycle analysis based in DNA content (Bradford et al. 2006) and can be used for Repli-seq sample preparation (Meyer-Nava et al. 2023). Because these dyes are not RNA sensitive, they allow precise DNA content measurement and S-phase gating of nuclei without the need for RNase treatment, thereby preserving nuclear RNA for downstream sequencing. Here, we processed isolated nuclei in order to enrich for nuclear RNA, which is closer to ongoing transcriptional activity and less influenced by cytoplasmic mRNA stability and export, as supported by analyses of intronic (nuclear) versus exonic reads in RNA-seq (Zaghlool et al. 2013; Gaidatzis et al. 2015; Bakken et al. 2018; Oh et al. 2022; Engström et al. 2025). In principle, our PARTAGE workflow can be adapted to whole-cell preparations to obtain steady-state RNA-seq from the same sorted populations, although specific library preparations would need validation. We demonstrated that PARTAGE sample preparation allows the copurification of DNA and RNA of high quality and in sufficient amounts for parallel processing for Repli-seq and RNA-seq. In fact, we found increasing yields of DNA that reach $2\times$ in G2/M as compared to G1, highlighting the accuracy of this method to collect cell populations throughout the cell cycle (Fig. 2).

Conventional methods for CNV detection rely on deep sequencing ($\geq 30\times$ coverage) to accurately identify variations by ensemble methods. Here, we demonstrated that PARTAGE can accurately detect major amplifications in the HepG2 genome with only $\leq 0.5\times$ coverage. Despite the shallow sequencing (~ 8 M single-end 100-bp reads), PARTAGE reproduced the large CNV detected by the ENCODE mapping (Fig. 4). This was achieved by the focused analysis of purified nuclei in G1-phase, which ensures the analysis of uniform diploid populations by removing S-phase

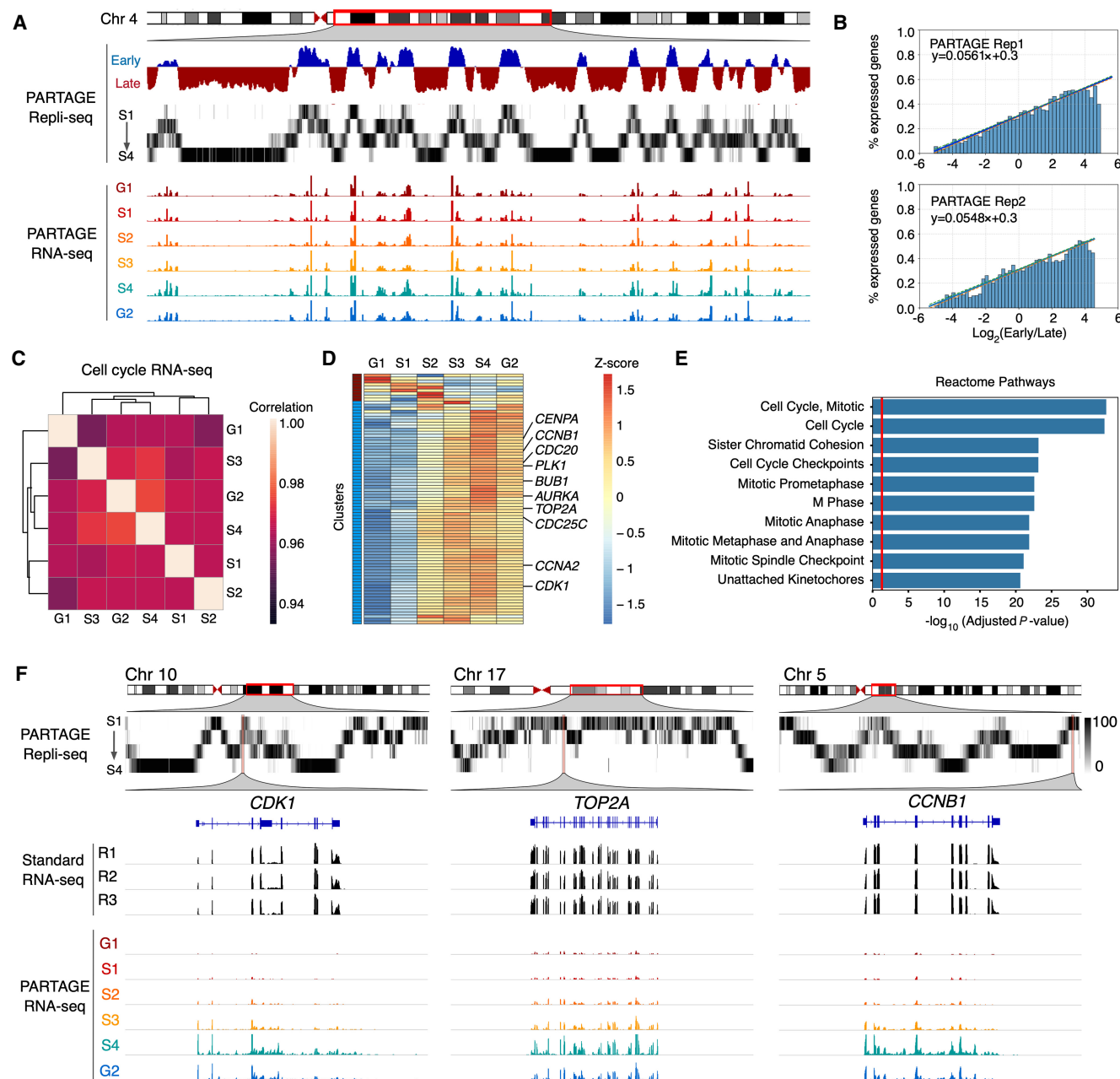


Figure 5. Cell cycle regulation of DNA replication and gene expression. (A) Exemplary genomic locus from Chromosome 4 showing enrichment of gene expression at the early replicating regions. Collapsed E/L ratios from PARTAGE Repli-seq data, PARTAGE multifraction signals, and coverage RNA-seq tracks are shown. (B) Genome-wide correlation between replication timing and gene expression. Expressed genes were identified (threshold of $\geq 2 \log_2 \text{TPM}$) and sorted by their RT. The y-axis represents the percentage of expressed genes within each bin. Logistic regression (inner line) and 95% confidence intervals (outer lines) display the correlation strength. Two PARTAGE replicates are shown. (C) Genome-wide correlation of transcriptome programs across the cell cycle. Average TPMs were calculated per cell cycle stage after batch correction (ComBat 2.0), and gene expression values $\geq 0.5 \log_2 \text{TPM}$ in at least three cell cycle stages were used. Pairwise Pearson's correlation was performed with hierarchical clustering based on correlation distance ($1 - r$). Note that the correlation scale is confined to a narrow range near 1.0 highlighting minor differences between cell cycle stages. (D) Heat map and clustering analysis of variable genes during the cell cycle. Genes are represented by row-wise z-scored $\log_2(\text{TPM})$ values and clustered by k -means ($k=2$). Only genes expressed in ≥ 2 cell cycle stages and with $\geq 4 \times \Delta \log_2 \text{TPM}$ were considered. (E) Reactome ontology analysis of the major cluster (blue) is shown. (F) Exemplary chromosome regions harboring variable genes that are induced during cell cycle progression (*CDK1*, *TOP2A*, and *CCNB1*).

and G2/M-phase nuclei, and reducing cell cycle-associated artifacts in CNV detection. Thus, PARTAGE enables the accurate detection of CNV, providing additional tools to better understand genome instability in development, disease, and cancer states. Importantly, PARTAGE's capacity to detect CNVs with this sequencing depth

makes this approach cost-effective and scalable for clonal tracking, genome integrity screenings, and cancer alterations studies, with the addition of joint profiling of RT and gene regulation.

For Repli-seq analysis in PARTAGE, we took advantage of our optimized Repli-seq workflow, which reduced DNA loss and

technical noise by minimizing transfer/handling steps, and increasing yields by 3× (Rivera-Mulia et al. 2022). We also applied the multifraction Repli-seq approach and exploited the data from the G1-purified nuclei to normalize signals and generate accurate heat maps of RT. Although standard E/L Repli-seq is also compatible with our PARTAGE workflow, it is sufficient to accurately detect RT and is more cost-effective; we selected multifraction Repli-seq to enable a more accurate detection of cell cycle regulation when integrating the data with the matching transcriptomes. The combined implementation of multifraction analyses and optimized library preparation strategies allowed us to generate highly reproducible RT data across replicates that recapitulate the RT data obtained from the standard method.

We also evaluated the robustness of the RNA-seq data generated in PARTAGE and found high reproducibility across replicates. Moreover, although PARTAGE RNA-seq was performed on low-input samples (20 K nuclei), it faithfully recapitulated the transcriptome obtained with standard RNA-seq from 0.5 M cells (Fig. 4). Moreover, PARTAGE sample preparation allows us to collect nuclei at distinct phases of the cell cycle. Thus, we asked whether PARTAGE could detect transcriptional changes across the cell cycle. Notably, PARTAGE profiles RNA in native conditions without the need of forced cell synchronization. Conventional studies of cell cycle-regulated transcriptional activity depend on synchronization strategies through serum starvation, mitotic shake-off, or using drugs that activate checkpoint mechanisms. However, these approaches trigger cellular stress responses, DNA damage pathways, and/or metabolic reprogramming that alter normal transcriptional regulation (Cooper 2003; Karlsson et al. 2017; Mahdessian et al. 2021). Thus, PARTAGE is a valuable alternative that does not require any external perturbation and enables a more accurate assessment of the cell cycle regulation of the transcriptome. Indeed, the simultaneous profiling of RT and transcriptomes in PARTAGE facilitated the integration of multimodal signals under physiological conditions and confirmed the strong enrichment of actively transcribed genes in early replicating regions (Fig. 5). Moreover, PARTAGE identified a subset of genes that increase transcriptional levels 4× as the cells progress in the cell cycle and confirmed that these genes are indeed involved in cell cycle regulation. The relatively small number of cell cycle-regulated genes detected in PARTAGE can be due to different factors: (1) the reduced synchronization-induced artifacts by processing cells in native conditions; (2) the processing of nuclei instead of whole cells to reduce confounding effects from cytoplasmic post-transcriptional regulation (e.g., mRNA stability and export); and (3) the use of stringent thresholds to identify variable genes.

Important open questions, such as whether RT changes precede or follow chromatin and transcriptome programs, can be addressed by jointly profiling these genomic properties across defined developmental trajectories. In its current implementation, PARTAGE provides a robust multiomic method that is directly applicable to differentiation systems of pluripotent stem cells. A limitation of PARTAGE is that it is an ensemble method, and further optimization and adaptations to low-input and single-cell formats will be necessary to better evaluate the temporal emergence of RT relative to chromatin and the transcriptome, especially during early embryogenesis.

Overall, PARTAGE is a robust method that provides accurate mapping of CNV, RT, and transcriptomes. We validated PARTAGE as a multiomic approach capable of generating highly accurate data consistent with those produced using the gold-standard separate methods (Repli-seq, RNA-seq, and CNV mapping).

This high consistency was quantitatively demonstrated by the near complete overlap of CNV calls, by strong genome-wide correlation ($r \geq 0.95$) between PARTAGE and standard Repli-seq RT profiles, and by the high concordance ($r \geq 0.945$) found between PARTAGE and standard transcriptome profiling, despite using significantly lower input material. This approach reduces the required sample input as compared to the corresponding separate approaches, minimizes heterogeneity and technical noise between batches of samples, and enables a direct assessment and integration of multimodal signals. Moreover, PARTAGE enables the characterization of cell cycle regulation of genome function in unperturbed conditions. PARTAGE is well suited for investigating coordinated changes in RT, genome stability, and transcriptional activity in dynamic system such as differentiation trajectories of pluripotent stem cells. In disease contexts, such as cancer and other conditions characterized by genome instability, PARTAGE provides a practical approach to concurrently monitor RT alterations, CNV burden, and transcriptional dysregulation, thereby enabling more comprehensive assessments of genome integrity and its functional implications.

Methods

Cell culture

The HepG2 cell line was obtained from ATCC and expanded according to the supplier's protocol. Briefly, cells were maintained in Eagle's Minimum Essential Medium (ATCC, #30-2003) supplemented with 10% FBS (Corning, #35011CV). Cell cultures up to 60% confluency were treated with BrdU (Sigma-Aldrich, #B5002) at a final concentration of 100 μ M at 37°C for 2 h and processed for nuclei preparation for PARTAGE analyses.

Intact nuclei preparation for PARTAGE

BrdU-labeled HepG2 cells were dissociated into single-cell suspensions by treatment with trypsin-EDTA (Life Technologies, #A011105-01), and intact nuclei were isolated as previously described (Meyer-Nava et al. 2023). Briefly, cell pellets were resuspended in 1 mL of DMEM/F-12 50/50 (1×), counted, and aliquoted into batches of 2.0×10^6 cells. Then, cells were triturated in a lysis buffer: 0.025% IGEPAL (Millipore Sigma, #542334) in 10 mM TrisHCl, 10 mM NaCl, 3 mM MgCl₂, and 1× PBS (Thermo Fisher Scientific, #20012027). Lysed cells were washed three times and resuspended in 1 mL PBS/1% FBS. Finally, nuclei were stained with Vybrant DyeCycle Violet Stain (Thermo Fisher Scientific, #V35003) at 1.5 μ L per 1×10^{-6} cells for 30 min at 37°C and filtered with a 36- μ m nylon mesh (Fisher Scientific, #NC05156770).

Cell cycle analysis and cell sorting by flow cytometry

Stained nuclei were processed in a SONY SH800 cell sorter. Single nuclei were identified and cell cycle stages determined based on DNA content (Vybrant DyeCycle Violet intensity). Twenty thousand cells were collected per fraction of G1, S1, S2, S3, S4, and G2M cell cycle stages and sorted directly into wells of a 96-well plate (Axygen, #14-223-345) containing 400 μ L of DNA/RNA Shield Buffer (Zymo, #R1100). If not processed immediately, the plates were sealed with an adhesive film and stored at −80°C.

Nucleic acid extraction

For DNA and RNA copurification, the samples were thawed at room temperature and nucleic acids isolated using the Quick DNA/RNA MicroPrep Plus kit (Zymo, #D7005) according to the

manufacturer's instructions. Purified DNA and RNA were quantified using Qubit kits (Invitrogen, #Q32852 and #Q32851).

CNV library preparation and whole-genome sequencing

For CNV analysis, we used the genomic DNA purified from G1 cells and performed WGS. Briefly, purified DNA from G1 sorted cells was directly processed using the KAPA HyperPlus library preparation kit (Roche, #07958978001) according to the manufacturer's instructions. Enzymatic fragmentation was performed at 37°C for 30 min, followed by end-repair and ligation of KAPA dual-indexed adapters (Roche, #08278555702). Libraries were purified using the Zymo DNA Clean & Concentrator-5 kit, amplified using the KAPA HiFi HotStart from the KAPA HyperPlus kit according to the manufacturer's instructions, and purified using AMPure XP beads. Low-coverage whole-genome sequencing libraries were sequenced at the University of Minnesota Genomics Center (UMGC), on the Illumina NextSeq P2 Illumina platform with 100 bp of single-end sequencing and ~0.5× genome coverage.

Repli-seq

DNA samples were processed using the optimized Repli-seq protocol as previously described (Rivera-Mulia et al. 2022). Briefly, purified DNA from sorted cells was directly processed using the KAPA HyperPlus library preparation kit (Roche, #07958978001) according to the manufacturer's instructions. Enzymatic fragmentation was performed at 37°C for 30 min, followed by end-repair and ligation of KAPA dual-indexed adapters (Roche, #08278555702). Libraries were purified using the Zymo DNA Clean & Concentrator-5 kit and processed for immunoprecipitation using antibodies against BrdU. Immunoprecipitated libraries were amplified using the KAPA HiFi HotStart from the KAPA HyperPlus kit according to the manufacturer's instructions and purified using AMPure XP beads. DNA concentration and size distribution were evaluated by Qubit quantification (Invitrogen, #Q32851), and size estimation was performed on TapeStation High Sensitivity DNA ScreenTape (Agilent, #5067–5592). Repli-seq libraries were pooled at a final concentration of 10 nM and sequenced at the University of Minnesota Genomics Center (UMGC), on the Illumina NextSeq P2 Illumina platform at ~8 M reads per library and 100 bp of single-end sequencing.

Standard RNA-seq

RNA samples were isolated from 0.5 million HepG2 cells using the RNeasy Plus kit (Qiagen, #74004), and libraries were prepared with the TruSeq Stranded mRNA Library Prep (Illumina) according to the manufacturer's instructions. Standard RNA-seq libraries were sequenced on the Illumina NovaSeq platform at ~20 M reads per library and 150 bp of pair-end sequencing.

PARTAGE RNA-seq

RNA samples were processed using the SMARTer Stranded Total RNA-seq Pico for ultralow inputs (Takara, #634411). RNA concentration and size distribution were evaluated by Qubit quantification (Invitrogen, #Q32852) and size measured using the Bioanalyzer RNA 6000 Nano kit (Agilent, #5067–1511). RNA-seq libraries were pooled at a final concentration of 10 nM and sequenced at the University of Minnesota Genomics Center (UMGC), on the Illumina NovaSeqX platform at ~40 M reads per library and 150 bp of pair-end sequencing.

Data analyses and visualization

Data analyses were performed with resources at the Minnesota Supercomputing Institute (MSI).

PARTAGE CNV data

Sequencing reads obtained from G1 fractions were trimmed for adapter sequences using cutadapt 4.9 (Martin 2011) and aligned to the reference genome with Burrows-Wheeler Aligner 0.7.18 (Li and Durbin 2010). Sequence Alignment Map (SAM) files were converted into Binary Alignment Maps (BAM) with SAMtools 1.18 (Danecek et al. 2021). Quality filter and removal of PCR duplicates were performed with SAMtools using the view and markdup functions, respectively. Mapped, filtered, and deduplicated reads were counted into genomic windows of 1 Mb using the intersect function from BEDTools 2.31.1 (Quinlan and Hall 2010). CNV tracks were visualized using the Integrative Genomics Viewer (IGV) browser (Robinson et al. 2023).

CNV calling

CNVs were called using CNVpytor (Suvakov et al. 2021). Briefly, aligned whole-genome sequencing reads from ENCODE (60× coverage) and from PARTAGE libraries (~0.5× coverage) were processed with CNVpytor following the developers' recommendations. Read-depth histograms were generated, normalized, and used to infer copy-number states, and CNV segments were defined using CNVpytor's built-in segmentation and significance thresholds. CNV calling for PARTAGE and ENCODE data sets was performed with the same pipeline to enable direct comparison.

CNV overlap analysis

To assess the concordance between CNVs detected in PARTAGE and those detected in the high-coverage ENCODE reference, we used ProOvErlap (Gualandi et al. 2025) to test the statistical significance of overlap between CNV segments. CNVs identified by CNVpytor in ENCODE were used as the reference set, and CNVs from each PARTAGE replicate were used as query sets. ProOvErlap was used to calculate the expected overlap under a random model and to obtain *P*-values for the observed overlap between reference and query CNVs.

Standard Repli-seq data

Sequencing reads were processed as for CNV (trimmed for adapter sequences, aligned with Burrows-Wheeler Aligner 0.7.18, and PCR duplicates removed). Mapped, filtered, and deduplicated reads were counted into genomic windows of distinct size (from 5- to 100-kb) using the intersect function from BEDTools 2.31.1 (Quinlan and Hall 2010). Downstream processing, including quantile normalization and LOESS smoothing, and data visualization, was performed in R (R Core Team 2017) as previously described (Marchal et al. 2018; Rivera-Mulia et al. 2022) and IGV browser (Robinson et al. 2023).

PARTAGE Repli-seq data

Binned RPKM data for each sample was obtained as above and further processed into 50-kb bins. For each set of samples, the $\log_2([\text{sample}/\text{G1_sample}] + 1)$ was calculated. Data were scaled linearly from 0 to 100 for visualization. Bedgraphs were generated and displayed using the IGV browser (Robinson et al. 2023).

For collapsed Repli-seq from PARTAGE data, FASTQ files were concatenated for S1 and S2 as the early sample and S3 and S4 as the late sample. Data were then processed as above, including loess

smoothing and quantile normalization for comparison to standard Repli-seq data.

RNA-seq data

Raw read counts were trimmed using Trimmomatic 0.39 (Bolger et al. 2014), aligned using HISAT2 2.1.0 (Kim et al. 2015), and alignment filtering and counts using SAMtools 1.18. Raw counts were converted to TPM counts by calculating reads per kilobase (RPK) and scaling by the sum of RPK/1,000,000. The $\log_2(\text{TPM} + 1)$ was used for downstream analyses. For analysis of expression patterns, samples were median-centered, low-expressed genes ($\text{TPM} < 1$) and all noncoding genes were removed, and batch correction was calculated using ComBat 2.0 (Johnson et al. 2007). Mean values were calculated across samples for each time point. Only genes expressed in at least two samples and with a fourfold TPM change were considered. *k*-means clustering was performed on the remaining genes with a cluster count of two. Gene set enrichment was performed using gseapy (Fang et al. 2023) and Enrichr (Chen et al. 2013) on the following terms: “ChEA_2022”, “ENCODE_and_ChEA_Consensus_TFs_from_ChIP-X”, “GO_Biological_Process_2025”, “GO_Cellular_Component_2025”, “GO_Molecular_Function_2025”, “Reactome_Pathways_2024”. Significantly enriched terms were filtered by gene count (> 1) and adjusted Benjamini–Hochberg *P*-value < 0.05 . For each category, the top 10 most significant gene categories were plotted using seaborn.

For comparisons of RNA and RT data, plots were created comparing the number of bins with expressed genes compared to their corresponding replication timing values. For each sample, genome wide RT data were sorted by value and fit into 100 bins. For each bin, gene expression values were binarized into 0 sec and 1 sec based on $\log_2(\text{TPM})$ expression > 1 . A linear regression model was fit to the data using statsmodels (Seabold and Perktold 2010).

Data access

All raw sequencing data generated in this study have been submitted to the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>) under accession number PRJNA1312882. Processed data have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE307022. Open-source code used in the analysis of PARTAGE data and generation of the figures is available at GitHub (<https://github.com/Rivera-Mulia-Lab/PARTAGE>) and as Supplemental Code.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

This work was supported by National Institutes of Health/National Institute of General Medical Sciences grants R35GM137950, R35GM137950-02S1, and R35GM137950-04S1 to J.C.R-M.; and by Regenerative Medicine Minnesota RMM-091621-DS-006 to J.C.R-M.; and institutional support from the University of Minnesota Medical School.

Author contributions: Conceptualization: J.C.R-M. Methodology: J.C.R-M. Investigation: L.S.S.M., Q.D., S.M.-N., and J.C.R-M. Formal analysis: Q.D. and J.C.R-M. Visualization: Q.D. and J.C.R-M. Data curation: Q.D. and J.C.R-M. Writing—original draft: L.S.S.M., Q.D., and J.C.R-M. Writing—review and editing: L.S.S.M., Q.D., and J.C.R-M. Funding acquisition: J.C.R-M. Resources: J.C.R-M. Supervision: J.C.R-M. Project administration: J.C.R-M.

References

- Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aevermann B, Barkan E, Bertagnolli D, Casper T, Dee N, et al. 2018. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* **13**: e0209648. doi:10.1371/journal.pone.0209648
- Bartlett DA, Dileep V, Baslan T, Gilbert DM. 2022. Mapping replication timing in single mammalian cells. *Curr Protoc* **2**: e334. doi:10.1002/cpz1.334
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170
- Bradford JA, Whitney P, Huang T, Pinson P, Cheung C-Y, Yue S, Godfrey WL. 2006. Novel Vybrant® DyeCycle™ stains provide cell cycle analysis in live cells using flow cytometry with violet, blue, and green excitation. *Blood* **108**: 4234. doi:10.1182/blood.V108.11.4234.4234
- Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, Daza RM, McFaline-Figueroa JL, Packer JS, Christiansen L, et al. 2018. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**: 1380–1385. doi:10.1126/science.aau0730
- Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**: 128. doi:10.1186/1471-2105-14-128
- Chen S, Lake BB, Zhang K. 2019. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat Biotechnol* **37**: 1452–1457. doi:10.1038/s41587-019-0290-0
- Cooper S. 2003. Rethinking synchronization of mammalian cells for cell cycle analysis. *Cell Mol Life Sci* **60**: 1099–1106. doi:10.1007/s00018-003-2253-2
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* **10**: giab008. doi:10.1093/gigascience/giab008
- Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306–1311. doi:10.1126/science.1067799
- Dileep V, Gilbert DM. 2018. Single-cell replication profiling to measure stochastic variation in mammalian replication timing. *Nat Commun* **9**: 427. doi:10.1038/s41467-017-02800-w
- Dileep V, Wilson KA, Marchal C, Lyu X, Zhao PA, Li B, Poulet A, Bartlett DA, Rivera-Mulia J, Qin ZS, et al. 2019. Rapid irreversible transcriptional reprogramming in human stem cells accompanied by discordance between replication timing and chromatin compartment. *Stem Cell Reports* **13**: 193–206. doi:10.1016/j.stemcr.2019.05.021
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380. doi:10.1038/nature11082
- Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, Yardımcı GG, Chakraborty A, Bann DV, Wang Y, et al. 2018. Integrative detection and analysis of structural variation in cancer genomes. *Nat Genet* **50**: 1388–1398. doi:10.1038/s41588-018-0195-8
- The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74. doi:10.1038/nature11247
- The ENCODE Project Consortium, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shores N, Adrian J, Kawi T, Davis CA, Dobin A, et al. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**: 699–710. doi:10.1038/s41586-020-2493-4
- Engström K, Nilsson Å, Ofori JK, Wierup N, Bacos K, Ling C. 2025. Evaluating cell-specific gene expression using single-cell and single-nuclei RNA-sequencing data from human pancreatic islets of the same donors. *Sci Rep* **15**: 36133. doi:10.1038/s41598-025-21595-1
- Fang Z, Liu X, Peltz G. 2023. GSEAPy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**: btac757. doi:10.1093/bioinformatics/btac757
- Flynn E, Almonte-Loya A, Fragiadakis GK. 2023. Single-cell multiomics. *Annu Rev Biomed Data Sci* **6**: 313–337. doi:10.1146/annurev-biodata-tasci-020422-050645
- Gaidatzis D, Burger L, Florescu M, Stadler MB. 2015. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat Biotechnol* **33**: 722–729. doi:10.1038/nbt.3269
- Gilbert DM. 2002. Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* **14**: 377–383. doi:10.1016/S0955-0674(02)00326-5
- Gnan S, Josephides JM, Wu X, Spagnuolo M, Saulebekova D, Bohec M, Dumont M, Baudrin LG, Fachinetti D, Baulande S, et al. 2022. Kronos scRT: a uniform framework for single-cell replication timing analysis. *Nat Commun* **13**: 2329. doi:10.1038/s41467-022-30043-x

- Gualandi N, Bertozzo A, Brancolini C. 2025. ProOvErlap: assessing feature proximity/overlap and testing statistical significance from genomic intervals. *J Biol Chem* **301**: 110209. doi:10.1016/j.jbc.2025.110209
- Halliwell JA, Martin-Gonzalez J, Hashim A, Dahl JA, Hoffmann ER, Lerdrup M. 2024. Sex-specific DNA-replication in the early mammalian embryo. *Nat Commun* **15**: 6323. doi:10.1038/s41467-024-50727-w
- Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, Dorschner MO, Gartler SM, Stamatoyannopoulos JA. 2010. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci* **107**: 139–144. doi:10.1073/pnas.0912402107
- Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang C-W, Lyou Y, Townes TM, Schübeler D, Gilbert DM. 2008. Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* **6**: e245. doi:10.1371/journal.pbio.0060245
- Hiratani I, Takebayashi S-I, Lu J, Gilbert DM. 2009. Replication timing and transcriptional control: beyond cause and effect—part II. *Curr Opin Genet Dev* **19**: 142–149. doi:10.1016/j.gde.2009.02.002
- Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B, Fussner E, Bazett-Jones DP, Plath K, Dalton S, et al. 2010. Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Res* **20**: 155–169. doi:10.1101/gr.099796.109
- Hulke ML, Massey DJ, Koren A. 2020. Genomic methods for measuring DNA replication dynamics. *Chromosome Res* **28**: 46–67. doi:10.1007/s10577-019-09624-y
- Hyrien O, Guilbaud G, Krude T. 2025. The double life of mammalian DNA replication origins. *Genes Dev* **39**: 304–324. doi:10.1101/gad.352227.124
- Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**: 118–127. doi:10.1093/biostatistics/kxj037
- Karlsson J, Kroneis T, Jonasson E, Larsson E, Ståhlberg A. 2017. Transcriptomic characterization of the human cell cycle in individual unsynchronized cells. *J Mol Biol* **429**: 3909–3924. doi:10.1016/j.jmb.2017.10.011
- Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357–360. doi:10.1038/nmeth.3317
- Klein KN, Zhao PA, Lyu X, Sasaki T, Bartlett DA, Singh AM, Tasan I, Zhang M, Watts LP, Hiraga S-I, et al. 2021. Replication timing maintains the global epigenetic state in human cells. *Science* **372**: 371–378. doi:10.1126/science.aba5545
- Lande-Diner L, Zhang J, Cedar H. 2009. Shifts in replication timing actively affect histone acetylation during nucleosome reassembly. *Mol Cell* **34**: 767–774. doi:10.1016/j.molcel.2009.05.027
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**: 589–595. doi:10.1093/bioinformatics/btp698
- 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. doi:10.1126/science.1181369
- Liu Z, Chen Y, Xia Q, Liu M, Xu H, Chi Y, Deng Y, Xing D. 2023. Linking genome structures to functions by simultaneous single-cell Hi-C and RNA-seq. *Science* **380**: 1070–1076. doi:10.1126/science.adg3797
- Liu Y, Zhangding Z, Liu X, Hu J. 2025. Chromatin-centric insights into DNA replication. *Trends Genet* **41**: 412–424. doi:10.1016/j.tig.2024.12.003
- Mahdessian D, Cesnik AJ, Gnann C, Danielsson F, Stenström L, Arif M, Zhang C, Le T, Johansson F, Schutten R, et al. 2021. Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**: 649–654. doi:10.1038/s41586-021-03232-9
- Marchal C, Sasaki T, Vera D, Wilson K, Sima J, Rivera-Mulia JC, Trevilla-García C, Noguez C, Nafie E, Gilbert DM. 2018. Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *Nat Protoc* **13**: 819–839. doi:10.1038/nprot.2017.148
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**: 10–12. doi:10.14806/ej.17.1.200
- Massey DJ, Koren A. 2022. High-throughput analysis of single human cells reveals the complex nature of DNA replication timing control. *Nat Commun* **13**: 2402. doi:10.1038/s41467-022-30212-y
- Meyer-Nava S, Shetty AV, Rivera-Mulia JC. 2023. Repli-seq sample preparation using cell sorting with cell-permeant dyes. *Curr Protoc* **3**: e945. doi:10.1002/cpz1.945
- Miura H, Takahashi S, Shibata T, Nagao K, Obuse C, Okumura K, Ogata M, Hiratani I, Takebayashi S-I. 2020. Mapping replication timing domains genome wide in single mammalian cells with single-cell DNA replication sequencing. *Nat Protoc* **15**: 4058–4100. doi:10.1038/s41596-020-0378-5
- Nakatani T. 2025. Dynamics of replication timing during mammalian development. *Trends Genet* **41**: 568–576. doi:10.1016/j.tig.2025.01.010
- Nakatani T, Schauer T, Altamirano-Pacheco L, Klein KN, Ettinger A, Pal M, Gilbert DM, Torres-Padilla M-E. 2024. Emergence of replication timing during early mammalian development. *Nature* **625**: 401–409. doi:10.1038/s41586-023-06872-1
- Nakatani T, Schauer T, Pal M, Ettinger A, Altamirano-Pacheco L, Zorn J, Gilbert DM, Torres-Padilla M-E. 2025. RIF1 controls replication timing in early mouse embryos independently of lamina-associated nuclear organization. *Dev Cell* **60**: 2149–2162.e7. doi:10.1016/j.devcel.2025.03.016
- Oh J-M, An M, Son D-S, Choi J, Cho YB, Yoo CE, Park W-Y. 2022. Comparison of cell type distribution between single-cell and single-nucleus RNA sequencing: enrichment of adherent cell types in single-nucleus RNA sequencing. *Exp Mol Med* **54**: 2128–2134. doi:10.1038/s12276-022-00892-z
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. doi:10.1093/bioinformatics/btq033
- R Core Team. 2017. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Rhind N. 2022. DNA replication timing: biochemical mechanisms and biological significance. *Bioessays* **44**: e2200097. doi:10.1002/bies.202200097
- Rivera-Mulia JC. 2025. The long-standing relationship between replication timing, gene expression, and chromatin accessibility is maintained in early mouse embryogenesis. *bioRxiv* doi:10.1101/2025.11.10.687227v1
- Rivera-Mulia JC, Gilbert DM. 2016a. Replicating large genomes: divide and conquer. *Mol Cell* **62**: 756–765. doi:10.1016/j.molcel.2016.05.007
- Rivera-Mulia JC, Gilbert DM. 2016b. Replication timing and transcriptional control: beyond cause and effect—part III. *Curr Opin Cell Biol* **40**: 168–178. doi:10.1016/j.ceb.2016.03.022
- 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. doi:10.1101/gr.187989.114
- Rivera-Mulia JC, Desprat R, Trevilla-García C, Cornacchia D, Schwerer H, Sasaki T, Sima J, Fells T, Studer L, Lemaitre J-M, et al. 2017. DNA replication timing alterations identify common markers between distinct progenitor diseases. *Proc Natl Acad Sci* **114**: E10972–E10980. doi:10.1073/pnas.1707143114
- Rivera-Mulia JC, Kim S, Gabr H, Chakraborty A, Ay F, Kahveci T, Gilbert DM. 2019a. Replication timing networks reveal a link between transcription regulatory circuits and replication timing control. *Genome Res* **29**: 1415–1428. doi:10.1101/gr.247049.118
- Rivera-Mulia JC, Sasaki T, Trevilla-García C, Nakamichi N, Knapp DJHF, Hammond CA, Chang BH, Tyner JW, Devidas M, Zimmerman J, et al. 2019b. Replication timing alterations in leukemia affect clinically relevant chromosome domains. *Blood Advances* **3**: 3201–3213. doi:10.1182/bloodadvances.2019000641
- Rivera-Mulia JC, Trevilla-García C, Martínez-Cifuentes S. 2022. Optimized Repli-seq: improved DNA replication timing analysis by next-generation sequencing. *Chromosome Res* **30**: 401–414. doi:10.1007/s10577-022-09703-7
- Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang X, Wang J, et al. 2015. Integrative analysis of 111 reference human epigenomes. *Nature* **518**: 317–330. doi:10.1038/nature14248
- Robinson JT, Thorvaldsdóttir H, Turner D, Mesirov JP. 2023. igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics* **39**: btac830. doi:10.1093/bioinformatics/btac830
- Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* **20**: 761–770. doi:10.1101/gr.099655.109
- Sasaki T, Rivera-Mulia JC, Vera D, Zimmerman J, Das S, Padgett M, Nakamichi N, Chang BH, Tyner J, Druker BJ, et al. 2017. Stability of patient-specific features of altered DNA replication timing in xenografts of primary human acute lymphoblastic leukemia. *Exp Hematol* **51**: 71–82.e3. doi:10.1016/j.exphem.2017.04.004
- Seabold S, Perktold J. 2010. Statsmodels: econometric and statistical modeling with Python. In *Proceedings of the 9th Annual Scientific Computing with Python Conference*, Austin, TX, pp. 92–96. <https://doi.org/10.2580/Majora-92bf1922-011>
- Sima J, Chakraborty A, Dileep V, Michalski M, Klein KN, Holcomb NP, Turner JL, Paulsen MT, Rivera-Mulia JC, Trevilla-García C, et al. 2019.

- Identifying *cis* elements for spatiotemporal control of mammalian DNA replication. *Cell* **176**: 816–830.e18. doi:10.1016/j.cell.2018.11.036
- Suvakov M, Panda A, Diesh C, Holmes I, Abyzov A. 2021. CNVpytor: a tool for copy number variation detection and analysis from read depth and allele imbalance in whole-genome sequencing. *GigaScience* **10**: giab074. doi:10.1093/gigascience/giab074
- Takahashi S, Kyogoku H, Hayakawa T, Miura H, Oji A, Kondo Y, Takebayashi S-I, Kitajima TS, Hiratani I. 2024. Embryonic genome instability upon DNA replication timing program emergence. *Nature* **633**: 686–694. doi:10.1038/s41586-024-07841-y
- Vouzas AE, Gilbert DM. 2023. Replication timing and transcriptional control: beyond cause and effect - part IV. *Curr Opin Genet Dev* **79**: 102031. doi:10.1016/j.gde.2023.102031
- Wheeler E, Mickelson-Young L, Wear EE, Burroughs M, Bass HW, Concia L, Thompson WF, Hanley-Bowdoin L. 2025. A comparison of genomic methods to assess DNA replication timing. *Sci Rep* **15**: 17761. doi:10.1038/s41598-025-02699-0
- Xu S, Wang N, Zuccaro MV, Gerhardt J, Iyyappan R, Scatolin GN, Jiang Z, Baslan T, Koren A, Egli D. 2024. DNA replication in early mammalian embryos is patterned, predisposing lamina-associated regions to fragility. *Nat Commun* **15**: 5247. doi:10.1038/s41467-024-49565-7
- Zaghlool A, Ameer A, Nyberg L, Halvardson J, Grabherr M, Cavelier L, Feuk L. 2013. Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues. *BMC Biotechnol* **13**: 99. doi:10.1186/1472-6750-13-99
- Zhang S, Plummer D, Lu L, Cui J, Xu W, Wang M, Liu X, Prabhakar N, Shrinet J, Srinivasan D, et al. 2022. DeepLoop robustly maps chromatin interactions from sparse allele-resolved or single-cell Hi-C data at kilobase resolution. *Nat Genet* **54**: 1013–1025. doi:10.1038/s41588-022-01116-w
- Zhao PA, Sasaki T, 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. doi:10.1186/s13059-020-01983-8
- Zhu C, Yu M, Huang H, Juric I, Abnoui A, Hu R, Lucero J, Behrens MM, Hu M, Ren B. 2019. An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat Struct Mol Biol* **26**: 1063–1070. doi:10.1038/s41594-019-0323-x

Received October 6, 2025; accepted in revised form March 2, 2026.