Superresolution imaging reveals spatiotemporal propagation of human replication foci mediated by CTCF-organized chromatin structures

Qian Peter Su,a,b,1,2,3 Qiziqing Winston Zhaoc,d,e,2,3 Luming Meng,1 Miao Ding,a,b Weiwei Zhang,a,b Yongzheng Lia,b Mengzhu Liu,a,b Rongqin Li,a,b Yi-Qin Gaof, Xiaoliang Sunney Xieb,g,3, and Yujie Suna,b,3,4

*aState Key Laboratory of Membrane Biology, Peking University, 100871 Beijing, China; bBiomedical Pioneering Innovation Center, School of Life Sciences, Peking University, 100871 Beijing, China; cDepartment of Chemistry, National University of Singapore, Singapore 117543, Singapore; dCentre for Biomolecular Sciences, National University of Singapore, Singapore 117557, Singapore; eCentre for Bioimaging Sciences, National University of Singapore, Singapore 138672, Singapore; fInstitute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, 100871 Beijing, China; and gBeijing Advanced Innovation Center for Genomics, Peking University, 100871 Beijing, China

Edited by Melike Lakadamyali, University of Pennsylvania, Philadelphia, PA, and accepted by Editorial Board Member Stephen J. Benkovic April 27, 2020 (received for review January 26, 2020)

Mammalian DNA replication is initiated at numerous replication origins, which are clustered into thousands of replication domains (RDs) across the genome. However, it remains unclear whether the replication origins within each RD are activated stochastically or preferentially near certain chromatin features. To understand how DNA replication in single human cells is regulated at the sub-RD level, we directly visualized and quantitatively characterized the spatiotemporal organization, morphology, and in situ epigenetic signatures of individual replication foci (RFi) across S-phase at superresolution using stochastic optical reconstruction microscopy. Importantly, we revealed a hierarchical radial pattern of RFi propagation dynamics that reverses directionality from early to late S-phase and is diminished upon caffeine treatment or CTCF knockdown. Together with simulation and bioinformatic analyses, our findings point to a “CTCF-organized Replication Propagation” (CoREP) model, which suggests a nonrandom selection mechanism for replication activation at the sub-RD level during early S-phase, mediated by CTCF-organized chromatin structures. Collectively, these findings offer critical insights into the key involvement of local epigenetic environment in coordinating DNA replication across the genome and have broad implications for our conceptualization of the role of multiscale chromatin architecture in regulating diverse cell nuclear dynamics in space and time.

DNA replication | chromatin organization | superresolution microscopy | epigenetic environment | spatiotemporal dynamics

DNA replication is a process of fundamental importance that maintains the integrity of genomic information in eukaryotic cells across the generations. After decades of biochemical and genetic studies, we now possess a comprehensive understanding of the mechanistic details of DNA replication and the various components of the cellular machinery that execute this process (1, 2). In addition, we have also gained substantial insights into the genomic organization of replication in sequence space, in which replication is initiated at multiple loci along the chromosomes termed replication origins, which demarcate the genome into numerous replicons (3, 4). Such organization is manifested in physical space as thousands of spatially clustered sites or puncta within the nucleoplasm, termed replication foci (RFi) or “factories,” as revealed by fluorescence imaging (4–6). Recent advances in high-throughput sequencing techniques have also enabled studies of DNA replication on a genome-wide scale, leading to the discovery of a replication timing program in which different regions of the genome tend to replicate in a temporally specific manner (7–9). Such temporally distinct regions, termed replication domains (RDs), could be considered as the counterpart of RFi in sequence space (6). In addition, cells have been found to respond and adapt to cellular environments by modulating the timing of RDs for specific genes, although the mechanisms of such modulation remain largely elusive (10, 11).

Several previous studies have suggested that the initiation of DNA replication is regulated at the RD level (9, 12, 13), while within an RD the activation of replication origins is rather random (14, 15). However, direct measurements on spread-out DNA fibers by DNA combing experiments have shown that an RD can harbor several replicons that fire synchronously (10, 16), thereby implicating a possible activation mechanism mediated by the spatial organization of the local chromatin environment. Such a proposal is also in line with the fact that the genomic boundaries of RDs align well with those of topologically associating domains (TADs), a prevalent structural feature of eukaryotic chromatin (17, 18).

In light of these findings, a fuller understanding of replication activation at the sub-RD level requires in situ visualization and

Significance

DNA replication in mammalian cells takes place at numerous replication origins clustered into thousands of replication domains (RDs) across the genome, which must be coordinated precisely to ensure that genomic information is duplicated once and only once during each cell cycle. Currently, how replication origins are chosen for firing within individual RDs remains poorly understood. We performed a quantitative characterization of the in situ epigenetic environment of replication foci across S-phase using superresolution microscopy. Our findings reveal a distinct pattern of replication propagation and point to a nonrandom mechanism for replication activation at early S-phase. These findings could have wide-ranging implications for our understanding of how local chromatin structures control dynamic processes within the cell nucleus.


The authors declare no competing interest.

This article is a PNAS Direct Submission. M.L. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

Data deposition: The localization coordinates used to generate the STORM images in Figs. 1–5 as well as all codes have been deposited in Zenodo (DOI: 10.5281/zenodo.3840656). 1Present address: Institute for Biomedical Materials & Devices, Faculty of Science, University of Technology Sydney, Ultimo NSW 2007, Australia. 2Q.P.S. and Z.W.Z. contributed equally to this work. 3To whom correspondence may be addressed. Email: qian.su@uts.edu.au, zhaozw@nus.edu.sg, sunneyxie@biopic.pku.edu.cn, or sun_yujie@pku.edu.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2001521117/-/DCSupplemental.

characterization of individual DNA replication sites in the context of their surrounding chromatin environment, with single-cell sensitivity and ultrahigh spatiotemporal resolutions. Among the technical tools available to address this problem, sequencing-based methods are mostly limited to analyzing populations of cells, and thus are more suitable for studying replication at the RD level. On the other hand, the majority of previous imaging studies were limited in their spatial resolution as well as their characterization of related chromatin features and functions and were thus unable to provide details into the regulation of replication activation within RDs.

More recently, superresolution microscopy has emerged as a powerful tool for chromatin biology and has been used to study a variety of chromatin structures ranging from nucleosome heterogeneity to the morphology of chromatin DNA in different epigenetic states (19–22). In this study, we use stochastic optical reconstruction microscopy (STORM) to probe the spatiotemporal organization of DNA replication within individual RFi at subdiffraction-limit resolution in single human cells. Using metabolic labeling, we quantitatively characterized the morphology, distribution, sequence length, and replicon number of individual RFi across different stages of S-phase. We also characterized the in situ epigenetic signatures of individual RFi and demonstrated the correlation between replication timing and chromatin modifications and organization at the RFi level. Importantly, we observed an intriguing spatiotemporal propagation pattern within individual RFi, with DNA that are replicated earlier spatially surrounded by DNA that are replicated later during early S-phase. The radial directionality of such a propagation pattern is reversed in late S-phase and diminished upon treatment with caffeine (an inhibitor of checkpoint kinase) or knockdown of CTCF (a key organizer of chromatin architecture). Together with simulations and bioinformatic analysis, our results point to a “CTCF-organized REplication Propagation” (CoREP) model, in which DNA replication activates nonrandomly within individual RFi, mediated by CTCF-organized chromatin structures at early S-phase. These findings offer critical insights into the regulation of origin activation at the sub-RD level and have broad implications for the spatiotemporal coordination of DNA replication across the genome, which may be applicable to a variety of mammalian systems.

Results

Superresolution Imaging and Quantitative Characterization of RFi across S-Phase. To label newly replicated DNA in situ, we first synchronized HeLa-S3 cells to the G1/S boundary as previously described (23, 24). At specific stages of S-phase, a short 30-min pulse of dye-labeled dUTP (2′-deoxyuridine 5′-triphosphate) or EdU (5-ethyl-2′-deoxyuridine) was supplied to the cells (see Materials and Methods for details), which incorporated these analog molecules into the DNA synthesized during the pulse (25). Imaging the labeled RFi with STORM (SI Appendix, Fig. S1A) revealed a punctate distribution as well as changing physical morphology across S-phase (Fig. L4), in line with previous observations using immunogold electron microscopy (26), conventional optical microscopy (16, 27), and other superresolution microscopy methods (20, 28–31). However, the superior spatial resolution afforded by STORM (∼20 nm) enabled us to perform more accurate quantitative characterization of RFi, particularly for subsequent findings based on multicolor colocalization analysis within tens of nanometers.

In order to assign the single-molecule localizations detected to individual RFi in a robust and unbiased fashion, we adopted a recently developed segmentation strategy, SR-Tesseler (32) (Fig. 1B; see Materials and Methods for details). With the RFi obtained from the segmentation, we found that while RFi density peaks during early-mid S-phase and slowly declines toward late S-phase (Fig. 1C), the size of RFi increases steadily from a mean value of 275.7 nm in early S-phase to 582.7 nm in late S-phase (Fig. 1D and SI Appendix, Fig. S1B). As a proof of the validity of our approach, we also ascertained that our synchronization and labeling procedures minimally impacted the morphology of RFi as well as cell growth and division (SI Appendix, Fig. S1 C and D).

As a further validation of our superresolution imaging approach, we correlated the RFi observed in Cartesian space with the actual replicated DNA in sequence space using a DNA combing approach (33). By imaging single stretched DNA fibers in vitro, we found that each single-molecule localization in the STORM images corresponded to 51.6 ± 12.6 base pairs of replicated DNA (Fig. 2A and SI Appendix, Fig. S2A; see Materials and Methods for details). Since identical imaging conditions and analysis algorithm were applied consistently across all imaging experiments, both in situ and in vitro, any overcounting of localizations from the same molecule will be present to the same extent in the images of both RFi in situ and DNA fibers in vitro, thereby ensuring the validity and accuracy of our calibration between single-molecule localizations and RFi sequence length. Such calibration in turn allowed us to quantify the sequence length of RFi in situ (Fig. 2B). Importantly, we found that the sequence length of RFi produced during the 30-min pulse increases from a mean value of 177.8 kb in early S-phase to 950.0 kb in late S-phase, in good agreement with previous sequencing measurements of 400 to 800 kb per RD (17) and 880 kb per TAD (18).

Moreover, by varying the labeling pulse duration (from 10 min to 60 min), we monitored the rate of expansion of the size and sequence length of individual RFi across S-phase (Fig. 2 C–F and SI Appendix, Fig. S2B). We also estimated the in situ replication speed, defined as sequence length per replication focus per minute (dashed linear fits in SI Appendix, Fig. S2B). We found that while the size and sequence length of RFi are similar for short labeling duration (10 min), they increase much faster during late S-phase for longer labeling durations (30 and 60 min). In addition, using the replication fork speeds (defined as sequence length per fork per minute) obtained from previous DNA combing experiments (34), we further found that on average 3.8–9.8 replicons exist within each replication focus at different stages of S-phase (Fig. 2G), and the number of replicons per focus increases from early to late S-phase, thereby lending support to the “increasing efficiency model” of DNA replication (35, 36).

RFi across S-Phase Are Associated with Distinct Stage-Specific Epigenetic Signatures In Situ. Even though the relationship between DNA replication timing and their chromatin environment has been studied using sequencing and biochemistry approaches, the measurements were often performed on cells in G1-phase (10, 21, 37, 38), and thus lacking in situ spatial and dynamic information for S-phase. To investigate the dynamic changes in the local chromatin environment of replicating DNA at subdiffraction-limit resolution, we colabeled RFi and seven key nuclear markers associated with a variety of epigenetic signatures in situ using two spectrally distinct dyes, Alexa 647 and Cy3B, and performed two-color STORM imaging to quantify their colocalization (Fig. 3 A–G and SI Appendix, Fig. S3A). Among these markers, CTCF is known to regulate three-dimensional (3D) chromatin architecture by forming chromatin loops and define the boundaries between euchromatic and heterochromatic DNA (39–41). Histone modifications H3K27ace and H3K4me3 are associated with transcriptionally active chromatin, with the former being a marker for active enhancers and the latter for active promoters (42, 43). In contrast, nuclear laminalocalized lamin A/C has been suggested to mostly associate with transcriptionally repressed DNA (44), while SUZ12 constitutes part of the polycomb repressive complex 2 (PRC2)/EED-EZH2 complex that methylates Lys-9 (H3K9me) and Lys-27 (H3K27me) of histone H3, leading to transcriptional repression of target genes.
Finally, proliferating cell nuclear antigen (PCNA) is an essential component of the mammalian replication fork and acts both as a processivity factor for DNA polymerase δ as well as a recruiter of other proteins involved in DNA replication (47). Thus, it serves as a positive control together with H2B, which provides an indication of the general histone background of the cell nucleus.

Quantitative analysis of STORM images revealed that while the densities and sizes of the protein foci (PFI) do not vary significantly (Fig. 3 H and I), the extent to which RFI colocalize

Fig. 1. STORM imaging and quantitative characterization of RFI across S-phase. (A) Conventional and STORM images of RFI in single HeLa cell nuclei at five different stages (from left to right, each with 30 min of labeling) across S-phase. Insets (red boxes) show zoomed-in representative images of individual RFI at each stage. The orange bars at the top indicate the relative temporal position of the labeling periods during S-phase (blue bars, not drawn to scale). (B) Unbiased hierarchical identification of RFI with SR-Tesseler, starting from raw single-molecule localizations in STORM images, to polygons in SR-Tesseler, to polygons upon using density-based filtering, and finally to RFI. Insets (red boxes) show zoomed-in areas where individual foci can be discerned. (C and D) RFI density (C) and diameter (D) at each of the five stages imaged in A. Data are shown as box-and-whisker plots. Each dot denotes a single RF, and each circle denotes a single cell. P values are determined by two-tailed Student’s t test; **** P < 0.0001; NS: not significant. n = 10–13 cells for C and 3,503, 4,556, 3,829, 2,566, and 1,509 RFI from 10–13 cells at each stage for D. Source data are provided as Dataset S1.
with each marker changes drastically across S-phase (Fig. 3J and SI Appendix, Fig. S3B). The relatively constant high colocalization ratio between RFi and PCNA throughout S-phase reflects the essential role of PCNA in DNA replication, and along with histone H2B lends support to the accuracy of our colocalization analysis. Specifically, the fractions of RFi that colocalize with H3K27ace, CTCF, or H3K4me3, all of which associate with the more open and transcriptionally active chromatin, show a marked decrease from early to late S-phase. In contrast, the fractions of RFi that colocalize with lamin A/C or SUZ12, which are involved in lamina-associating domains and transcriptional repression (48), display a pronounced increase from early to late S-phase. These results indicate that DNA near open and transcriptionally active chromatin tend to be replicated during early S-phase, while those near closed and transcriptionally inactive chromatin tend to be replicated during mid and late S-phase. Collectively, our imaging results constitute an in situ characterization of the dynamic epigenetic signatures of RFi, providing critical insights into the...
Fig. 3. Dynamics of RFi-associated epigenetic signatures in situ across S-phase. (A–G) Colocalization analysis between RFi and seven key nuclear markers reveals distinct local chromatin states of RFi across S-phase. Cells with newly replicated DNA (pink) labeled at the beginning of early (Top), mid (Middle), and late (Bottom) S-phase were subsequently immunolabeled with antibody against each marker (green) and imaged with dual-color STORM. Contrast between the two colors is enhanced here for better visualization; white indicates colocalization between RFi and epigenetic marker. See SI Appendix, Fig. S3 for corresponding images with normal contrast as well as zoomed-in areas from each image. (H–J) Density (H) and diameter (I) of protein foci (PFi) associated with each of the seven markers imaged in A–G, as well as fraction of RFi that colocalize with the PFi of each marker (J) during early, mid, and late S-phase, respectively. Data are shown as box-and-whisker plots. Each dot denotes a single PF, and each circle denotes a single cell. P values are determined by two-tailed Student’s t test; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; NS: not significant. n = 1,197, 1,088, and 969 (PCNA); 907, 859, and 699 (H3K27ace); 2,117, 1,712, and 1,570 (CTCF); 972, 765, and 608 (H3K4me3); 522, 473, and 431 (Lamin A/C); 253, 451, and 675 (SUZ12); and 1,804, 2,263, and 2,450 (Histone H2B) PFi from 5–8 cells at each stage for I and 5–8 cells for H and J. Source data are provided as Dataset S1.

<table>
<thead>
<tr>
<th>PCNA</th>
<th>Active</th>
<th>Repressive</th>
<th>Histone H2B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H3K27ace</td>
<td>CTCF</td>
<td>H3K4me3</td>
</tr>
<tr>
<td>Early</td>
<td>RFI</td>
<td>PCNA</td>
<td>RFI</td>
</tr>
<tr>
<td>Mid</td>
<td>RFI</td>
<td>H3K27ace</td>
<td>RFI</td>
</tr>
<tr>
<td>Late</td>
<td>RFI</td>
<td>CTCF</td>
<td>RFI</td>
</tr>
</tbody>
</table>

5 µm
correlation between the timing of DNA replication and its local chromatin environment.

**The Spatiotemporal Dynamics of RFI Propagation Are Directionally Specific.** To probe the spatial organization and dynamics of individual RFI, we next set out to investigate the propagation pattern of DNA replication in space and time. To do so, newly replicated DNA during two consecutive 30-min time windows at three different stages of S-phase were labeled with two spectrally distinct dyes, Alexa 647 and Atto 550 (Materials and Methods and SI Appendix, Fig. S4A). In addition, lamin A/C in the nuclear lamina was immunofluorescently labeled with a third dye, Atto 488 (Fig. 4 A–C). Upon multicolor STORM imaging of the labeled cells, the relative spatial distributions of the foci during both time windows reveal a unique spatiotemporal pattern of RFI progression. In early S-phase, DNA synthesized during the second window occupies a larger area and envelopes DNA synthesized during the first window, a trend that could be better visualized by enhancing the contrast between the two colors or directly plotting the coordinates of RFI localizations output by SR-Tesseler without pixel rendering (Fig. 4 A, Insets). Such contrast enhancement serves solely as a visualization aid and does not introduce additional artifacts in STORM images (SI Appendix, Fig. S5); all subsequent quantitative analyses, however, were performed on coordinates of raw localizations output by SR-Tesseler. Three-dimensional STORM imaging excluded the possibility that the colocalization pattern observed in two-dimensional images results from structures that did not overlap in the z-direction (SI Appendix, Fig. S4B). Significantly, this spatial propagation pattern becomes less pronounced during mid S-phase (Fig. 4 B) and reverses its radial directionality in late S-phase (Fig. 4 C), during which the DNA synthesized in the first window envelopes the DNA synthesized in the second window and occupies a larger spatial spread.

Such propagation dynamics can be more quantitatively illustrated by measuring the sizes of RFI in both time windows (Fig. 4 D). To better quantify the radial directionality of RFI propagation, we further defined a wrapping index (WI) based on the extent of colocalization between the RFI from the two time windows as

\[
WI = \left[1 - \sqrt{\left(1 - \frac{S_{\text{st}}}{S_{\text{st}} - S_{\text{sd}}}\right)\left(1 - \frac{S_{\text{sd}}}{S_{\text{sd}} - S_{\text{st}}}\right)}\right] \times \frac{S_{\text{st}} - S_{\text{sd}}}{S_{\text{st}} - S_{\text{sd}}}
\]

where \(S_{\text{st}}\) and \(S_{\text{sd}}\) denote the areas occupied by the RFI from the first and second time windows, respectively, and \(S_{\text{st}}\) denotes their colocalized area. A WI of 1 indicates that the RFI in the first time window completely wraps around those in the second window, whereas a WI of -1 indicates the opposite (see SI Appendix, Fig. S4C for schematic illustration of this definition). From our STORM images, we found that the mean value of WI changed from -0.88 in early S-phase to +0.83 in late S-phase (Fig. 4 E), clearly illustrating the reversal in the radial directionality of RFI propagation. A similar reversal was observed when the order of the two dye-labeled nucleotides used was swapped (SI Appendix, Fig. S4 D and E), thereby eliminating the possibility that such a propagation pattern could be the consequence of labeling and detection artifacts associated with specific dyes. Moreover, when cells were treated with 5 mM caffeine, a known inhibitor of cell cycle checkpoint that shifts origin firing from late to early S-phase (49), the RFI distribution pattern in early S-phase became similar to that in mid S-phase, especially around nuclear periphery and nucleoli (Fig. 4 F). In addition, the size of RFI in both time windows decreased drastically (Fig. 4 G), and the radial propagation pattern observed previously for early S-phase was abrogated, as indicated by a shift in the mean value of WI from -0.84 to -0.09 upon caffeine treatment (Fig. 4 H).

**CTCF Regulates RFI Morphology and Spatiotemporal Propagation Dynamics.** In order to understand the mechanistic basis under-lying the unique pattern of RFI propagation, we remind ourselves of the prominent role that CTCF plays in regulating 3D genome architecture through forming chromatin loops and hypothesize that CTCF-organized loop structures could serve both as structural anchor for the replicating DNA as well as spatial constraints for their propagation. To that end, we knocked down CTCF expression in HeLa cells using RNA interference, which we confirmed at both messenger RNA and protein levels using qPCR, immunostaining, and Western blotting, respectively (Fig. 5 A and SI Appendix, Fig. S6). Down-regulation of CTCF enlarged early S-phase RFI size without significantly reducing their density (Fig. 5 B and C). More importantly, the radially specific propagation pattern of RFI described in Fig. 4 was diminished upon CTCF knockdown (Fig. 5 D), as evidenced by significant changes in the relative sizes of RFI (Fig. 5 E) as well as WI (Fig. 5 F), which exhibits a drastic reversal in mean value from -0.79 to +0.22. These results strongly suggest that CTCF and the chromatin loop structures it mediates act as a key regulator of RFI morphological and propagation dynamics.

To further correlate the RFI propagation dynamics mediated by CTCF-organized chromatin structures with replication timing, we mapped the replication timing profile along human chromosomes onto chromatin immunoprecipitation sequencing (ChIP-Seq) data for both control input and CTCF (Fig. 5 G and H; only a specific region along chromosome 10 is shown). The CTCF binding sequences exhibit a substantially higher level of correlation with the replication timing profile as compared to the control sequence (Fig. 5 I). These bioinformatic analyses, along with our superresolution imaging results, point to the critical role played by CTCF-organized chromatin structures in facilitating the activation of DNA replication and its radial propagation within individual RFI.

**CoREP Model for Nonrandom Replication Activation within Early S-Phase RD.** To account for the observed link between the propagation pattern of RFI and CTCF-mediated chromatin organization during early S-phase, we propose a spatiotemporal model, CoREP, for the nonrandom activation and propagation of DNA replication at the sub-RD level (Fig. 6). Our model is based on previous findings that an RD contains one or several TADs [each with a mean size of 880 kb (17)] and that each TAD contains 4–6 CTCF-organized DNA loops [each with a mean size of 185 kb (40)]. These loops, occupied by promoters, enhancers, and active genes, are organized at their anchor sites by CTCF and cohesin (Fig. 6 A). Our observed spatiotemporal pattern of RFI propagation suggests that replication activation events may take place preferentially at CTCF-mediated loop anchors within each RD, which then propagate to the periphery of the domain, thereby giving rise to the radially outward propagation pattern observed during early S-phase (Fig. 6 B). Alternatively, it is also possible that DNA replication origins could be selectively activated at the periphery of each RD, and chromatin structural reorganization subsequently drives the DNAs that are replicated earlier toward the interior of the domain and those replicated later to the exterior (SI Appendix, Fig. S7). While our current data cannot differentiate between these two models, they unambiguously argue against the random activation of replication origins within RD, regardless of the structural architecture adopted by the chromatin (Fig. 6 C and D).

To further validate our model, we simulated the propagation of RFI for each of these models and calculated the radius of gyration of the simulated RFI (see Materials and Methods and SI Appendix, Figs. S8 and S9 for details). We found that the experimentally
Fig. 4. RFI during early and late S-phase display opposing spatiotemporal patterns of propagation dynamics. (A–C) Newly replicated DNA during two consecutive 30-min windows (purple and green) at the beginning of early (A, 0–1.0 h), mid (B, 2.5–3.5 h), and late (C, 5.0–6.0 h) S-phase were labeled (purple: first window; green: second window; white: colocalized RFI), together with lamin A/C (blue) to demarcate the nuclear boundary of each cell. Multicolor STORM images of single nuclei are shown in comparison with conventional images (Insets) of the same nuclei. The purple and green bars at the top indicate the relative temporal positions of the labeling periods during S-phase (blue bars, not drawn to scale). Insets (numbered 1 through 6) show zoomed-in areas (in red boxes), displaying both normal contrast (Top) and enhanced contrast (Middle) between the two colors for better visualization, as well as direct plotting of coordinates of RFI localizations output by SR-Tesseler without pixel rendering (Bottom) for comparison; white indicates colocalization. (D and E) Diameter (D) and WI (E) for RFI produced during the two labeling windows at early, mid, and late S-phase. (F) Images of RFI replicated in two consecutive 30-min windows (purple and green) at the beginning of early S-phase without (control, Top) and with (Bottom) 5 mM caffeine treatment. (G and H) Impact of caffeine treatment on RFI diameter (G) and WI (H). Data are shown as box-and-whisker plots. Each dot denotes a single RFI. P values are determined by two-tailed Student’s t test; ***P < 0.001; ****P < 0.0001; n = 3,365 and 3,356 (early), 3,743 and 3,637 (mid), and 1,391 and 1,273 (late) RFI from 13 cells at each stage for D; 2,535, 3,241, and 1,180 RFI from 13 cells at each stage for E; 2,070 and 2,051 (− Caffeine) and 2,357 and 2,304 (+ Caffeine) RFI from 8 cells for G; and 1,446 and 1,954 RFI from 8 cells for H. Source data are provided as Dataset S1.
observed RFi size (Fig. 4D) agree much better with the simulated radius of gyration using the CoREP model as compared to those using random firing models (Fig. 6E). Collectively, these results strongly support a chromatin structure-mediated mechanism for the nonrandom activation and propagation of DNA replication.

**Discussion**

DNA replication in mammalian cells takes place at thousands of RDs across the genome and must be coordinated precisely to ensure the complete and faithful duplication of the genomic information in the cell. While previous studies have shed much light on the regulation of replication at the RD level and found the activation of replication to be correlated with a number of factors such as G-quadruplex motifs (14), chromatin modifications (10, 50, 51), transcriptional activity (9), and cis-regulatory elements (52), within an RD the choice of which replication origins to fire remains poorly understood. In this study, we fill a critical gap in our understanding of the spatiotemporal coordination of DNA replication by performing quantitative characterization of the in situ epigenetic environment of RFi in relation to replication activation and propagation across different stages of S-phase at the single-cell level. Our superresolution imaging data reveal a distinct radial pattern of replication propagation within individual RDs and strongly suggest a non-random selection mechanism for replication activation at early S-phase. The CoREP model we proposed further establishes the role that CTCF plays in this process and points to the key involvement of local chromatin structures in controlling replication activation at early S-phase.

While our findings are in line with previous reports of the role of CTCF and cohesin in organizing the genome into numerous "contact domains" (40, 41, 53, 54) as well as the enrichment of CTCF and cohesin at TAD boundaries (18) and replication sites (53), a recent study has found that CTCF depletion minimally impacts replication timing at the RD level (52). Such a seeming discrepancy could potentially be reconciled by the fact that the propagation pattern we have observed results from the cumulative DNA replicated in situ within individual RFi over a 30- to 60-min time window and thus may not fully correspond with the replication timing profile captured at rather different temporal and spatial (either Cartesian space or sequence space) scales using high-throughput sequencing. In fact, the distinct propagation pattern we revealed highlights the ability of in situ superresolution imaging in providing a powerful approach complementary to sequencing-based
methods for probing DNA replication at unprecedented level of detail. In summary, our study paves the way for further interrogations into the spatiotemporal organization and dynamics of DNA replication and other associated intranuclear processes (55). For instance, to pinpoint the precise relationship between the RFi observed in Cartesian space and the RDs in sequence space revealed by replication timing profiles, specific RD sequences could be colabeled using DNA fluorescence in situ hybridization and their colocalization with RFi quantitatively measured by STORM. The roles of CTCF, cohesin, as well as other factors implicated in replication timing during mid and late S-phase that have not been fully accounted for in our CoREP model, also warrant further systematic investigation, especially in light of the finding that cohesin and CTCF could differentially affect chromatin architecture (54). Moreover, as chromatin structures organized by CTCF/cohesin have been found to also spatially coordinate RNA polymerase II transcription (41), our findings also provide a framework for understanding the coupling between replication and transcription, since open chromatin around transcriptionally active promoters could facilitate the selection of replication origins in the vicinity, which might in turn explain the observation that transcriptionally active genes tend to replicate early while inactive ones tend to replicate late (41). In fact, replication initiation sites have often been found in close proximity to transcription initiation sites (14, 56). Combining the approaches developed here with those reported previously for probing RNA polymerase II-mediated transcription (57, 58) and transcription factor binding (59, 60) will allow us to elucidate the extent to which replication and transcription are correlated in space and time (61), thus shedding important light on the mechanistic cooperation between these two processes.

**Materials and Methods**

**Cell Culture and Synchronization.** HeLa-S3 immortalized cell line (PubMed ID: 5733811) was obtained from Dr. Wei Guo, University of Pennsylvania, Philadelphia, PA, and cultured in high-glucose Dulbecco’s modified Eagle’s medium (Invitrogen) supplemented with 10% (vol/vol) fetal bovine serum (Gibco), 100 U/mL penicillin (Thermo Fisher), and 100 mg/mL streptomycin (Thermo Fisher) at 37 °C with 5% CO2. The duration of the cell cycle was determined by counting the number of cells every 8 h for a continuous period of 72 h. Cells were synchronized to the G1/S boundary with sequential treatment of 2 mM thymidine (Sigma-Aldrich) for 15 h, normal culture medium for 10 h, and 2 µg/mL aphidicolin (Sigma-Aldrich) for 15 h. The synchronization efficiency was determined by measuring the cellular DNA content profile in a population of DAPI-stained cells, fixed at 30 min after release into S-phase, using fluorescence-activated cell sorter (MoFlo; Beckman Coulter) at excitation wavelength of 340 to 380 nm.
RNA Interference. Hela-S3 cells were transfected with the following small interfering RNA (siRNA) oligos (GenePharma) using Lipofectamine 2000 (Thermo Fisher) according to manufacturer’s instructions: Nonspecific control: CGUAUCGGGAAUUCUGGATT (sense) and UCCAGAUAAUCCCGGUAACGT (anti-sense); CTCF: GAAGGCGUCGGAAGAUAUTT (sense) and AAUUCUUAACGGCAGCUCTT (anti-sense).

Cells were harvested at 48 h after transfection, followed by Western blot and qPCR quantification of CTCF expression.

RFI Labeling. At specific times after releasing the cells from G1/S boundary, RFI in the cell nucleus were labeled by introducing pulses of thymidine analogs of different durations (10, 30, or 60 min). Dye-labeled dUTPs (either with Alexa 647 [Invitrogen] or Atto 550 [Molbi Tech]) were delivered into the nucleus via transfection using FuGENE 6 reagent (Promega), while EdU (Life Technologies), being uncharged, was directly added into the culture medium and subsequently conjugated with a dye using the Click-IT Edu Imaging Kit (Life Technologies) according to the manufacturer’s instructions. When performing two-color labeling, the cells were washed with culture medium three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.

Upon labeling, the cells were incubated with extraction buffer (0.1 M Pipes at pH 7.5, 1 M NaCl, 50 mM ethylenediaminetetraacetic acid, and 0.5% [vol/vol] sodium dodecyl sulfate), and diluted fourfold with unlabeled cells at the same concentration. Cells were then air-dried and fixed with 3:1 methanol/acetic acid (vol/vol), followed by then washing with PBS three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.

Upon labeling, the cells were incubated with extraction buffer (0.1 M Pipes at pH 7.5, 1 M NaCl, 50 mM ethylenediaminetetraacetic acid, and 0.5% [vol/vol] sodium dodecyl sulfate), and diluted fourfold with unlabeled cells at the same concentration. Cells were then air-dried and fixed with 3:1 methanol/acetic acid (vol/vol), followed by then washing with PBS three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.

Upon labeling, the cells were incubated with extraction buffer (0.1 M Pipes at pH 7.5, 1 M NaCl, 50 mM ethylenediaminetetraacetic acid, and 0.5% [vol/vol] sodium dodecyl sulfate), and diluted fourfold with unlabeled cells at the same concentration. Cells were then air-dried and fixed with 3:1 methanol/acetic acid (vol/vol), followed by then washing with PBS three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.

Upon labeling, the cells were incubated with extraction buffer (0.1 M Pipes at pH 7.5, 1 M NaCl, 50 mM ethylenediaminetetraacetic acid, and 0.5% [vol/vol] sodium dodecyl sulfate), and diluted fourfold with unlabeled cells at the same concentration. Cells were then air-dried and fixed with 3:1 methanol/acetic acid (vol/vol), followed by then washing with PBS three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.

Upon labeling, the cells were incubated with extraction buffer (0.1 M Pipes at pH 7.5, 1 M NaCl, 50 mM ethylenediaminetetraacetic acid, and 0.5% [vol/vol] sodium dodecyl sulfate), and diluted fourfold with unlabeled cells at the same concentration. Cells were then air-dried and fixed with 3:1 methanol/acetic acid (vol/vol), followed by then washing with PBS three times after incubating with the first dye before adding in the second dye. For caffeine treatment, a final concentration of 5 mM caffeine (Sigma-Aldrich) was added into the culture medium 30 min before the release and during the 30-min labeling time.
Acknowledgments. We thank Prof. Xiaowei Zhuang and Dr. Hazen Babcock (Harvard University) for help with optical setup of the STORM microscope, Prof. Bo Huang (University of California, San Francisco) for the gift of the Insight software, and the Core Facilities at School of Life Sciences, Peking University for imaging support. This work is supported by Australia National Health and Medical Research Council Grant APP1177374 and Australia National Heart Foundation Grant 102592 to P.S.; National Medical Research Council of Singapore Open Fund-Young Individual Research Grant OFYR18G0000019 and National University of Singapore Presidential Young Professorship start-up funds to Z.W.Z.; Beijing Advanced Innovation Center for Genomics at Peking University to X.S.X.; National Key R&D Program of China Grant 2017YFA0503500; and National Science Foundation of China Grant 21825401 to Y.S.