bioRxiv, 2025, pp. 1–12

doi: xxxxxxx
xxxxxx
Paper

PAPER

Chromatin as a Coevolutionary Graph: Modeling the Interplay of Replication with Chromatin Dynamics

Sevastianos Korsak,[1,2,*] Krzysztof H. Banecki,[1,2] Karolina Buka,[2] Piotr J. Górski[3] and Dariusz Plewczynski[1,2,*]

[1]Laboratory of Bioinformatics and Computational Genomics, Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, 00-662, Warsaw, Poland, [2]Laboratory of Functional and Structural Genomics, Centre of New Technologies, University of Warsaw, Stefana Banacha 2c, 02-097, Warsaw, Poland and [3]Group of Physics in Economy and Social Sciences, Faculty of Physics, Warsaw University of Technology, Koszykowa 75, 00-662, Warsaw, Poland
[*]Corresponding authors. s.korsak@datascience.edu.pl, d.plewczynski@datascience.edu.pl
FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Modeling DNA replication poses significant challenges due to the intricate interplay of biophysical processes and the need for precise parameter optimization. In this study, we explore the interactions among three key biophysical factors that influence chromatin folding: replication, loop extrusion, and compartmentalization. Replication forks, known to act as barriers to the motion of loop extrusion factors, also correlate with the phase separation of chromatin into A and B compartments. Our approach integrates three components: (1) a numerical model that takes into advantage single-cell replication timing data to simulate replication fork propagation; (2) a stochastic Monte Carlo simulation that captures the interplay between the biophysical factors, with loop extrusion factors binding, unbinding, and extruding dynamically, while CTCF barriers and replication forks act as static and moving barriers, and a Potts Hamiltonian governs the spreading of epigenetic states driving chromatin compartmentalization; and (3) a 3D OpenMM simulation that reconstructs the chromatin's 3D structure based on the states generated by the stochastic model. To our knowledge, this is the first framework to dynamically integrate and simulate these three biophysical factors, enabling insights into chromatin behavior during replication. Furthermore, we investigate how replication stress alters these dynamics and affects chromatin structure.

Key words: DNA Replication, Loop Extrusion, Chromatin, 3D Simulation, Monte Carlo

Introduction

This study explores the multi-scale nature of chromatin, a dynamic and hierarchical structure fundamental to genome organization. Chromatin is organized across several levels, beginning with nucleosomes, the basic units of DNA packaging, and extending to loops and topologically associating domains (TADs), which are dynamic, non-equilibrium structures shaped by loop extrusion factors (LEFs) and boundary elements (BEs) (1–3). Beyond these, chromatin exhibits phase separation into A and B compartments, which can further subdivide into subcompartments (4–7), culminating in the spatial arrangement of chromosomal territories (8). The complexity inherent in these scales highlights the challenges of chromatin modeling, both from a physical and biological perspective. A key feature of chromatin, central to this study, is its stochastic nature: chromatin's structure is not only multi-scale (9–13) but also dynamic (14–16), with time-dependent processes continuously altering its organization. For instance, the loop extrusion mechanism involves LEFs binding and unbinding along the chromatin (1; 2), forming transient graphs of short- and long-range loops. This process embodies an interplay of order, driven by stabilizing proteins like CTCF, and randomness, arising from the stochastic binding, unbinding, and movement of LEFs. Moreover, compartmentalization has often been associated with the spread of epigenetic marks (17), and is often modeled with Potts-like models (18–24).

Within the field of chromatin modeling, a wide range of methodologies has been developed, each targeting different structural scales and biophysical processes, often with distinct objectives. A common class of models, known as *top-down* approaches, uses experimental data -typically Hi-C contact maps- to optimize bead-to-bead distances, ensuring agreement between simulations and observed data (25–34). In contrast, *bottom-up* approaches (13; 21; 35–44) aim to predict chromatin structure based on underlying physical laws and dynamical principles. Although bottom-up methods are more complex to implement and require deeper biophysical understanding, they often yield novel insights into the fundamental mechanisms of chromatin folding. At different levels of resolution, various modeling strategies exist: nucleosome-scale models often incorporate solvent effects, loop extrusion is typically modeled using stochastic frameworks (41–45),

© The Author 2025. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

compartments are represented through block-copolymer interactions (46–48), and specific interactions with the nuclear lamina are also considered. While some efforts have attempted to integrate these multiscale phenomena, such unified models remain rare due to the considerable complexity and number of assumptions required. In this study, we introduce a stochastic modeling framework designed to generate a statistical ensemble of chromatin structures, in the spirit of loop extrusion models. Notably, current models do not yet fully integrate loop extrusion dynamics, compartmentalization -interpreted as the spreading of epigenetic marks- and DNA replication via replication fork propagation. Our approach aims to address this gap.

The study of DNA replication has long captivated physicists, who have sought to develop analytical models of this essential biological process (49–56). These processes have been modeled using diffusion or transport differential equations, with researchers aiming to derive analytical solutions for various quantities of interest, such as the average length of replicated or non-replicated regions and the density of replication forks. However, these models typically rely on simplifications, such as assuming uniform replication fork speeds and homogeneous firing probabilities (50; 57). More recently, some studies have sought to model inhomogeneous scenarios (51), where replication variables change across the chromosome. Stochastic approaches, including Monte Carlo simulations, have been employed to simulate the stochastic firing of replication origins and the subsequent propagation of replication forks (49; 51; 52; 56). Although extensive research has been conducted on the physical modeling of replication, the relationship between the one-dimensional propagation of replication forks and the three-dimensional structure of chromatin remains largely unexplored. This intriguing connection has only recently garnered attention from a small group of researchers (58–60). Moreover, models that integrate additional biophysical processes, such as loop extrusion and compartmentalization, and explore their interplay with replication, are still lacking.

In this study, we introduce RepliSage, an extension of the LoopSage model (44), originally designed to simulate the process of loop extrusion. Our primary objective is to advance the understanding of DNA replication by demonstrating the feasibility of simulation frameworks that integrate replication dynamics with other biophysical processes, specifically loop extrusion and compartmentalization. Unlike traditional models that focus solely on the one-dimensional propagation of replication forks along the chromatin fiber, RepliSage also incorporates the dynamic alterations of the 3D chromosomal structure during replication. Therefore, chromatin can be modeled in two distinct ways: as a graph $\mathcal{G}$, where nodes and links possess specific biophysical attributes, or as a molecular dynamics model, which represents the 3D implementation of the graph by applying a force field that depends on the states of its nodes and links. In this perspective, RepliSage is a multi-scale stochastic model integrating a Monte Carlo energy framework that simultaneously modifies topology -via LEF bonds- and the node state of the chromatin graph (Fig. 1(c)), reflecting epigenetic compartmentalization. Such models are mathematically challenging, as they deviate from the common approach of fixing either node or link states (23; 24; 61–64). Only recent studies have begun to address the dynamics of graphs with evolving structure and state (65–67), which highlights the complexity of this type of modeling.

The RepliSage model consists of three distinct simulation components. First, replication timing data is imported to simulate the propagation of replication forks (*Replikator*). Next, the fork trajectories are fed into a stochastic model, derived from LoopSage, with additional energy terms for compartmentalization and DNA replication. This model operates on the graph's nodes and links, representing system states. Finally, RepliSage generates a pseudo-time trajectory of these states and imports it into an OpenMM potential to produce 3D structures.

Biological Background

DNA replication is a fundamental biological process through which the DNA polymer is duplicated to create two identical copies of genetic material. In eukaryotic cells, DNA replication starts at multiple sites called replication origins, which are progressively activated during the S phase of the cell cycle (68). Accurate regulation of origins is fundamental to ensure precise duplication of a genome. Errors during this process may result in genomic instability and lead to human disease including cancer (69).

The initiation of DNA replication is temporally separated into two steps, (1) origin licensing and (2) origin activation, which ensures that replication can only occur once per cell cycle (70). Origin licensing occurs in late mitosis/early G1 and consists of loading the MCM helicase into the origins in an inactive form. In the second step, which occurs in S phase, the active form of the helicase and other factors, including polymerases, form the replisome machinery responsible for DNA synthesis. Once initiated, two replisomes move away from the origin in a bidirectional fashion synthesizing the DNA. Replication is terminated when two converging replisomes meet or when a replisome reaches the end of a chromosome (71).

In mammalian cells, it is estimated that only a small fraction of licensed origins is activated during S phase (72; 73). The rest are passively replicated by ongoing forks (74). The excess of potential replication start sites gives the cell flexibility in origin selection. In addition, they were shown to play an important role in maintaining genomic stability in situations of replication stress (75; 76). Activation of origins throughout S is temporally regulated by a replication timing (RT) programme that coordinates when certain genomic regions are replicated throughout S phase. Replication timing profiles reveal that replication can be classified into early and late replicating regions.

A fundamental concept in modeling DNA replication is the regulation of origin firing (56). The simplest assumption is that the firing of one origin occurs independently of others (51; 52; 57; 77). On the contrary, some studies report the 'domino effect' (78; 79) of origin firing, where the activation of one origin triggers the firing of nearby origins in spatial proximity. Supporting the idea of correlated firing is the concept of replication factories (80–82) - clusters of replication forks that initiate replication simultaneously in specific regions, suggesting a higher-order organization of synchronized activity.

When examining the interaction between replication forks and loop extrusion factors (LEFs) (83–89), two possible scenarios has been proposed. Early studies suggested that cohesins (acting as LEFs) act as barriers to replication forks (86–88). Nevertheless, recent findings suggest that replication forks may function as moving barriers for LEFs (84; 85). This conclusion is supported by experiments showing weakened TAD borders and loops during DNA replication in 3C-based experiments (84). This indicates that DNA replication introduces new barriers, reducing loop enrichment. Biophysically, this could mean that replication forks push LEFs, causing them to either sink or unbind from chromatin. An essential aspect of the interplay between loop extrusion and DNA replication is the comparison of their characteristic times. The speed of loop extrusion ranges from $u_{\mathrm{le}} \sim 0.5 - 2$ kb/sec (90; 91), whereas the characteristic speed for DNA replication is significantly slower, at $u_{\mathrm{rep}} \sim 2$ kb/min (92; 93). This stark difference in timescales suggests a clear separation into fast and slow processes.

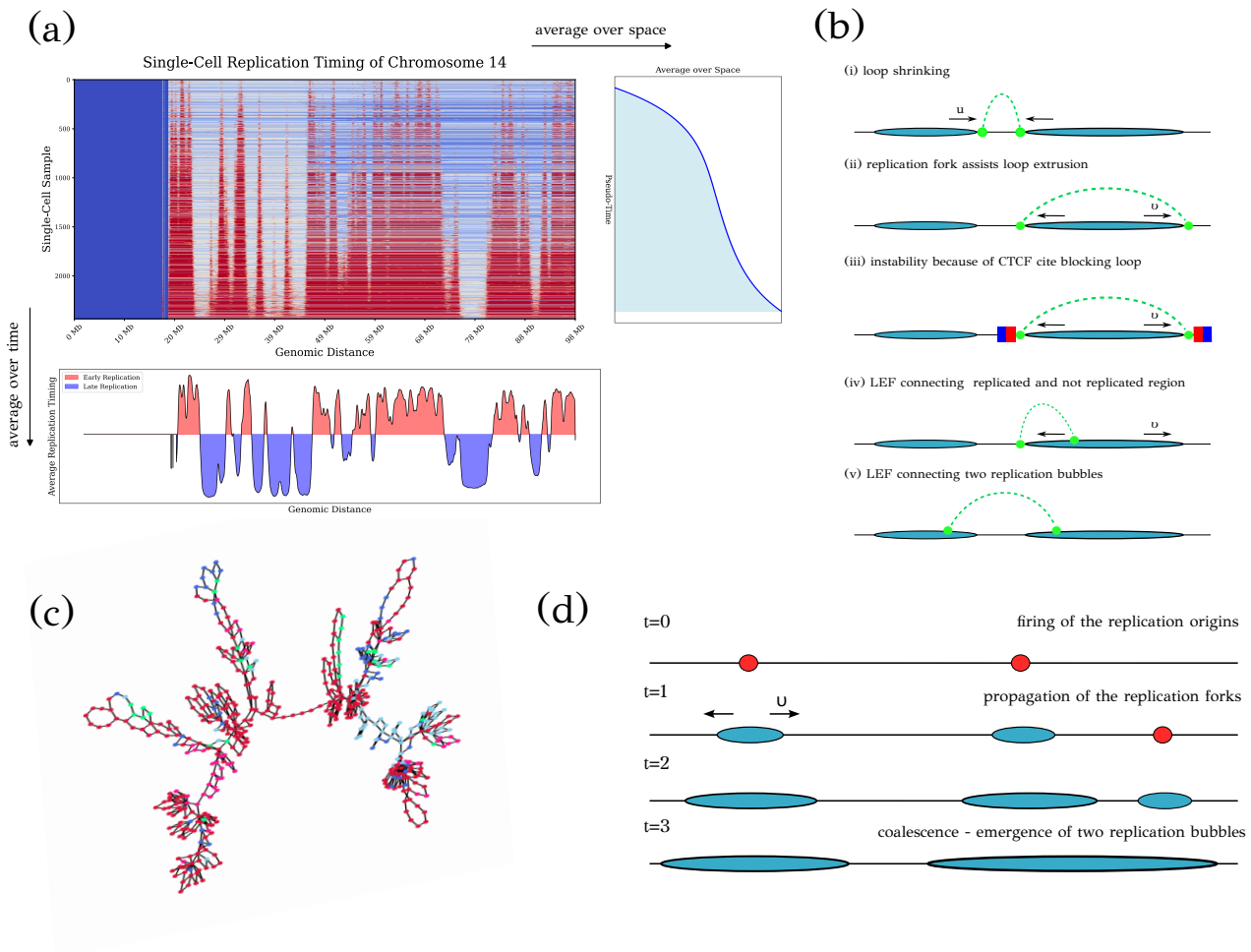

Fig. 1. Explanation of DNA replication. (a) Single-cell replication timing data. In this example we show the replication timing data of the chromosome 14. In a single sample of single cell, a particular loci can be either replicated (1) or not replicated (0). If we compute the average over time we should take the average replication timing which should be correlated with the replication fraction function $f(x, t)$. Averaging over space can produce a sigmoid-like curve, representing the transition from the unreplicated to the fully replicated state. (b) Cases of interaction between LEFs and replication forks: The first three cases are acceptable, where a LEF can bind to unreplicated regions, even if interrupted by a replicated segment. The third case illustrates a scenario where CTCF halts LEF propagation, potentially destabilizing the LEF to allow the replication fork to pass through, resulting in loop removal. The last two cases are more complex, as they involve interactions between replicated and unreplicated regions, which should not occur if replication forks act as moving barriers. Interactions between two replicated regions are even more puzzling. However, for this study, we assume they are not allowed, as a LEF initiating extrusion with zero loop length would unlikely connect two replicated regions without crossing the fork. (c) A chromatin coevolutionary graph, where links represent loops and node states compartments. (d) An example of how the DNA replication Monte Carlo simulation (Replikator) works: First, an origin has a probability of firing to initiate replication (red). The replication fork then propagates at speed v (blue). When two forks meet, coalescence occurs. Unreplicated regions between replication forks are referred to as *gaps*. The simulation concludes when all forks have merged, and the entire DNA is replicated.

Furthermore, A and B compartments are closely linked to early and late replication, respectively (94–97). This is significant for biophysical modeling because compartments exhibit long-range block-copolymer forces that cluster regions within the same compartment. Observations show that during replication, interactions related to compartmentalization are strengthened, whereas loop extrusion interactions become weaker (84). This suggests that replication forks may play a role in organizing compartmentalization.

Regarding replication stress (RS), it is a cellular condition in which replication forks slow down or stall in response to challenges to replication fork progression, such as: difficult-to-replicate sequences, obstacles caused by the transcription process, depletion of dNTP pools, and other circumstances (98). During RS, the cell alters its replication dynamics, compensating for the slowing replication fork speed by increasing origin activation to ensure complete replication (99). If prolonged, RS can be detrimental to cells, leading to the genomic instability that is a hallmark of cancer cells (100; 101).

Input Data

For input, two types of datasets were used: (i) single-cell replication timing data from the A. Koren laboratory (102), and (ii) loops from ChIA-PET experimental data filtered where there is a CTCF motif. Despite the fact that any dataset of loops derived from a 3C-type experiment can be used, in this study, we used the dataset

from ENCODE with codename ENCSR184YZV produced by the laboratory of Yijun Ruan.

Data Preprocessing

Before we run the replication simulation and RepliSage it is important to apply a pre-processing analysis so as to extract the parameters of interest. The implementation of a replication simulation is quite simple. Having the probabilities of an origin to fire $I(x,t)$ for a location x and time t, and the velocities of the replication forks v_k it is easy to run a simple Monte Carlo algorithm. However, the main difficulty of this task is to estimate the initiation rate probabilities $I(x,t)$. Therefore, we followed an approach similar to (49). Initially, we calculate the replication fraction $f(x,t)$ by computing the average of the single cell replication timing over the time (Fig.1(a)). We then assume that the initiation rate function follows a Gaussian distribution:

$$I(x,t) \propto \exp\left(-\frac{(t-\mu(x))^2}{2\sigma_I^2}\right), \tag{1}$$

where the parameter $\mu(x)$ represents the mean of normal distribution and is derived directly from the average replication timing curve, whereas the variance σ_I is constant and set proportionally to the *total replication time* T_r. In our replication model, the total replication time is specified as an input parameter. At first glance, this may seem counterintuitive, as one would expect the completion time to vary under different conditions of slow and fast replication. However, RepliSage is an equilibrium stochastic model. As we will see in the following sections, we allow sufficient steps both before and after replication to reach thermodynamic equilibrium. Therefore, *replication is primarily determined by the speed and number of replication forks*. The probabilities in $I(x,t)$ were later adjusted so as to take into account that the replication firing time was not drawn directly from the normal distribution but rather iteratively tested for potential firing over the course of simulation process. The adjustment was made so as to preserve the resulting firing time distribution same as in Equation (1).

The task of accurately estimating the parameters of the normal distribution only from the binary single cell states is extremely challenging. Therefore, our approach described above follows a very similar one described elsewhere (49). Consequently, we assume that our model is good enough as long as averaged independent experiments of the replication simulation, which we will describe in the following section, can reconstruct indeed the averaged replication timing curve.

The replication fork speed was calculated as follows. According to the paper (49) we can calculate the distribution of speeds directly from the slopes of the replication curve. Certainly, such an approximation is again a very simplistic one because the variability of the curve in certain places can be very high due to its complexity and possible experimental errors that may introduce discontinuities. In this case, we assume that the replication fork speed—meaning both sides of the fork move at the same rate—follows a normal distribution with mean μ_v and standard deviation σ_v:

$$p(v) \propto \exp\left(-\frac{(v-\mu_v)^2}{2\sigma_v^2}\right). \tag{2}$$

The parameters μ_v and σ_v are determined as the mean and standard deviation of the slopes between consecutive optima of the replication curve f. Despite the fact that our pre-processing approach is relatively simple, we were able to reconstruct the average replication curve only with these assumptions (Fig. 2).

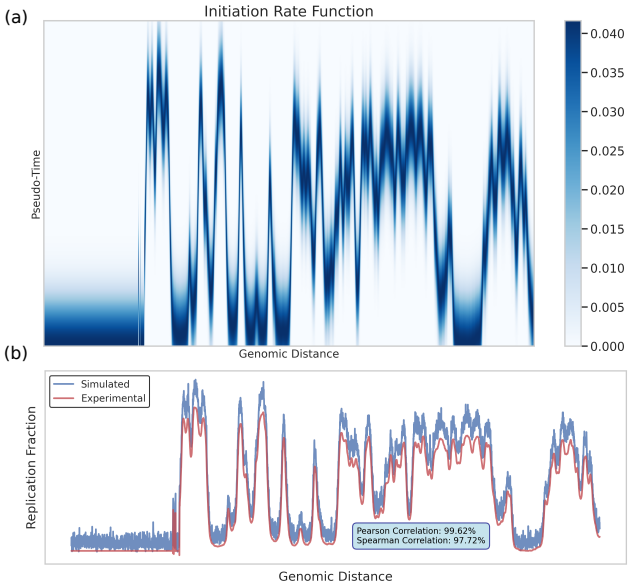

Fig. 2. Validation of the replication simulation for the human chromosome 14. (a) The initiation rate function is shown, peaking near the average replication timing and characterized by a Gaussian-like decision interval around this peak. (b) Validation of the replication simulation is primarily achieved by comparing the average replication timing curve with the replication fraction averaged over N_{trials} simulations. In this example, $N_{\text{trials}} = 100$ simulations were performed, yielding Pearson and Spearman correlations of 99.62% and 97.72%, respectively, with experimental data. This approach allows for stochastic simulations that produce realistic trajectories of replication fork dynamics.

Method

For the simulation of the interplay of replication with loop extrusion and compartmentalization, we used a complex model that can be split into three parts: a replication simulation, a stochastic loop extrusion simulation that takes into account the propagation of replication forks as moving barriers and a molecular dynamics OpenMM simulation for the production of the 3D structures. Each part of RepliSage is analyzed carefully in the following sections.

Replication Simulation (Replikator)

Initially, the initiation rate and replication fork speed are imported into the replication simulation. Replikator is a simplified model designed to simulate the stochastic propagation of replication forks (Fig.1(d)). The primary assumption of this model is the stochastic nature of origin firing. In a single run of Replikator, only a subset of origins initiates replication, while other origins are passively replicated. Despite this simplification, our objective is to ensure that the model accurately reconstructs the average replication timing curves used as input in RepliSage. Consequently, we have designed Replikator to produce highly correlated average replication fork trajectories that align with experimental data (Fig.2).

In each epoch of the simulation, we evaluate all monomers by drawing a random number in the interval $[0,1]$ for each monomer. If the number is less than the initiation rate $I(x,t)$ for the given time t, the Monte Carlo move is accepted, and the origin fires. Upon origin firing, we draw a velocity from the probability distribution $p(v)$ defined in Equation (2). In the second phase of the epoch, we propagate each replication fork a distance $r = v(t-t_0)$, where t_0 is the initiation time of a specific origin. The simulation terminates when the entire DNA is replicated. The trajectories of replication

forks and the simulation-estimated replication fraction function are obtained as outputs of the simulation.

Stochastic Model

The stochastic model's energy is composed by three factors,

$$\mathcal{H} = \mathcal{H}_{\text{le}} + \mathcal{H}_{\text{rep}} + \mathcal{H}_{\text{potts}}, \tag{3}$$

where each one of them aims to model different biophysics. In this sense, $\mathcal{H}_{le}$ factor is split into folding, binding and crossing term as it is explained in LoopSage (44). The term $\mathcal{H}_{\text{rep}}$ models the barrier activity of replication forks, and $\mathcal{H}_{\text{potts}}$ models their role as organizers of compartmentalization. In this paper, we take advantage of the Hamiltonian formalism, which allows us to explore the preferred states of the system by minimizing the functional of generalized energy which is called Hamiltonian.

A key assumption of our model is that the distribution of states, $\mathcal{S}$, follows a Boltzmann distribution:

$$p(\mathcal{S}) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{\mathcal{H}}{kT}\right) \tag{4}$$

where $\mathcal{Z}$ is the normalization constant, known in statistical physics as the *partition function*. The Boltzmann constant is set to $k = 1$ without loss of generality, while T represents the *temperature* or *order parameter* of the system. Although T plays a crucial role in system dynamics, its interpretation in biological systems remains elusive. In this study, we use T to parameterize the system's distribution: higher T results in greater randomness, whereas lower T leads to more ordered states. Assuming a Boltzmann distribution enables us to apply the well-established formalism of Monte Carlo methods.

Finally, we consider N_{lef} loop-extruding factors (LEFs), each defined by two degrees of freedom, (m_i, n_i), representing the link states of the graph, along with epigenetic color states s_i for each monomer. As previously discussed, chromatin compartmentalization is a biophysical phenomenon associated with regions sharing similar epigenetic marks (17). In this study, we interpret compartmentalization as a *macrostate* in the sense of statistical physics, emerging from underlying *microstates* defined by these local epigenetic features (103). The system state is thus given by $\mathcal{S} \equiv \{m_i, n_i\} \otimes s_k$. This notation implicitly accounts for the time dependence of states, as typically encountered in Monte Carlo simulations: $\mathcal{S}(t) \equiv \{m_i(t), n_i(t)\} \otimes s_k(t)$. The graph $\mathcal{G}$ consists of links either between consecutive nodes $i, i \pm 1$ and between nodes m_i and n_i due to LEF-mediated interactions. Node states s_k can take five discrete values, $\{0, \pm 1, \pm 2\}$, which depend on the graph's structure, determined by the distribution of links. The model's resolution is set by the user-defined number of beads, N_{beads}, ensuring that input data is downsampled to match this specified granularity.

Loop Extrusion with LoopSage

Since RepliSage is a fork of LoopSage (44), we adopt the same energy function for modeling loop extrusion. However, to simplify the notation, we omit the explicit time dependence. Consequently, we can write down the following formula,

$$\mathcal{H}_{le} = \mathcal{C}_{\text{fold}} \sum_{i=1}^{N_{\text{lef}}} \log(n_i - m_i) + \mathcal{C}_{\text{cross}} \sum_{i,j} K(m_i, n_i; m_j, n_j) + \mathcal{C}_{\text{bind}} \sum_{i=1}^{N_{\text{lef}}} \left(L(m_i) + R(n_i)\right). \tag{5}$$

The meaning of this formula is that the energy is minimized in two ways: (i) by loop extrusion because of the first term ($\mathcal{C}_{\text{fold}} < 0$), or (ii) by CTCF binding because of the last term ($\mathcal{C}_{\text{bind}} < 0$). The potentials $L(\cdot)$ and $R(\cdot)$ are minimum in location where there is a loop in experimental data, and specifically a CTCF loop. The remaining term is the crossing energy, which is a penalty ($\mathcal{C}_{\text{cross}} > 0$) in the case that two LEFs i and j cross each other $m_i < m_j < n_i < n_j$. RepliSage allows for the testing of hypotheses involving multiple populations of loop extrusion factors (LEFs). In particular, it is possible to introduce a secondary population of LEFs, denoted $N_{\text{lef},2}$, characterized by a distinct folding factor $\mathcal{C}_{\text{fold},2}$. This extension is especially relevant for modeling the mitotic phase of the cell cycle, which follows DNA replication. During mitosis, condensin complexes bind to chromatin and act as more processive LEFs, extruding longer loops at higher speeds compared to their interphase counterparts.

Replication and Compartmentalization

The replication-related energy term, $\mathcal{H}_{\text{rep}}$, is a time-dependent component that models the interaction between LEFs and replication forks. Specifically, Fig.1(b) illustrates various configurations of LEFs encountering replication forks. RepliSage assigns penalties to configurations (iv) and (v), as these configurations are physically prohibited when LEFs are unable to traverse replication forks. Furthermore, we assume that the whole information of replication is embedded into the replication fraction function $f_{\text{rep}}(x, t)$ which is produced by a single run of a replication simulation and it is 1 if x is replicated at time t and 0 otherwise. Therefore, we can write,

$$\mathcal{H}_{\text{rep}}(t) = \mathcal{C}_{\text{rep}} \sum_{i=1}^{N_{lef}} \mathcal{R}(m_i, n_i; f_{\text{rep}}) \tag{6}$$

where the function $\mathcal{R}(\cdot)$ is a penalty that takes non-zero values in case that a cohesin crosses a replication fork or connects two distinct replicated locis,

$$\mathcal{R}(m_i, n_i; f_{\text{rep}}) = \begin{cases} 1, & \text{if } f_{\text{rep},m_i} \neq f_{\text{rep},n_i} \\ 1, & \text{if } f_{\text{rep},m_i} = f_{\text{rep},n_i} = 1 \\ & \text{and } \exists\, \ell \in (m_i, n_i) : f_{\text{rep},\ell} = 0 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

and $\mathcal{C}_{\text{rep}} > 0$ is a normalization term. The biological meaning of this term is that a replication fork acts as barrier for the motion of LEFs and thus if a LEF will be located close to a replication fork, it should prefer to stay there and not propagate further. This term is particularly complicated because the replication fork trajectories (which are represented by $f_{\text{rep}}(t)$) derived from Replikator are time-dependent as well, and therefore the energy landscape changes over time.

Following Potts model, each node of the graph is in one of five epigenetic states $si \in \{0, \pm 1, \pm 2\}$. The node's state may change in time. States represents possible subcompartments. Our model differs from the traditional Potts model, because it includes an additional epigenetic field which depends on the replication timing,

$$\mathcal{H}_{\text{potts}} = \mathcal{C}_{p,1} \sum_k \left(\frac{h_k + h_{t,k}}{2}\right) s_k + \mathcal{C}_{p,2} \sum_{i>j} J_{ij} \left|s_i - s_j\right| \tag{8}$$

where h_k is an '*epigenetic field*' ($\mathcal{C}_{p,1} < 0$) which is proportional to the replication timing of the current region, normalized to take values from -1 to 1 and assume that A compartment is linked to the negative values of the field. The term $h_{t,k}$ represents the spread

of the epigenetic state due to the propagation of the replication fork and it is,

$$h_{t_r,k} = 2\sum_{t=0}^{t_r}\left(1 - 2\frac{t}{T_r}\right)\left(f_{\mathrm{rep},k}(t) - f_{\mathrm{rep},k}(t-1)\right) \quad (9)$$

where t_r is the relative time since the replication process started, and T_r is the total duration of replication. Before the replication starts this epigenetic field is zero, and the average epigenetic field h_k is the dominant one. The term $h_{t,k}$ depends on the propagation of the replication forks in the current stochastic experiment rather than their statistical average. It takes a positive value when the replication starts (early replication) and a negative value when the replication ends (late replication). The biological meaning of the term $h_{t,k}$ is that the replication forks are capable of spreading the epigenetic states of the system. This assumption is aligned with the study of C. Alabert and A. Groth (104) which states that despite the fact that the entire epigenome is stripped down and reassembled at the replication fork, it happens in a way that the epigenetic integrity is maintained. The final term ($\mathcal{C}_{p,2} > 0$) represents the interactions across the polymer graph assuming that J_{ij} takes value 1 in locations of LEFs $(m_i(t), n_i(t))$, and zero in location where are no LEFs. The difference $|s_i - s_j|$ increases sizes of subcompartments by penalizing interactions of polymers with varying states.

Parameter Balancing

In the updated version of LoopSage, we slightly modified the model parametrization:

$$\mathcal{C}_{\mathrm{fold}} = -\frac{N_{\mathrm{beads}} f}{N_{\mathrm{lef}}\log(N_{\mathrm{beads}}/N_{\mathrm{lef}})}, \quad (10)$$

$$\mathcal{C}_{\mathrm{bind}} = -\frac{N_{\mathrm{beads}} b}{\sum_i \left(L(m_i) + R(n_i)\right)}, \quad (11)$$

$$\mathcal{C}_{\mathrm{cross}} = \kappa \times 10^4. \quad (12)$$

This re-parametrization ensures that when $f = b = \kappa = 1$, the energy terms are balanced, allowing for a more systematic tuning of parameters. The balancing is achieved by setting C_{fold} and C_{bind} to yield comparable equilibrium energies for $f = b = 1$, while κ takes high values so as to penalize improbable configurations of LEFs. The same principle is applied to subsequent energy terms.

Therefore, we can write for the replication and Potts model parameters,

$$\mathcal{C}_{\mathrm{rep}} = c_{\mathrm{rep}} \times 10^5 \quad (13)$$

$$\mathcal{C}_{p,1} = -c_{p,1} \quad (14)$$

$$\mathcal{C}_{p,2} = -c_{p,2} \quad (15)$$

the choice of the sign for the Potts model is made in such a way that the energy is negative so as to be minimized when two interacting beads have the same node state. The replication term is multiplied by a large value of 10^5 so as to penalize improbable configurations.

The algorithm

To define the algorithm of RepliSage we need to decompose the energy difference into two parts: parts related (i) to rewiring (moves of LEFs), or (ii) to changing the epigenetic color state. Therefore, we can write,

$$\Delta\mathcal{H} = \Delta\mathcal{H}_{\mathrm{link}} + \Delta\mathcal{H}_{\mathrm{nodes}}. \quad (16)$$

and thus we can easily handle the energies related with our Monte Carlo moves.

Therefore, the stochastic simulation of RepliSage can be summarized in the following steps,

1. The user defines the input `.bedpe` and replication curve data and from these data RepliSage determines the binding potentials and simulated replication fraction.
2. Initialize random initial configrations of LEFs (m_i, n_i) and states s_i.
3. Initialize the matrix J_{ij} so as to be equal to 1 when i and j are connected by a LEF, and 0 otherwise.
4. We choose with probability $p_1 = 0.5$ if it will be a node state change or rewiring move.
5. If a rewiring move was chosen, we may choose one LEF (m_i, n_i) and with probability $p_2 = 0.5$ either to (i) randomly slide $m_i \to m_i \pm 1$ and $n_i \to n_i \pm 1$ as random walk (with probability $p = 0.5$), or (ii) unbind and rebind in different random location. Then,
   - We check if the move minimizes the energy $\Delta\mathcal{H}_{\mathrm{link}} < 0$ and accept if the condition is true, or if it is not true with probability $e^{-\Delta\mathcal{H}_{\mathrm{link}}/kT}$.
   - If the move is accepted change the adjacency matrix J_{ij}.
6. In case that a node state change is chosen, then we propose a change in node states from the values $s_i' \in \{0, \pm1, \pm2\}$.
   - We compute the energy difference $\Delta\mathcal{H}_{\mathrm{node}}$.
   - Check if the move minimizes the energy $\Delta\mathcal{H}_{\mathrm{node}} < 0$ and accept if the condition is true, or if it is not true with probability $e^{-\Delta\mathcal{H}_{\mathrm{node}}/kT}$.
7. Repeat steps 4–6 for N_{steps} RepliSage epochs. In each epoch, perform N_{beads} steps to ensure that, on average, each node updates its node or link state at least once per epoch.

In our Monte Carlo simulation we use sampling frequency f_{mc} so as to avoid correlated models, and a burnin period $\tau_{\text{burn-in}}$ so as to reach the equillibrium. The tricky part of RepliSage is that it is an out-of-equillibrium simulation in the sense that when the replication forks start their action, the system moves to a different state after the replication. Therefore, we expect that at least during replication RepliSage acts out of equillibrium.

Molecular Dynamics with OpenMM

The molecular modeling approach assumes two polymer chains, each composed of N_{beads} monomers, where N_{beads} defines the granularity of the stochastic simulation. The total potential energy of the system is given by:

$$U = U_{\mathrm{bk}} + U_{\mathrm{le}}(t) + U_{\mathrm{rep}}(t) + U_{\mathrm{block}}(t), \quad (17)$$

with each term representing a specific physical contribution. The backbone potential, U_{bk}, includes strong covalent bonds between consecutive beads ($k_{\mathrm{bk}} = 3 \times 10^5$ kJ/(mol · nm^2)) with an equilibrium distance $d_{\mathrm{bk}} = 0.1$ nm, angular forces ($k_\phi = 100$ kJ/(mol · rad^2), $\theta = \pi$), and an excluded volume term acting only within the same chain:

$$U_{\mathrm{ev}} = \epsilon\left(\frac{r_{\mathrm{ev}}}{r+\delta}\right)^3, \quad (18)$$

where $\epsilon = 100$ kJ/(mol · nm^2), $r_{\mathrm{ev}} = 0.1$ nm, and $\delta = 0.01$ nm.

The loop-formation potential, $U_{\mathrm{le}}(t)$, is time-dependent and models LEF-mediated loop formation via harmonic bonds:

$$U_{\mathrm{le}} = k_{\mathrm{le}}\sum_{i=1}^{N_{\mathrm{lef}}}(r_i - \ell)^2, \quad (19)$$

Parameter Name	Symbol	Units	Value
Temperature	T	K	310
Backbone bond strength	k_{bk}	$\mathrm{kJ/(mol \cdot nm^2)}$	3×10^5
Backbone bond length	d_{bk}	nm	0.1
Backbone angle force constant	k_ϕ	$\mathrm{kJ/(mol \cdot rad^2)}$	100
Backbone angle equilibrium	θ	rad	π
Excluded volume strength	ϵ	$\mathrm{kJ/(mol \cdot nm^2)}$	100
Excluded volume radius	r_{ev}	nm	0.1
Excluded volume shift	δ	nm	0.01
LEF bond strength	k_{le}	$\mathrm{kJ/(mol \cdot nm^2)}$	5×10^4
LEF bond length	ℓ	nm	0.1
Replication bond strength (on)	k_{rep}	$\mathrm{kJ/(mol \cdot nm^2)}$	5×10^4
Replication bond strength (off)	k_{rep}	$\mathrm{kJ/(mol \cdot nm^2)}$	10^2
Replication bond length (on)	d_{rep}	nm	0.1
Replication bond length (mid)	d_{rep}	nm	$d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}$
Replication bond length (off)	d_{rep}	nm	$5d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}$
Compartment interaction range	σ	nm	$0.5d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}$
Compartment equilibrium distance	ℓ_b	nm	0.2
Compartment energy (state $s = 2$)	E_2	$\mathrm{kJ/(mol \cdot nm^2)}$	-0.1
Compartment energy (state $s = 1$)	E_1	$\mathrm{kJ/(mol \cdot nm^2)}$	-0.2
Compartment energy (state $s = 0$)	E_0	$\mathrm{kJ/(mol \cdot nm^2)}$	-0.3
Compartment energy (state $s = -1$)	E_{-1}	$\mathrm{kJ/(mol \cdot nm^2)}$	-0.4
Compartment energy (state $s = -2$)	E_{-2}	$\mathrm{kJ/(mol \cdot nm^2)}$	-0.5

Table 1. Default OpenMM model parameters used in the molecular simulation, along with corresponding symbols, units, and values.

where $\ell = 0.1$ nm is the equilibrium bond length and $r_i(t) = |\vec{r}_{m_i(t)} - \vec{r}_{n_i(t)}|$ is the distance between LEF-bound beads. These interactions are weaker than backbone bonds, with strength $k_{\mathrm{le}} = 5 \times 10^4$ $\mathrm{kJ/(mol \cdot nm^2)}$.

The replication potential, $U_{\mathrm{rep}}(t)$, maintains strong harmonic bonds between paired monomers i and $i + N_{\mathrm{beads}}$ across the two chains prior to replication. As replication progresses, these bonds weaken, allowing chain separation:

$$U_{\mathrm{rep}} = \sum_i k_{\mathrm{rep},i} \left(r_i - d_{\mathrm{rep},i}\right)^2, \tag{20}$$

where both the bond strength and equilibrium distance depend on the replication state of each monomer i:

$$k_{\mathrm{rep},i} = \begin{cases} 5 \times 10^4 \ \mathrm{kJ/(mol \cdot nm^2)}, & \text{if not replicated,} \\ 10^2 \ \mathrm{kJ/(mol \cdot nm^2)}, & \text{if replicated,} \end{cases} \tag{21}$$

$$d_{\mathrm{rep},i} = \begin{cases} 0.1 \text{ nm}, & \text{if not replicated,} \\ d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}, & \text{if replicated,} \\ 5d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}, & \text{after replication terminates.} \end{cases} \tag{22}$$

With this parameterization, DNA replication results in the gradual separation of the two polymer chains through the formation of long-range bonds, effectively modeling the de-attachment process. The factor $\sqrt{N_{\mathrm{beads}}}$ is used as an approximation of the characteristic length scale, based on the average distance of a random walk, which scales with the square root of the number of steps.

The epigenetic compartment potential (U_{block}) is a Gaussian-shaped potential (46–48) that enforces attraction between monomers belonging to the same compartment, based on their epigenetic state,

$$U_{\mathrm{block}} = E \exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{23}$$

where,

$$E = \sum_{s \in \{0, \pm 1, \pm 2\}} E_s \delta(s'_i - s)\delta(s'_j - s) \tag{24}$$

which takes the value E_s when both monomers are in the same compartment, i.e., $s'_1 = s'_2$. The interaction energy E_s is stronger for the B compartment, which correlates with late replication timing. The range of excluded volume interactions is defined as $\sigma = 0.5 \cdot d_{\mathrm{bk}}\sqrt{N_{\mathrm{beads}}}$, meaning that the interaction range is half the average end-to-end distance of a random walk. Accordingly, the width of the Gaussian potential is set to 2σ. We assume that the equilibrium distance for compartment interactions is twice the loop equilibrium distance, $\ell_b = 0.2$ nm, as compartmentalization reflects more distal, non-local interactions compared to loop extrusion. Compartment interactions are relatively weak, with A compartments considered more loose than B compartments (6; 105). We assign the following interaction strengths: $E_2 = -0.1$ $\mathrm{kJ/\left(mol \cdot nm^2\right)}$), $E_1 = -0.2$ $\mathrm{kJ/\left(mol \cdot nm^2\right)}$, $E_0 = -0.3$ $\mathrm{kJ/\left(mol \cdot nm^2\right)}$, $E_{-1} = -0.4$ $\mathrm{kJ/\left(mol \cdot nm^2\right)}$ and $E_{-2} = -0.5$ $\mathrm{kJ/\left(mol \cdot nm^2\right)}$.

Results

We investigated the region $[78835000, 97674700]$ of human chromosome 14 to study the relationship between DNA replication and chromatin dynamics. Simulations used two polymer chains of length $N_{\mathrm{beads}} = 2000$ under two conditions: regular replication and replication stress. For regular replication, $\sim 10\%$ of replication origins were fired. Under replication stress, initiation rates and standard deviations doubled, while replication fork speed was reduced by a factor of four. Each simulation ran for $N_{\mathrm{steps}} = 10^5$, with $\tau_{\mathrm{rep}} = 2 \times 10^4$ steps for system equilibration, $T_r = 4 \times 10^4$ steps for replication, and additional steps post-replication to reach equilibrium.

Figure 3(a) demonstrates that loop extrusion factor (LEF) trajectories remain stable under normal replication conditions, showing only minimal perturbations throughout the replication process. LEFs predominantly undergo diffusive motion and are halted primarily by two mechanisms: (i) encountering CTCF proteins, which function as directional barriers and are visualized as stable horizontal lines indicating prolonged LEF residence at

CTCF sites, and (ii) interactions with progressing replication forks, which can displace LEFs. Despite this interaction, replication forks do not substantially alter LEF dynamics. This limited influence arises from the spatiotemporal nature of replication forks - they act as transient boundaries in time and exert only localized spatial effects. Moreover, when two converging replication forks meet, the previously unreplicated region between them becomes fully replicated, and the forks cease to function as boundaries. Consequently, the impact of replication forks on LEF-mediated chromatin organization remains both temporally constrained and spatially restricted.

Consequently, Figure 3(b) reveals a pronounced increase in chromatin folding energy under replication stress. This increase is attributed to the barrier activity of replication forks, which hinder LEF progression, causing premature loop termination and a consequent reduction in the average loop size. The discrete steps observed in the Potts model energy at the initiation and completion of replication emphasize the regulatory influence of replication forks on the spatial organization of epigenetic states. Assuming a nonzero field parameter ($h_t \neq 0$), replication forks act as epigenetic organizers, thereby potentially driving the compartmentalization of chromatin. To evaluate the validity of this assumption, we examined changes in chromatin structure as a response to altered compartmentalization. Figure 3(c) illustrates a phase transition in both the gyration radius and the average epigenetic state across the replication event. Under stress conditions, early-activated replication origins produce higher average epigenetic states ($s_i > 0$), effectively diminishing the block-copolymer potential and resulting in a more decondensed chromatin conformation. This leads to a paradox: if replication fork trajectories are assumed to cause compartmentalization, then replication stress, which alters fork dynamics, should in turn modify compartmentalization and the global replication timing program. This contradiction suggests two possibilities: either replication forks do not serve as primary propagators of epigenetic marks, or their role in driving chromatin compartmentalization is secondary rather than causal.

To investigate whether spontaneous changes in the average epigenetic state can occur in the absence of an external epigenetic field, we set $h_t = 0$. The results indicate that the interaction matrix J_{ij} alone -modulated by changes in loop distribution caused by replication fork activity- is insufficient to induce significant phase transitions. This outcome is attributed to the inherently local and transient nature of replication forks, which produce only minor perturbations in loop architecture. These findings support the conclusion that the role of replication forks as organizers of epigenetic marks should be considered independent from their function as physical barriers to loop extrusion factors (LEFs). Changes in the distribution of LEFs can only meaningfully affect the average epigenetic state if there are substantial shifts in either their loop length distributions or their spatial density along the chromosome.

Finally, Fig.3 illustrates changes in the 3D chromatin structure during replication. Initially, replication begins in the less condensed A compartment and progresses into the more condensed B compartment. Replication forks continue until two distinct copies of the polymer chain are formed.

Conclusion

In this study, we introduced a novel computational framework to model chromatin as a coevolutionary graph, in which both the epigenetic states of chromatin segments (nodes) and the physical connections between them (edges) evolve over time. This dual-level dynamical system captures the interplay between chromatin structure and epigenetic regulation, particularly during the complex biological process of DNA replication.

Our model offers several significant advancements over previous approaches. First, it is the first chromatin simulation framework to employ a Hamiltonian formalism that integrates multiple fundamental biophysical processes—compartmentalization, loop extrusion, and DNA replication—within a unified energy-based description. This formalism enables the generation of thermodynamically consistent ensembles of chromatin conformations, bridging the gap between local chromatin interactions and global architectural features.

A key innovation of our framework is its capacity to simulate the dynamical influence of replication forks on chromatin structure. While replication forks introduce transient local perturbations - acting as dynamic boundaries to loop extrusion factors (LEFs) - our results suggest that these effects alone are insufficient to induce large-scale epigenetic phase transitions. The study demonstrates that, in the absence of an external epigenetic field ($h_t = 0$), the interaction matrix J_{ij}, shaped by replication-driven loop configurations, fails to produce significant changes in global epigenetic organization. This observation underscores the limited role of replication forks as sole organizers of epigenetic states, highlighting the need for additional regulatory mechanisms or external fields to explain chromatin compartmentalization.

Moreover, our approach offers a flexible platform for testing a wide range of biological hypotheses. It supports modular modifications, enabling researchers to integrate new mechanistic details such as binding probabilities of loop extrusion factors, heterogeneous populations of LEFs with varying velocities, differential strengths of epigenetic mark propagation, frequencies of replication origin initiation, and the dynamic speed of replication fork progression. Importantly, the model accommodates both one-dimensional (1D) and three-dimensional (3D) representations of chromatin, offering outputs in the form of epigenetic state trajectories, graph topologies, and spatial conformations. This multimodal capability enables integration with tools from graph theory, polymer physics, and spatial modeling, promoting interdisciplinary applications in genomics and biophysics.

In summary, our framework provides a powerful and extensible tool to study chromatin dynamics during DNA replication. It not only captures key features of chromatin behavior—such as LEF-mediated looping and replication-induced structural remodeling—but also facilitates exploration of how local interactions propagate into large-scale genome organization. Future work may focus on incorporating additional molecular factors (e.g., cohesin dynamics, histone modifiers, transcriptional activity) and on refining model parameters using high-resolution experimental data. We believe this approach lays the groundwork for a deeper, mechanistic understanding of genome architecture and its regulation through the cell cycle.

Data Availability

RepliSage model can be found as an open-source project in `https://github.com/SFGLab/RepliSage`. The updated version of LoopSage can be found in `https://github.com/SFGLab/pyLoopSage` or via PyPI in `https://pypi.org/project/pyLoopSage/`.

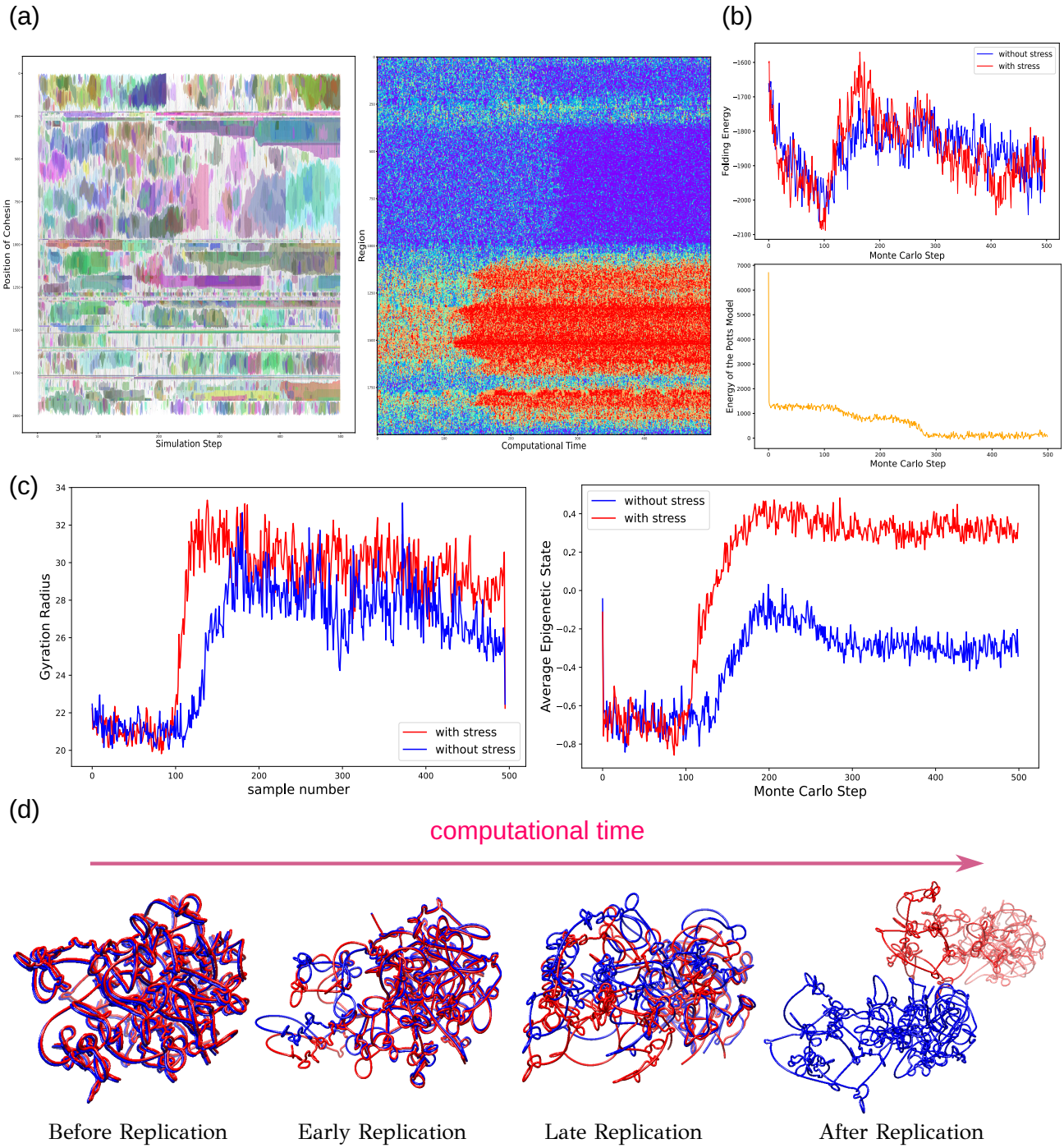

Fig. 3. Results of RepliSage for the region [78835000, 97674700] of chromosome 14: (a) Evolution of node and link states (y-axis) over computational time (x-axis). Links represent loops (left panel), and nodes represent compartments (right panel). In the left panel, each color denotes the extruded region of a specific LEF. In the right panel, red indicates the A compartment (positive epigenetic marks), while blue corresponds to the B compartment (negative epigenetic marks). (b) Folding energy -representing the entropic cost term in LoopSage that maximizes the average logarithm of loop length- and total Potts energy over time. Folding energy remains low before replication and increases during replication, as forks displace LEFs, leading to shorter loops. The Potts energy shows two distinct transitions, reflecting structural changes during DNA replication. (c) Structural metrics: radius of gyration and average epigenetic state. Under replication stress, the radius of gyration is slightly higher due to earlier replication of more regions. This is captured by a more positive average epigenetic state after replication under stress, as more loci reside in the A compartment. (d) Structural visualization across different replication phases. Each polymer chain is shown in a different color. Early replication targets the A compartment, followed by the B compartment.

Acknowledgments

This work has been supported by Polish National Science Centre (2020/37/B/NZ2/03757). Research was co-funded by Warsaw University of Technology within the Excellence Initiative: Research University (IDUB) programme. Computations were performed thanks to the Laboratory of Bioinformatics and Computational Genomics, Faculty of Mathematics and Information Science, Warsaw University of Technology using Artificial Intelligence HPC platform financed by Polish Ministry of Science and Higher Education (decision no. 7054/IA/SP/2020 of 2020-08-28). The work was co-supported by National Institute of Health USA 4DNucleome grant 1U54DK107967-01 and "Nucleome Positioning System for Spatiotemporal Genome Organization and Regulation".

References

1. A. S. Hansen, "Ctcf as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism," *Nucleus*, vol. 11, no. 1, pp. 132–148, 2020. PMID: 32631111.
2. A. Agarwal, S. Korsak, A. Choudhury, and D. Plewczynski, "The dynamic role of cohesin in maintaining human genome architecture," *BioEssays*, vol. 45, no. 10, p. 2200240, 2023.
3. M. Lazniewski, W. K. Dawson, A. M. Rusek, and D. Plewczynski, "One protein to rule them all: The role of ccctc-binding factor in shaping human genome in health and disease," *Seminars in Cell and Developmental Biology*, vol. 90, pp. 114–127, 2019. 3D Genome and Diseases.
4. K. Rippe, "Liquid–liquid phase separation in chromatin," *Cold Spring Harbor perspectives in biology*, vol. 14, no. 2, p. a040683, 2022.
5. F. Erdel and K. Rippe, "Formation of chromatin subcompartments by phase separation," *Biophysical journal*, vol. 114, no. 10, pp. 2262–2270, 2018.
6. J.-P. Fortin and K. D. Hansen, "Reconstructing a/b compartments as revealed by hi-c using long-range correlations in epigenetic data," *Genome biology*, vol. 16, pp. 1–23, 2015.
7. H. L. Harris, H. Gu, M. Olshansky, A. Wang, I. Farabella, Y. Eliaz, A. Kalluchi, A. Krishna, M. Jacobs, G. Cauer, *et al.*, "Chromatin alternates between a and b compartments at kilobase scale for subgenic organization," *Nature communications*, vol. 14, no. 1, p. 3303, 2023.
8. T. Cremer and M. Cremer, "Chromosome territories," *Cold Spring Harbor perspectives in biology*, vol. 2, no. 3, p. a003889, 2010.
9. G. Ozer, A. Luque, and T. Schlick, "The chromatin fiber: multiscale problems and approaches," *Current opinion in structural biology*, vol. 31, pp. 124–139, 2015.
10. C. Weinreb and B. J. Raphael, "Identification of hierarchical chromatin domains," *Bioinformatics*, vol. 32, no. 11, pp. 1601–1609, 2016.
11. G. Li and D. Reinberg, "Chromatin higher-order structures and gene regulation," *Current opinion in genetics & development*, vol. 21, no. 2, pp. 175–186, 2011.
12. C. L. Woodcock and R. P. Ghosh, "Chromatin higher-order structure and dynamics," *Cold Spring Harbor perspectives in biology*, vol. 2, no. 5, p. a000596, 2010.
13. S. Korsak, K. Banecki, and D. Plewczynski, "Multiscale molecular modeling of chromatin with multimm: From nucleosomes to the whole genome," *Computational and Structural Biotechnology Journal*, vol. 23, pp. 3537–3548, 2024.
14. B. Fierz and M. G. Poirier, "Biophysics of chromatin dynamics," *Annual review of biophysics*, vol. 48, no. 1, pp. 321–345, 2019.
15. M. R. Hübner and D. L. Spector, "Chromatin dynamics," *Annual review of biophysics*, vol. 39, no. 1, pp. 471–489, 2010.
16. E. I. Prieto and K. Maeshima, "Dynamic chromatin organization in the cell," *Essays in biochemistry*, vol. 63, no. 1, pp. 133–145, 2019.
17. M. Di Pierro, R. R. Cheng, E. Lieberman Aiden, P. G. Wolynes, and J. N. Onuchic, "De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture," *Proceedings of the National Academy of Sciences*, vol. 114, no. 46, pp. 12126–12131, 2017.
18. H. Zhang, X.-J. Tian, A. Mukhopadhyay, K. Kim, and J. Xing, "Statistical mechanics model for the dynamics of collective epigenetic histone modification," *Physical review letters*, vol. 112, no. 6, p. 068101, 2014.
19. Q. Mo and F. Liang, "A hidden ising model for chip-chip data analysis," *Bioinformatics*, vol. 26, no. 6, pp. 777–783, 2010.
20. Q. Mo and F. Liang, "Bayesian modeling of chip-chip data through a high-order ising model," *Biometrics*, vol. 66, no. 4, pp. 1284–1294, 2010.
21. D. Michieletto, E. Orlandini, and D. Marenduzzo, "Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3d organization of chromatin domains," *Physical Review X*, vol. 6, no. 4, p. 041047, 2016.
22. K. Adachi and K. Kawaguchi, "Chromatin state switching in a polymer model with mark-conformation coupling," *Physical Review E*, vol. 100, no. 6, p. 060401, 2019.
23. B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model*. Cambridge, MA and London, England: Harvard University Press, 1973.
24. L. Beaudin, "A review of the potts model," *Rose-Hulman Undergraduate Mathematics Journal*, vol. 8, no. 1, p. 13, 2007.
25. E. Lieberman-Aiden, N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, *et al.*, "Comprehensive mapping of long-range interactions reveals folding principles of the human genome," *science*, vol. 326, no. 5950, pp. 289–293, 2009.
26. G. Le Treut, F. Képès, and H. Orland, "A polymer model for the quantitative reconstruction of chromosome architecture from hic and gam data," *Biophysical journal*, vol. 115, no. 12, pp. 2286–2294, 2018.
27. S. Shinkai, M. Nakagawa, T. Sugawara, Y. Togashi, H. Ochiai, R. Nakato, Y. Taniguchi, and S. Onami, "Phi-c: deciphering hi-c data into polymer dynamics," *NAR genomics and bioinformatics*, vol. 2, no. 2, p. lqaa020, 2020.
28. S. Shinkai, H. Itoga, K. Kyoda, and S. Onami, "Phi-c2: interpreting hi-c data as the dynamic 3d genome state," *Bioinformatics*, vol. 38, no. 21, pp. 4984–4986, 2022.
29. C. Zou, Y. Zhang, and Z. Ouyang, "Hsa: integrating multi-track hi-c data for genome-scale reconstruction of 3d chromatin structure," *Genome biology*, vol. 17, pp. 1–14, 2016.
30. O. Oluwadare, M. Highsmith, and J. Cheng, "An overview of methods for reconstructing 3-d chromosome and genome structures from hi-c data," *Biological procedures online*, vol. 21, pp. 1–20, 2019.
31. L. Fiorillo, S. Bianco, A. M. Chiariello, M. Barbieri, A. Esposito, C. Annunziatella, M. Conte, A. Corrado, A. Prisco, A. Pombo, *et al.*, "Inference of chromosome 3d structures from gam data by a physics computational approach," *Methods*, vol. 181, pp. 70–79, 2020.
32. M. Di Stefano, J. Paulsen, T. G. Lien, E. Hovig, and C. Micheletti, "Hi-c-constrained physical models of human chromosomes recover functionally-related properties of genome organization," *Scientific reports*, vol. 6, no. 1, p. 35985, 2016.

33. D. Bau and M. A. Marti-Renom, "Genome structure determination via 3c-based data integration by the integrative modeling platform," *Methods*, vol. 58, no. 3, pp. 300–306, 2012.
34. G. Shi and D. Thirumalai, "A maximum-entropy model to predict 3d structural ensembles of chromatin from pairwise distances with applications to interphase chromosomes and structural variants," *Nature Communications*, vol. 14, no. 1, p. 1150, 2023.
35. M. Biswas, J. Langowski, and T. C. Bishop, "Atomistic simulations of nucleosomes," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 3, no. 4, pp. 378–392, 2013.
36. J. Huertas and V. Cojocaru, "Breaths, twists, and turns of atomistic nucleosomes," *Journal of molecular biology*, vol. 433, no. 6, p. 166744, 2021.
37. C. Zhang and J. Huang, "Interactions between nucleosomes: From atomistic simulation to polymer model," *Frontiers in Molecular Biosciences*, vol. 8, p. 624679, 2021.
38. M. Nicodemi and A. Prisco, "Thermodynamic pathways to genome spatial organization in the cell nucleus," *Biophysical Journal*, vol. 96, no. 6, pp. 2168–2177, 2009.
39. M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, "Complexity of chromatin folding is captured by the strings and binders switch model," *Proceedings of the National Academy of Sciences*, vol. 109, no. 40, pp. 16173–16178, 2012.
40. L. Liu, M. H. Kim, and C. Hyeon, "Heterogeneous loop model to infer 3d chromosome structures from hi-c," *Biophysical journal*, vol. 117, no. 3, pp. 613–625, 2019.
41. R. Rossini, V. Kumar, A. Mathelier, T. Rognes, and J. Paulsen, "Modle: high-performance stochastic modeling of dna loop extrusion interactions," *Genome Biology*, vol. 23, 11 2022.
42. E. J. Banigan and L. A. Mirny, "The interplay between asymmetric and symmetric dna loop extrusion," *eLife*, vol. 9, p. e63528, dec 2020.
43. E. J. Banigan, A. A. van den Berg, H. B. Brandão, J. F. Marko, and L. A. Mirny, "Chromosome organization by one-sided and two-sided loop extrusion," *eLife*, vol. 9, p. e53558, apr 2020.
44. S. Korsak and D. Plewczynski, "Loopsage: An energy-based monte carlo approach for the loop extrusion modelling of chromatin," *Methods*, 2024.
45. M. Crippa, Y. Zhan, and G. Tiana, "Effective model of loop extrusion predicts chromosomal domains," *Physical Review E*, vol. 102, no. 3, p. 032414, 2020.
46. D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, "Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains," *Nucleic acids research*, vol. 42, no. 15, pp. 9553–9561, 2014.
47. F. S. Bates and G. H. Fredrickson, "Block copolymer thermodynamics: theory and experiment," *Annual review of physical chemistry*, vol. 41, no. 1, pp. 525–557, 1990.
48. R. Zhou and Y. Q. Gao, "Polymer models for the mechanisms of chromatin 3d folding: review and perspective," *Physical Chemistry Chemical Physics*, vol. 22, no. 36, pp. 20189–20201, 2020.
49. R. Yousefi and M. Rowicka, "Stochasticity of replication forks' speeds plays a key role in the dynamics of dna replication," *PLoS computational biology*, vol. 15, no. 12, p. e1007519, 2019.
50. S. C.-H. Yang, M. G. Gauthier, and J. Bechhoefer, "Computational methods to study kinetics of dna replication," *DNA Replication: Methods and Protocols*, pp. 555–573, 2009.
51. M. G. Gauthier, P. Norio, and J. Bechhoefer, "Modeling inhomogeneous dna replication kinetics," *PLoS one*, vol. 7, no. 3, p. e32053, 2012.
52. A. P. de Moura, R. Retkute, M. Hawkins, and C. A. Nieduszynski, "Mathematical modelling of whole chromosome replication," *Nucleic acids research*, vol. 38, no. 17, pp. 5623–5633, 2010.
53. A. Baker and J. Bechhoefer, "Inferring the spatiotemporal dna replication program from noisy data," *Physical Review E*, vol. 89, no. 3, p. 032703, 2014.
54. R. Retkute, C. A. Nieduszynski, and A. De Moura, "Mathematical modeling of genome replication," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 86, no. 3, p. 031916, 2012.
55. H. Zhang and J. Bechhoefer, "Reconstructing dna replication kinetics from small dna fragments," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 5, p. 051903, 2006.
56. J. Bechhoefer and N. Rhind, "Replication timing and its emergence from stochastic processes," *Trends in Genetics*, vol. 28, no. 8, pp. 374–381, 2012.
57. J. Herrick, S. Jun, J. Bechhoefer, and A. Bensimon, "Kinetic model of dna replication in eukaryotic organisms," *Journal of molecular biology*, vol. 320, no. 4, pp. 741–750, 2002.
58. D. Löb, N. Lengert, V. Chagin, M. Reinhart, C. Casas-Delucchi, M. Cardoso, and B. Drossel, "3d replicon distributions arise from stochastic initiation and domino-like dna replication progression," *Nature communications*, vol. 7, no. 1, p. 11207, 2016.
59. J. H. Yang, H. B. Brandão, and A. S. Hansen, "Dna double-strand break end synapsis by dna loop extrusion," *Nature communications*, vol. 14, no. 1, p. 1913, 2023.
60. D. D'Asaro, M. M. Tortora, C. Vaillant, J.-M. Arbona, and D. Jost, "Dna replication and polymer chain duplication reshape the genome in space and time," *Physical Review X*, vol. 14, no. 4, p. 041020, 2024.
61. D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.
62. R. N. Onody and P. A. de Castro, "Nonlinear barabási–albert network," *Physica A: Statistical Mechanics and its Applications*, vol. 336, no. 3-4, pp. 491–502, 2004.
63. M. Barahona and L. M. Pecora, "Synchronization in small-world systems," *Physical review letters*, vol. 89, no. 5, p. 054101, 2002.
64. M. Kuperman and G. Abramson, "Small world effect in an epidemiological model," *Physical review letters*, vol. 86, no. 13, p. 2909, 2001.
65. T. Gross and B. Blasius, "Adaptive coevolutionary networks: a review," *Journal of the Royal Society Interface*, vol. 5, no. 20, pp. 259–271, 2008.
66. J. Toruniewska, K. Suchecki, and J. A. Hołyst, "Unstable network fragmentation in co-evolution of potts spins and system topology," *Physica A: Statistical Mechanics and its Applications*, vol. 460, pp. 1–15, 2016.
67. P. Fronczak, A. Fronczak, and J. A. Hołyst, "Self-organized criticality and coevolution of network structure and dynamics," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 4, p. 046117, 2006.
68. Y. Hu and B. Stillman, "Origins of dna replication in eukaryotes," *Molecular cell*, vol. 83, no. 3, pp. 352–372, 2023.
69. A.-S. Boyer, D. Walter, and C. S. Sørensen, "Dna replication and cancer: From dysfunctional replication origin activities to therapeutic opportunities," in *Seminars in cancer biology*, vol. 37, pp. 16–25, Elsevier, 2016.
70. A. Costa and J. F. Diffley, "The initiation of eukaryotic dna replication," *Annual review of biochemistry*, vol. 91, no. 1, pp. 107–131, 2022.

71. E. Zlotorynski, "A class of their own," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 2, pp. 71–71, 2017.
72. D. McIntosh and J. J. Blow, "Dormant origins, the licensing checkpoint, and the response to replicative stresses," *Cold Spring Harbor perspectives in biology*, vol. 4, no. 10, p. a012955, 2012.
73. M. Méchali, "Eukaryotic dna replication origins: many choices for appropriate answers," *Nature reviews Molecular cell biology*, vol. 11, no. 10, pp. 728–738, 2010.
74. S. Kagiwada, K. Kurimoto, T. Hirota, M. Yamaji, and M. Saitou, "Replication-coupled passive dna demethylation for the erasure of genome imprints in mice," *The EMBO journal*, vol. 32, no. 3, pp. 340–353, 2013.
75. A. Ibarra, E. Schwob, and J. Méndez, "Excess mcm proteins protect human cells from replicative stress by licensing backup origins of replication," *Proceedings of the National Academy of Sciences*, vol. 105, no. 26, pp. 8956–8961, 2008.
76. X. Q. Ge, D. A. Jackson, and J. J. Blow, "Dormant origins licensed by excess mcm2–7 are required for human cells to survive replicative stress," *Genes & development*, vol. 21, no. 24, pp. 3331–3341, 2007.
77. S. C.-H. Yang, N. Rhind, and J. Bechhoefer, "Modeling genome-wide replication kinetics reveals a mechanism for regulation of replication timing," *Molecular systems biology*, vol. 6, no. 1, p. 404, 2010.
78. F. G. Brunet, B. Audit, G. Drillon, F. Argoul, J.-N. Volff, and A. Arneodo, "Evidence for dna sequence encoding of an accessible nucleosomal array across vertebrates," *Biophysical Journal*, vol. 114, no. 10, pp. 2308–2316, 2018.
79. G. Guilbaud, P. Murat, H. S. Wilkes, L. K. Lerner, J. E. Sale, and T. Krude, "Determination of human dna replication origin position and efficiency reveals principles of initiation zone organisation," *Nucleic Acids Research*, vol. 50, no. 13, pp. 7436–7450, 2022.
80. P. Hozák and P. R. Cook, "Replication factories," *Trends in cell biology*, vol. 4, no. 2, pp. 48–52, 1994.
81. H. Leonhardt, H.-P. Rahn, P. Weinzierl, A. Sporbert, T. Cremer, D. Zink, and M. C. Cardoso, "Dynamics of dna replication factories in living cells," *The Journal of cell biology*, vol. 149, no. 2, pp. 271–280, 2000.
82. E. Guillou, A. Ibarra, V. Coulon, J. Casado-Vela, D. Rico, I. Casal, E. Schwob, A. Losada, and J. Méndez, "Cohesin organizes chromatin loops at dna replication factories," *Genes & development*, vol. 24, no. 24, pp. 2812–2822, 2010.
83. J. J. van Schie and J. de Lange, "The interplay of cohesin and the replisome at processive and stressed dna replication forks," *Cells*, vol. 10, no. 12, p. 3455, 2021.
84. T. Nagano, Y. Lubling, C. Várnai, C. Dudley, W. Leung, Y. Baran, N. Mendelson Cohen, S. Wingett, P. Fraser, and A. Tanay, "Cell-cycle dynamics of chromosomal organization at single-cell resolution," *Nature*, vol. 547, no. 7661, pp. 61–67, 2017.
85. M. Minamino, C. Bouchoux, B. Canal, J. F. Diffley, and F. Uhlmann, "A replication fork determinant for the establishment of sister chromatid cohesion," *Cell*, vol. 186, no. 4, pp. 837–849, 2023.
86. M.-E. Terret, R. Sherwood, S. Rahman, J. Qin, and P. V. Jallepalli, "Cohesin acetylation speeds the replication fork," *Nature*, vol. 462, no. 7270, pp. 231–234, 2009.
87. S. Villa-Hernández and R. Bermejo, "Cohesin dynamic association to chromatin and interfacing with replication forks in genome integrity maintenance," *Current genetics*, vol. 64, no. 5, pp. 1005–1013, 2018.
88. H. Bierne and B. Michel, "When replication forks stop," *Molecular microbiology*, vol. 13, no. 1, pp. 17–23, 1994.
89. K. E. Duderstadt, R. Reyes-Lamothe, A. M. van Oijen, and D. J. Sherratt, "Replication-fork dynamics," *Cold Spring Harbor Perspectives in Biology*, vol. 6, no. 1, p. a010157, 2014.
90. F. Corsi, E. Rusch, and A. Goloborodko, "Loop extrusion rules: the next generation," *Current Opinion in Genetics & Development*, vol. 81, p. 102061, 2023.
91. M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker, "Real-time imaging of dna loop extrusion by condensin," *Science*, vol. 360, no. 6384, pp. 102–105, 2018.
92. C. Conti, B. Saccà, J. Herrick, C. Lalou, Y. Pommier, and A. Bensimon, "Replication fork velocities at adjacent replication origins are coordinately modified during dna replication in human cells," *Molecular biology of the cell*, vol. 18, no. 8, pp. 3059–3067, 2007.
93. R. Milo and R. Phillips, *Cell biology by the numbers*. Garland Science, 2015.
94. C. Marchal, J. Sima, and D. M. Gilbert, "Control of dna replication timing in the 3d genome," *Nature Reviews Molecular Cell Biology*, vol. 20, no. 12, pp. 721–737, 2019.
95. P. A. Zhao, J. C. Rivera-Mulia, and D. M. Gilbert, "Replication domains: genome compartmentalization into functional replication units," *DNA replication: from old principles to new discoveries*, pp. 229–257, 2017.
96. T. Ryba, I. Hiratani, J. Lu, M. Itoh, M. Kulik, J. Zhang, T. C. Schulz, A. J. Robins, S. Dalton, and D. M. Gilbert, "Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types," *Genome research*, vol. 20, no. 6, pp. 761–770, 2010.
97. B. Moindrot, B. Audit, P. Klous, A. Baker, C. Thermes, W. de Laat, P. Bouvet, F. Mongelard, and A. Arneodo, "3d chromatin conformation correlates with replication timing and is conserved in resting cells," *Nucleic acids research*, vol. 40, no. 19, pp. 9470–9481, 2012.
98. H. Técher, S. Koundrioukoff, A. Nicolas, and M. Debatisse, "The impact of replication stress on replication dynamics and dna damage in vertebrate cells," *Nature Reviews Genetics*, vol. 18, no. 9, pp. 535–550, 2017.
99. M. K. Zeman and K. A. Cimprich, "Causes and consequences of replication stress," *Nature cell biology*, vol. 16, no. 1, pp. 2–9, 2014.
100. D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *cell*, vol. 144, no. 5, pp. 646–674, 2011.
101. M. Macheret and T. D. Halazonetis, "Dna replication stress as a hallmark of cancer," *Annual Review of Pathology: Mechanisms of Disease*, vol. 10, no. 1, pp. 425–448, 2015.
102. D. J. Massey and A. Koren, "High-throughput analysis of single human cells reveals the complex nature of dna replication timing control," *Nature Communications*, vol. 13, no. 1, p. 2402, 2022.
103. M. Kardar, *Statistical physics of particles*. Cambridge University Press, 2007.
104. C. Alabert and A. Groth, "Chromatin replication and epigenome maintenance," *Nature reviews Molecular cell biology*, vol. 13, no. 3, pp. 153–167, 2012.
105. A. Oji, L. Choubani, H. Miura, and I. Hiratani, "Structure and dynamics of nuclear a/b compartments and subcompartments," *Current Opinion in Cell Biology*, vol. 90, p. 102406, 2024.