Nucleic Acids Research, 2026, **54**, gkag618
https://doi.org/10.1093/nar/gkag618
Methods

RepliSage: a stochastic graph-based framework for 3D chromatin modeling across the cell cycle

Sevastianos Korsak [1,2], Krzysztof H. Banecki [1,2], Abhishek Agarwal [2], Joanna Borkowska [2], Piotr J. Górski [3], Haoxi Chai [4], Yijun Ruan [4], Karolina Buka [2,*], Dariusz Plewczynski [1,2,*]

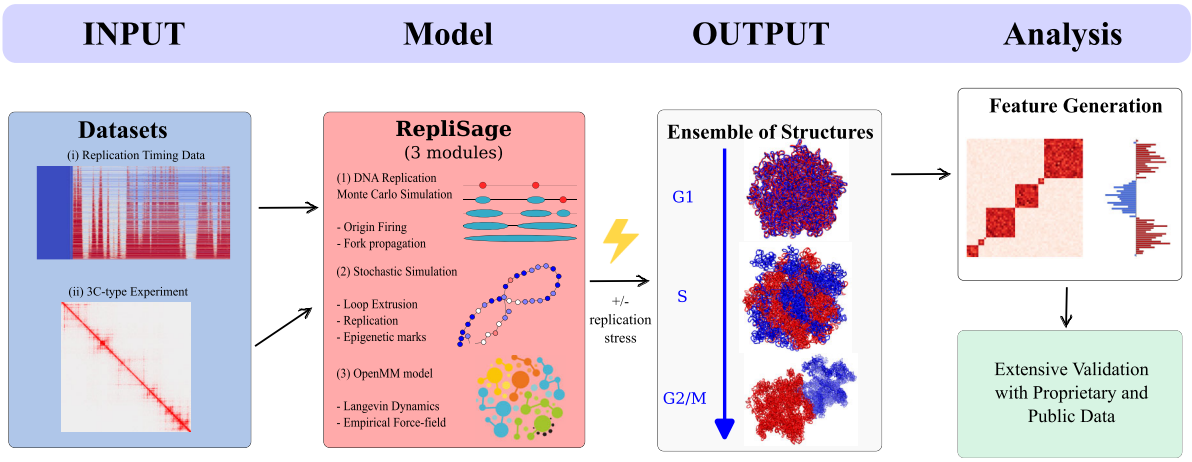
[1]Warsaw University of Technology, Faculty of Mathematics and Information Science, Laboratory of Bioinformatics and Computational Genomics, 00-662 Warsaw, Poland
[2]University of Warsaw, Centre of New Technologies, Laboratory of Functional and Structural Genomics, 02-097 Warsaw, Poland
[3]Warsaw University of Technology, Faculty of Physics, Group of Physics in Economy and Social Sciences, 00-662 Warsaw, Poland
[4]Zhejiang University, Life Sciences Institute and The Second Affiliated Hospital, Zhejiang 310058, China

*To whom correspondence should be addressed. Email: d.plewczynski@cent.uw.edu.pl
Correspondence may also be addressed to Karolina Buka. Email: k.buka@cent.uw.edu.pl

Abstract

Understanding chromatin dynamics across the cell cycle is crucial, as the structural transitions of chromosomes are fundamental to processes including transcriptional regulation, DNA replication, and faithful chromosome segregation. Although chromatin undergoes extensive reorganization throughout the cell cycle, no existing biophysical model describes its transitions from G1 through S and G2 to the completion of mitosis. To address this limitation, we present *RepliSage*, a multi-scale framework that integrates three fundamental processes shaping chromatin architecture: DNA replication, loop extrusion, and compartmentalization. In our model, replication forks are modeled as dynamic barriers that interact with loop extrusion factors, altering chromatin architecture during S phase. The framework integrates three complementary components: (i) replication fork progression simulated from single-cell replication timing data, (ii) Monte Carlo modeling of loop extrusion and epigenetic state transitions, and (iii) 3D reconstruction in OpenMM. Unlike previous approaches, *RepliSage* captures chromatin dynamics across the full cell cycle. In G1, random loop extrusion dominates; during S phase, replication forks interact with extrusion factors; and in mitosis, condensins drive long-range loop formation, facilitating chromosome segregation and polymer compaction. Chromatin is represented as a dynamic graph whose node states and connectivity evolve continuously over time. By tuning parameters across phases, *RepliSage* reproduces known structural transitions and provides a mechanistic platform to investigate how replication stress perturbs genome organization. The model was extensively validated using both publicly available and proprietary datasets. To our knowledge, this is the first framework to dynamically couple DNA replication, loop extrusion, and compartmentalization throughout the entire cell cycle.

Graphical abstract



Received: August 25, 2025. **Revised:** May 26, 2026. **Accepted:** May 31, 2026

Introduction

This study explores the multi-scale nature of chromatin, a dynamic and hierarchical structure fundamental to genome organization. Chromatin architecture spans from nucleosomes, the basic units of DNA packaging, to higher-order features such as loops and topologically associating domains (TADs). These structures are thought to emerge from the process of loop extrusion, driven by loop extrusion factors (LEFs) and constrained by boundary elements (BEs) [1–3]. The main proteins involved in loop formation in mammalian cells are cohesin and CTCF (CCCTC-binding factor), which act as the LEF and the BE, respectively [1, 4]. Beyond these, chromatin exhibits phase separation into A and B compartments, which can further subdivide into subcompartments [5–10]. Finally, interphase chromosomes occupy specific regions of the nucleus, known as chromosome territories [11]. Understanding and modeling the coupling between these organizational scales remains a central challenge in chromatin biology, as it ultimately determines the 3D structure of the genome.

A key feature of chromatin, central to this study, is its stochastic nature: chromatin structure is not only multi-scale [12–16] but also dynamical [17–19], 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 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. Notably, cohesin itself exists in two functional states: an extrusion-competent pool that generates chromatin loops and a cohesion-establishing pool that tethers sister chromatids, each contributing differently to 3D genome organization [20–24]. Furthermore, compartmentalization is closely linked to the propagation of epigenetic marks [25] and is frequently modeled using Potts-like frameworks [26–32], in which polymer loci are represented as agents with discrete (binary or categorical) epigenetic states. Interactions among these loci allow local state information to spread along the polymer, leading to collective behaviors that correspond to compartment formation.

Chromatin structure undergoes dynamic changes throughout the cell cycle, becoming condensed during mitosis and decondensed during interphase [33]. Extensive chromatin reorganization occurs in the S phase, when DNA replication takes place [34, 35]. During DNA replication, chromatin is first disassembled ahead of the replication machinery, and then restored after the passage of the replication fork. The tight coordination of DNA replication with the complex process of epigenome maintenance ensures that genetic and epigenetic information is faithfully passed on to daughter cells [34]. In mammals, replication is initiated by the licensing of replication origins in the G1 phase, when inactive MCM helicases are loaded onto the origins. This is followed by their activation and DNA synthesis in the S phase [36, 37]. Importantly, a small proportion of the licensed origins fire, while the rest remain dormant [38, 39]. The activation of replication origins occurs in a specific temporal order known as replication timing (RT), which correlates strongly with chromatin modifications and 3D chromatin structure at the level of compartments. Early-replicating regions are typically associated with A compartments, while late-replicating regions correspond to B compartments [40].

Recent findings provide insights into the interplay between cohesin and replication forks. *In vitro* studies of cohesin–replisome interactions in the context of cohesion establishment show that multiple outcomes can occur upon encounter, including the replisome pushing cohesin along the DNA, the replisome passing through the cohesin ring, and unloading of cohesin from the DNA [41–43]. A recent study [44] in a prokaryotic system, specifically investigating loop-extruding cohesin, showed that its activity is frequently impeded by replication forks. Furthermore, the inactive MCM helicases can restrict cohesin mediated loop extrusion, as demonstrated in G1-phase–arrested mammalian cells [45]. Together, these findings suggest that the replication machinery can act as a barrier to LEFs. In line with this, the weakening of TAD borders during the S phase [35] further indicates that the replication process may hinder loop extrusion. Since replication forks progress significantly more slowly ($u_{\mathrm{rep}} \sim$ 1–2 kb/min [46–50]) than loop extrusion ($u_{\mathrm{le}} \sim 0.5$–2 kb/sec [51, 52]), the replisome could act as a "slow-moving barrier" that pushes approaching LEF along the DNA. A key challenge in chromatin biophysics is to develop a modeling framework that incorporates a unified description of loop extrusion, replication fork propagation, and compartmental organization, with the ultimate goal of modeling chromatin structural transitions throughout the cell cycle.

Various endogenous or exogenous factors can interfere with replication, causing replication forks to slow or stall. This condition is known as replication stress (RS) [53]. Sustained RS is commonly found in cancer cells and was proposed to be a driver of tumor development [54–56]. To ensure complete genome duplication upon RS, the cell compensates for altered replication dynamics by activating additional (dormant) origins [38, 39]. Under this condition, epigenetic reorganisation at the replication forks contributes to fork protection [57, 58]. The effects of RS on higher-order chromatin structure are poorly defined, highlighting the need for Chromatin Conformation Capture (3C), based experiments in this context. Moreover, the impact of RS on higher-order chromatin structure remains unexplored in current 3D computational models, motivating the development of a new approach.

At the onset of mitotic phase, DNA replication is complete, but sister chromatids are still intertwined and contain DNA entanglements, which must be actively resolved to allow the separation of sister-chromatids [59, 60]. There are two major factors involved in mitotic chromosome assembly: (i) condensin complexes, both I and II, which extrude loops more rapidly than cohesin, folding chromosomes into the elongated, coiled structures typical of mitosis [61–63]; and (ii) Topoisomerase II (TOP2A), which resolves DNA catenations by transiently cleaving and rejoining DNA strands, ensuring efficient disentanglement of sister chromatids [64–69]. Condensin and residual cohesin briefly coexist: cohesin maintains cohesion, particularly at centromeres [70], while condensin and TOP2A drive large-scale reorganization of replicated chromatin [60]. In this process, condensin can bypass cohesin complexes, allowing loop extrusion and chromatin compaction to proceed even when cohesin remains bound [63, 70]. This reorganization forms mechanically stable, coiled mitotic chromosomes, which facilitates the separation of sister chromatids.

Simulations of DNA replication have progressed from 1D stochastic models of origin activation and fork progression [71–81] to 3D frameworks that explicitly incorporate spatial organization, capturing phenomena such as replication

fork clustering and the progression of replication from early- to late-replicating regions [82–85]. However, these models generally do not account for the interplay between distinct biophysical processes, such as interactions between replication forks and LEFs or the coupling of epigenetic states to chromatin compartmentalization, and are typically restricted to a single phase of the cell cycle.

To address these limitations, we introduced RepliSage, an extension of LoopSage [86] that integrates DNA replication dynamics, loop extrusion, and chromatin compartmentalization across G1, S, and G2/M. Chromatin is represented as a biophysically attributed graph coupled to molecular dynamics, while a multiscale Monte Carlo framework updates topology and epigenetic states, capturing coevolutionary graph dynamics [87–89]. The model focuses on mammalian genome and comprises a replication fork propagation Monte Carlo simulation (Replikator), a stochastic graph-based loop extrusion simulation engine, and 3D structure reconstruction via OpenMM software [90, 91]. RepliSage is validated against diverse experimental data. It reproduces RT and Hi-C contact patterns, including the characteristic compartment structure. Single-cell Hi-C comparisons [92] show that its loop dynamics follow the same qualitative trends seen in experiments: with shorter loops occurring during S phase and the largest loops in G2/M phase. Benchmarks against NucDynamics [93] confirm that the resulting 3D conformations are geometrically consistent. Importantly, by performing cohesin HiChIP experiment in normal and stress conditions, we show for the first time that RS triggers decrease in average loop size. This structural response was captured effectively by RepliSage. The main advantage of our model over previous approaches is its ability to span all phases of the cell cycle, capturing the complex interplay of biophysical processes that shape 3D chromatin organization before and after replication. It is provided as a user-friendly tool, easily installable via PyPI.

Materials and methods

This section elaborates on the experimental procedures and offers an in-depth description of the RepliSage model methodology. RepliSage operates on two primary types of input data: (i) *RT profiles* that guide the progression of replication forks along the genome, and (ii) *chromatin loop datasets* derived from 3C-based experiments such as ChIA-PET, which define CTCF-mediated interactions. Understanding how these inputs shape simulation outcomes is essential for interpreting RepliSage results and for applying the framework to other genomic datasets.

Cell line and culture conditions

The human GM12878 lymphoblastoid cell line was grown in RPMI 1640 (Biowest) supplemented with 15% fetal bovine serum (Biowest) and 2 mM L-glutamine (Gibco). Cells were cultured at 37 °C in a humidified environment containing 5% CO_2. Mild RS was induced by 18 h treatment with either aphidicolin (0.4 μM) or hydroxyurea (0.2 mM). As controls, cells were treated with dimethylsulfoxide (DMSO) and water (H_2O), respectively. RS induction was confirmed by analyzing cell cycle distribution, apoptosis, and DNA damage using the Apoptosis, DNA Damage and Cell Proliferation Kit (BD Biosciences, AB_2869407) in combination with flow cytometry (Supplementary Fig. S1).

HiChIP

HiChIP experiment was performed using the Arima-HiC+ Kit (HiChIP User Guide for Mammalian Cells and Tissues using Transcription Factors and Histones, A101020) with modifications. Cell fixation was carried out using FA-EGS cross-linking protocol [94]. The chromatin immunoprecipitation (ChIP) step was performed with 10 μg of SMC1 antibody (Bethyl Laboratories, A300-055A). Library preparation was performed as described in the HiChIP library preparation protocol of the Arima-HiC+ Kit using the Swift Biosciences Accel-NGS 2S Plus DNA Library Kit. Sequencing was conducted on the Illumina NovaSeq 6000 platform at the Genomics Core Facility, CeNT, University of Warsaw. One biological replicate of the HiChIP experiment was generated for each condition.

Raw paired-end HiChIP reads were processed using the nf-HiChIP pipeline [94]. In this pipeline, the raw reads were aligned to the hg38 reference genome using the BWA-MEM algorithm (version 0.7.17). Subsequently, the aligned reads were filtered based on a mapping quality (MapQ) score >30 and deduplicated using SAMtools. For each sample, peak calling was performed with a q-value cutoff of 0.01 using the `--nomodel` option. Interaction calling was carried out using default parameters, which define an anchor width of 5 kilobases (kb).

RepliSage model structure

The RepliSage framework comprises three interconnected simulation modules: (i) Replication simulator (Replikator): takes RT data as input to model stochastic origin firing and the propagation of replication forks. (ii) Stochastic loop extrusion simulator: models the dynamic interplay between loop extrusion, replication forks, and epigenetic mark spreading. It uses replication fork trajectories from Replikator and CTCF locations from a BEDPE-formatted 3C-type experiment to generate ensembles of chromatin state graphs. (iii) Physical modeling module: based on OpenMM [90, 91], which reconstructs the 3D chromatin conformation. RepliSage supports simulation of different cell cycle phases through adjustable parameters. All components are fully reproducible and publicly available via GitHub and PyPI. For ease of use with CUDA, a Docker container is also provided, enabling third parties to deploy the model with minimal effort.

Replication timing data pre-processing

Before running the replication simulation in RepliSage, it is essential to perform a pre-processing step to extract the necessary parameters. We define the initiation rate $I(x,t)$ as the probability per unit time and per unit length that a replication origin at position x initiates replication at time t. Given this initiation landscape $I(x,t)$, together with the replication fork velocities v_k, the replication process can be simulated using a straightforward Monte Carlo algorithm. Following an approach similar to [72], we first compute the replication fraction $f(x,t)$ by averaging single-cell RT data over cells (Fig. 1), thereby constructing a *pseudo-bulk* representation of single-cell RT. We then assume that the initiation rate at each genomic position follows a Gaussian profile in time:

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

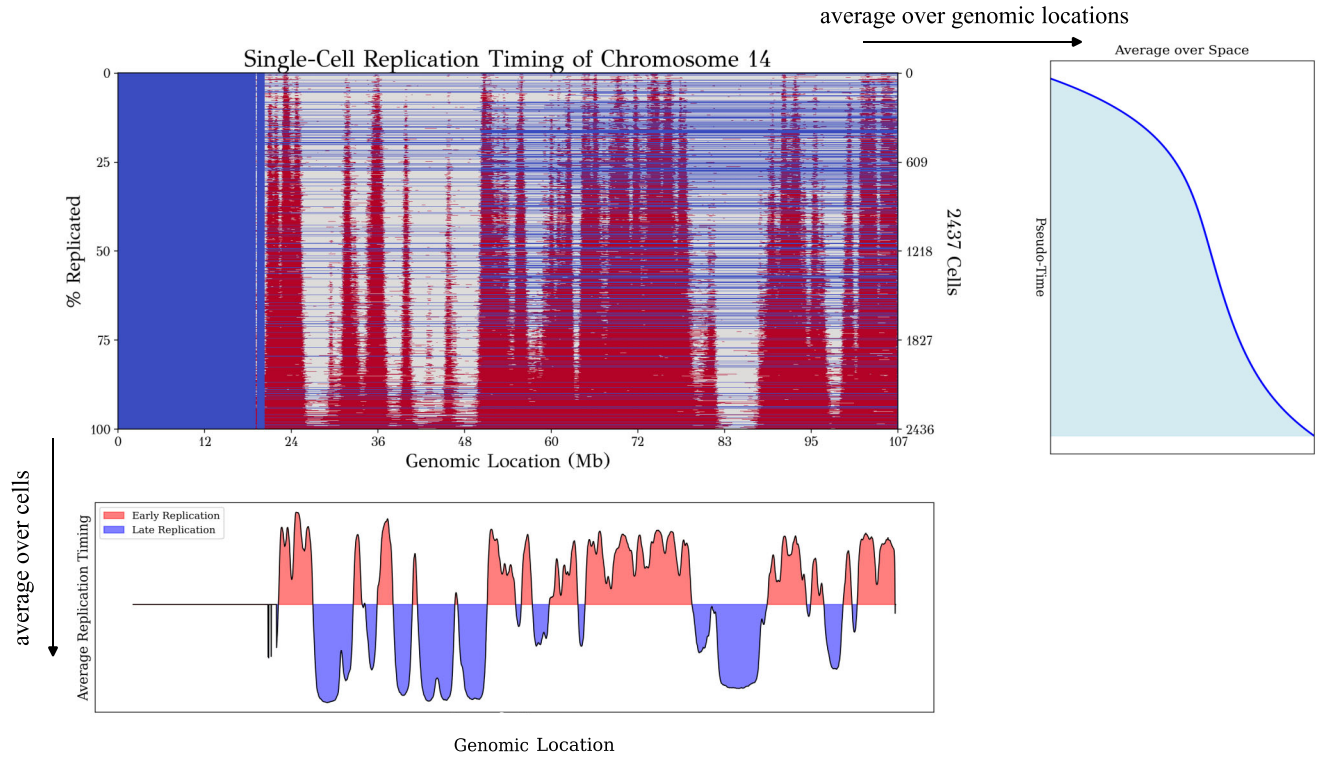


Figure 1. Single-cell RT data for chromosome 14. At each genomic locus, replication is represented as a binary state (0 or 1). Temporal averaging across cells yields a replication fraction function $f(x, t)$, whereas spatial averaging produces a characteristic sigmoid-like transition. Notably, RT is highly correlated with chromatin compartmentalization: loci in the A compartment tend to replicate early, while those in the B compartment replicate late.

where $\mu(x)$ denotes the mean replication time at position x, obtained directly from the average RT curve. The variance σ_I is assumed constant across the genome and is set proportionally to the total replication time T_r. In our replication model, T_r is provided as an input parameter. Consequently, the progression of replication is primarily controlled by the fork speed and the number of active replication forks. It is important to note that the initiation probabilities $I(x, t)$ are not sampled directly from the Gaussian distribution (Fig. 2A). Instead, replication origins are stochastically tested for firing at discrete time steps throughout the simulation. To ensure that this iterative sampling procedure yields a firing time distribution consistent with Equation (1), the values of $I(x, t)$ were renormalized accordingly so as that $\sum_t I(x, t) = 1$.

Single-cell RT experiments only indicate whether a chromatin locus has been replicated in a given cell, making it difficult to accurately estimate the parameters of the underlying normal distribution from these binary states alone. Additionally, quantifying the replication fork speed is required. To address this, we adopt an approach similar to that of Yousefi *et al.* [72]. Our model is considered adequate if the averaged results of independent replication simulations, described in the next section, can reliably reproduce the experimentally observed average RT curve.

Replication fork speeds can be estimated directly from the slopes of the replication curve [72], and in our model we assume that both sides of a fork move at the same rate following a normal distribution with mean μ_v and standard deviation σ_v:

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

The parameters μ_v and σ_v are calculated from the slopes between consecutive optima of the replication curve f, allowing us to reconstruct the average replication curve (Fig. 2B). Notably, single-cell and high-resolution studies [47, 48] show that fork speed varies dynamically during S phase, influenced by transcriptional activity, chromatin context, and RS. To create a simple, controllable working model, we prioritize the time-scale separation between loop extrusion and replication fork progression and do not incorporate temporal variability across S phase. This allows us to track and adjust the average fork speed directly, facilitating parameter studies in a systematic way. Variability is still captured by sampling fork speeds randomly across different genomic loci, and replication fork conditions under stress can be modeled by decreasing the average speed.

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. 2D). The primary assumption of this model is the stochastic nature of origin firing [95]. In a single run of Replikator, only a subset of origins initiates replication, while other origins are passively replicated. Our objective is to ensure that the model accurately reconstructs the average RT 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. 2B).

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 $I(x, t)$

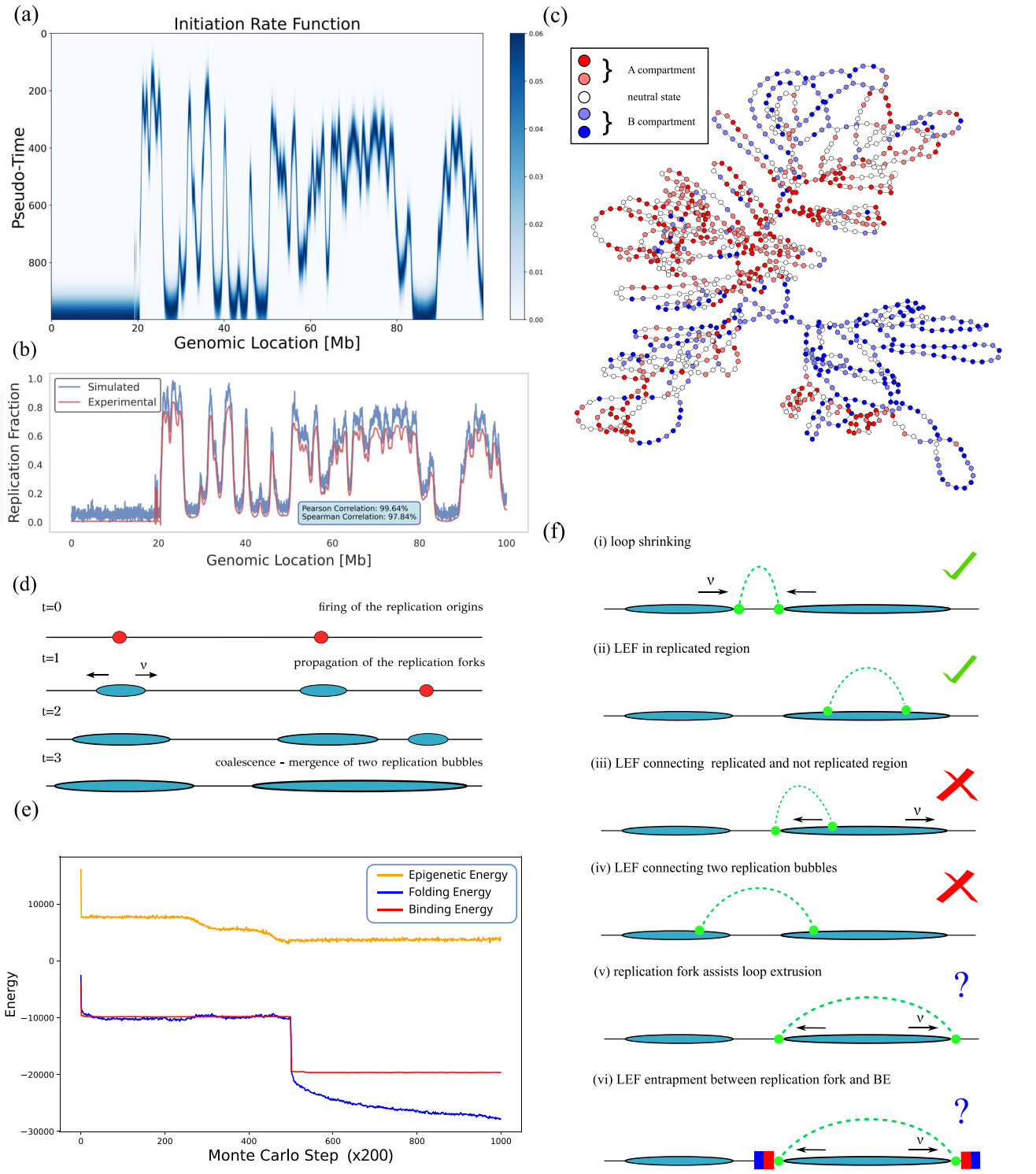


Figure 2. Overview of the RepliSage methodology. (**a**) The initiation rate function $I(x, t)$ follows a Gaussian distribution centered on the RT $\mu(x)$, defining the probability of origin firing over time. (**b**) Replication simulations produce average replication fraction curves highly correlated with experimental data. (**c**) Example chromatin graph where node colors indicate epigenetic states (blue: negative, red: positive) and links represent either sequential connections or LEF-mediated loops. Node states reflect epigenetic marks, whose collective behavior can drive compartment-like organization, though they do not directly correspond to compartments. (**d**) Schematic of stochastic replication: origins fire with probability $I(x, t)$, generating bidirectional forks. (**e**) Interplay of RepliSage energy components: loop extrusion and binding terms take similar values, whereas the epigenetic state term has an opposite sign. (**f**) Allowed and forbidden LEF–replication fork interactions in the model: (i) A LEF encountering a replication bubble is pushed by the replication fork. (ii) A LEF connecting two fragments within the same replication bubble is allowed. (iii) A LEF cannot connect a replicated region with a non-replicated region. (iv) A LEF attempting to connect separate replication bubbles should be prohibited, as this would imply extrusion through the unreplicated regions separating the two bubbles. (v) A LEF connecting two unreplicated regions with a replicated region in between is generally acceptable; however, if the replicated region is long, the probability decreases due to the temporal separation between replication and loop extrusion. (vi) If a LEF is trapped between an advancing replication fork and a BE, the fork-driven pushing while extrusion is blocked creates mechanical frustration, which is likely to promote LEF unbinding according to the assumptions of our model.

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 firing 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:

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

where each one of them aims to model different aspects of biophysics of chromatin. In this sense, $\mathcal{H}_{le}$ factor is related to loop extrusion and can be decomposed into folding, binding and crossing terms as it is explained in LoopSage [86]. The term $\mathcal{H}_{\text{rep}}$ represents the "barrier effect" imposed by replication forks [41–44], while $\mathcal{H}_{\text{epi}}$ captures the spreading of epigenetic states, which may contribute to chromatin compartmentalization. In this paper, we adopt a Hamiltonian formalism, which enables us to explore the system's preferred states by minimizing a generalized energy functional [96].

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

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

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_{mc} represents the *Monte Carlo temperature* or *order parameter* of the stochastic system. Although T_{mc} plays a crucial role in system dynamics, its interpretation in biological systems remains elusive. In this study, we use T_{mc} to parameterize the system's distribution: higher T_{mc} results in greater randomness, whereas lower T_{mc} leads to more ordered states. Assuming a Boltzmann distribution enables us to apply the well-established formalism of Monte Carlo methods [97].

Finally, we consider a system composed of N_{beads} monomers, each characterized by an epigenetic state s_k and N_{lef} LEFs. Each LEF is specified by two degrees of freedom, (m_i, n_i), representing the indices of monomers it connects. Chromatin compartmentalization is a biophysical phenomenon associated with contacting regions sharing similar epigenetic marks [25, 98, 99]. 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 [100]. The system configuration is thus given by $\mathcal{S} = \left(\{(m_i, n_i)\}_{i=1}^{N_{\text{LEF}}},\ \{s_k\}_{k=1}^{N_{\text{beads}}}\right)$. This notation implicitly accounts for the time dependence of states, as typically encountered in Monte Carlo simulations: $\mathcal{S}(t) = \left(\{(m_i(t), n_i(t))\}_{i=1}^{N_{\text{LEF}}},\ \{s_k(t)\}_{k=1}^{N_{\text{beads}}}\right)$. The graph $\mathcal{G}$ contains links connecting consecutive nodes $k, k \pm 1$ together with links between 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 (Fig. 2C). The model's resolution is set by the user-defined number of beads, N_{beads}, ensuring that input data are downsampled to match this specified granularity.

Loop extrusion with LoopSage

Since RepliSage is a fork of LoopSage [86], 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:

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

The meaning of this formula is that the energy is minimized in two ways: (i) by loop extrusion $\mathcal{H}_{\text{fold}}$ because of the first term ($\mathcal{C}_{\text{fold}} < 0$), or (ii) by BE binding $\mathcal{H}_{\text{bind}}$ because of the last term ($\mathcal{C}_{\text{bind}} < 0$). The potentials $L(\cdot)$ and $R(\cdot)$ reach their minima at positions corresponding to loops observed in experimental data, specifically at CTCF loops as defined in LoopSage [86], which are determined from a BEDPE file representing interactions measured by 3C-type experiments such as Hi-C, ChIA-PET, or Hi-ChIP. The remaining term is the crossing energy $\mathcal{H}_{\text{cross}}$, 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 LEFs. In particular, RepliSage can include a second population of loop-extruding factors, with size $N_{\text{lef}}^{(2)}$ and its own folding coefficient $\mathcal{C}_{\text{fold}}^{(2)}$. The two LEF populations are treated as independent [63, 70], such that the mixed crossing penalty between a LEF of population 1 and a LEF of population 2 is set to zero, $K_{\text{cross}}^{(1,2)} = 0$. This extension is especially useful for modeling mitosis, where condensin complexes function as a distinct, more processive class of LEFs, extruding longer loops at higher speeds compared to their interphase counterparts.

Replication

Evidence from studies investigating cohesin–replisome encounters during cohesion formation [41–43] and cohesin extrusion [44] suggests that the replication fork can impede the movement of cohesin. Moreover, in at least some experimental settings, this outcome of such encounters is predominant [42–44]. Since replication forks move much more slowly than extrusion, the replisome may therefore act as a "*slow-moving barrier*," pushing LEFs along the DNA. Based on these findings, RepliSage proposes that the replisome predominantly acts as a "physical barrier" for cohesin upon encountering a LEF.

Our model makes the following minimal assumption: loops are formed entirely within unreplicated or replicated regions [Fig. 2F, scenarios (i) and (ii)], since LEFs are unable to cross active replication forks [Fig. 2F, (iii) and (iv)]. Cohesin complexes typically unbind within 15–30 min [1, 101] much faster than the 8-h duration of S phase [102, 103]. Loop extrusion is also locally constrained by CTCF boundaries, so individual loops rarely span large regions. As replication progresses, a loop connecting two unreplicated regions is increasingly unlikely to contain long replicated segments. Because of this separation of timescales *between replication and extrusion rate*, in our model we assume that replication forks rarely remain inside a single loop for long. Consequently, LEFs spanning replicated regions are penalized proportionally to the length of the replicated segment [Fig. 2F, (v) and (vi)].

Therefore, 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:

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

where the function $\mathcal{R}(\cdot)$ is a penalty that takes non-zero values in the case that a cohesin crosses a replication fork or connects two distinct replicated loci (Fig. 2F):

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

where $\ell_{\text{rep}} = \left|\{\lambda \in (m_i, n_i) : f_\lambda > 1\}\right|$ represents the number of replicated loci within the loop, and $\mathcal{C}_{\text{rep}} > 0$ is a normalization term whose magnitude encodes the permeability of the replisome. In this sense, large values of C_{rep} correspond to a strong barrier effect of the replication forks, whereas smaller values reflect a weaker, more permissive obstruction. By preventing LEFs from connecting replicated and unreplicated regions [42, 45, 104, 105], replication forks effectively function as "moving barriers," directly resulting from this assumption. 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.

Although RepliSage primarily considers the replication fork to be a barrier to LEFs, it does not rule out the other possible outcomes of LEF–replisome interactions occurring simultaneously, such as cohesin unloading or bypass [41–44]. In this regard, RepliSage allows for a certain degree of flexibility to accommodate these alternative outcomes. The cohesin unloading scenario arises from the stochastic dynamics of the model. While the ring cannot bypass the replisome in principle, an LEF may transiently attempt such a move, leading to a higher-energy state. If accepted, this is typically followed by an unbinding event in subsequent MCs, resulting in loop removal. The ability of cohesin to bypass the replisome can be modulated using a replisome permeability coefficient C_{rep}. Adjusting this parameter partially relaxes the strict barrier assumption.

Epigenetic energy and compartmentalization

Following the Potts model framework [106], each node in the graph represents a chromatin segment and can exist in one of five discrete epigenetic states, $s_k \in \{0, \pm1, \pm2\}$. These states correspond to different combinations of epigenetic marks (e.g. histone modifications or DNA accessibility) and can dynamically change over time. Biologically, regions of chromatin bearing similar epigenetic signatures tend to spatially associate, forming domains known as subcompartments (such as A1, A2, B1, B2, and B3 in Hi-C data) [8, 107]. In our model, the alignment of multiple nodes in the same state increases their effective attraction, promoting the spatial segregation characteristic of compartmentalization. Therefore, we introduce five distinct states to capture a spectrum of possible chromatin identities and their contribution to 3D genome organization. We consider a model with ordered discrete states defined on a graph, with interaction energy proportional to the absolute difference between connected epigenetic states. In this sense, the model extends Potts-like models from statistical physics by allowing the interaction strength to depend on differences between ordered states rather than solely on their equality. The state energy term also includes an additional epigenetic field which depends on the time of replication:

$$\mathcal{H}_{\text{epi}} = \mathcal{C}_{p,1} \sum_{k=1}^{N_{\text{beads}}} \left(\frac{h_k + h_{t,k}}{2} \right) s_k + \mathcal{C}_{p,2} \sum_{k>l} J_{kl} \left| s_k - s_l \right|, \tag{8}$$

where h_k is an "*epigenetic field*" ($\mathcal{C}_{p,1} < 0$) which is proportional to the RT of the current region, normalized to take values from -1 to 1. It is assumed that the B compartment is linked to the negative values and the A compartment to positive values of the field. The term $h_{t,k}$ represents the reorganization of the epigenetic states due to the propagation of the replication fork [57, 58] and it is:

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

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 $h_{t,k}$ is zero, and the epigenetic field h_k is the dominant one. The term $h_{t,k}$ captures the influence of replication forks on the epigenetic landscape during a specific stochastic simulation. It is zero in unreplicated regions, positive when replication initiates (in early replicated regions) and negative when replication concludes (in late replicated regions). Biologically, $h_{t,k}$ represents the potential of replication forks to reorganize epigenetic states. Overall, this term reflects the progressive strengthening of compartmentalization from G1 to G2 and reproduces B-to-A compartment transitions during RS, corresponding to increased early origin firing [108–110]. However, the precise mechanisms by which replication forks propagate or modify epigenetic marks remain incompletely understood. To account for this uncertainty, RepliSage provides users with the option to disable this time-dependent effect by setting $h_{t,k} \equiv h_k$, allowing expert users to tailor the model according to their domain-specific assumptions or hypotheses. The final term ($\mathcal{C}_{p,2} > 0$) represents the interactions across the polymer graph. The interaction matrix is defined such that $J_{kl} = 1$ for monomer pairs that are connected either by a LEF, i.e. $(k,l) \equiv (m_i, n_i)$, or by backbone adjacency, i.e. $k = l \pm 1$ and $J_{kl} = 0$ otherwise. The difference $|s_k - s_l|$ penalizes interactions of monomers with varying epigenetic states, which leads to larger sizes of subcompartments.

Parameter balancing

When performing Monte Carlo simulations involving multiple energy components, it is important to understand how each term contributes to the model's behavior. In RepliSage, quantifying the influence of individual energy terms is particularly challenging due to the lack of experimental data to constrain their relative weights. To address this, we propose an intuitive tuning strategy based on achieving a rough balance between energy contributions at equilibrium. In the following, we provide a set of default parameters that users can adopt with confidence, ensuring that the energy terms remain ap-

proximately comparable in magnitude under typical simulation conditions (Fig. 2E).

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

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

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

$$\mathcal{C}_{\text{cross}} = \kappa \times 10^8. \tag{12}$$

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. The reparametrization is constructed so that the user specifies the lowercase parameters, which are then automatically normalized within the simulation into the corresponding uppercase variables.

Similarly, we can write for the replication and epigenetic energy parameters:

$$\mathcal{C}_{\text{rep}} = c_{\text{rep}} \times 10^8 \tag{13}$$

$$\mathcal{C}_{p,1} = -c_{p,1} \tag{14}$$

$$\mathcal{C}_{p,2} = c_{p,2}, \tag{15}$$

where the signs of the epigenetic energy parameters are chosen so that the energy is minimized when two interacting beads share the same node state. Specifically, $\mathcal{C}_{p,1}$ is negative to favor identical states, while $\mathcal{C}_{p,2}$ is positive because the interaction energy depends on the absolute difference of states, $|s_k - s_l|$, ensuring that larger differences contribute positively to the energy. The replication term is multiplied by a large value of 10^8 so as to penalize improbable configurations. With this reparametrization, setting $c_{\text{rep}} = c_{p,1} = c_{p,2} = 1$ yields comparable effective contributions from replication and epigenetic terms, ensuring that the different physical mechanisms compete on a balanced energetic scale (for a broader discussion about hyper-parameters of RepliSage, see the user manual in the Supplementary Information and Supplementary Table S2).

The algorithm

RepliSage employs the Metropolis Monte Carlo algorithm. The total energy change associated with a proposed update is decomposed into two components:

$$\Delta\mathcal{H} = \Delta\mathcal{H}_{\text{link}} + \Delta\mathcal{H}_{\text{nodes}}, \tag{16}$$

where $\Delta\mathcal{H}_{\text{link}}$ arises from network rewiring, corresponding to LEF movements and $\Delta\mathcal{H}_{\text{nodes}}$ arises from changes of epigenetic states. When a LEF move is proposed (i.e. modification of network links), two types of updates are possible: either a local random walk, shifting the LEF to neighboring positions $(m_i, n_i) \hookrightarrow (m_i \pm 1, n_i \pm 1)$, or a rebinding move, in which the LEF detaches and rebinds at a new random position with minimal loop length, i.e. $m_i = n_i$. If an epigenetic update is proposed, the state s_k at a given node may transition to any state different from its current one.

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

(1) The user defines the input `.bedpe` and RT data and from these data RepliSage determines the binding potentials and simulated replication fraction.
(2) Initialize random initial configurations of LEFs (m_i, n_i) and epigenetic states s_k.
(3) Initialize the matrix J_{kl} so as to be equal to 1 when k and l are connected by a LEF, or when $k = l \pm 1$ and 0 otherwise.
(4) With probability $p_1 = 0.5$, choose a node state change or rewiring move, otherwise.
(5) Compute the total energy difference $\Delta\mathcal{H}$ resulting from the proposed move, including contributions from the state interaction term. Accept the move if $\Delta\mathcal{H} < 0$, or with probability $e^{-\Delta\mathcal{H}/kT_{\text{mc}}}$ otherwise.
(6) Change the interaction matrix if a rewiring move was accepted.
(7) Repeat steps 4–6 for N_{steps} Monte Carlo steps (MCs). Each MC consists of N_{sweep} attempted Metropolis update proposals, as specified by the user. To ensure adequate mixing of the system, it is generally recommended that N_{sweep} is comparable to the system size N_{beads}, so that, on average, a substantial fraction of the system's states is proposed for change during each MC.

In the Metropolis simulation, configurations were sampled with a frequency f_{mc} in order to reduce correlations between successive states, and an initial burn-in period $\tau_{\text{burn-in}}$ was applied to allow the system to reach equilibrium. It is important to note that RepliSage is inherently an out-of-equilibrium simulation in the sense that when the replication forks start their action, the system is driven away from its pre-replication state and evolves toward a new configuration. Therefore, at least throughout the S phase, RepliSage describes an intrinsically non-equilibrium process.

Molecular dynamics

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

$$U = U_{\text{bk}} + U_{\text{ev}} + U_{\text{le}}(t) + U_{\text{rep}}(t) + U_{\text{pull}} + U_{\text{block}}(t). \tag{17}$$

Each term encodes a distinct physical contribution. The backbone potential U_{bk} enforces bonds between consecutive beads, maintaining approximately fixed distances along the chain. The excluded volume term U_{ev} ensures self-avoidance. The time-dependent potential $U_{\text{le}}(t)$ captures long-range, pair-specific loop interactions. The pulling term U_{pull} drives the spatial separation of the two chains during the G2/M phase. Finally, $U_{\text{block}}(t)$ is a non-bonded potential that models effective attraction between beads belonging to the same compartment. We use `OpenMM` [90, 91] to perform molecular dynamics simulations for 3D chromatin reconstruction; the parameters listed below define the corresponding polymer model and integration scheme.

The backbone potential, $U_{\text{bk}} = k_{\text{bk}} \sum_k^{N_{\text{beads}}} (r_k - d_{\text{bk}})^2$, includes strong covalent bonds between consecutive beads ($k_{\text{bk}} = 5 \times 10^5$ kJ/(mol · nm^2)) with an equilibrium distance $d_{\text{bk}} = 0.1$ nm, angular forces ($k_\phi = 200$ kJ/(mol · rad^2), $\theta = \pi$). Moreover, the excluded volume term is defined as:

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

where $\epsilon = 200$ kJ/mol, $r_{\text{ev}} = 0.1$ nm, and $\delta = 0.01$ nm. By setting $\epsilon = 0$ at specific positions along the polymer chain, we model the activity of topoisomerases. This mechanism is cru-

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

Parameter name	Symbol	Units	Value
Temperature	T	Kelvin	310
Backbone bond strength	k_{bk}	kJ/(mol·nm^2)	5×10^5
Backbone bond length	d_{bk}	nm	0.1
Backbone angle force constant	k_ϕ	kJ/(mol·rad^2)	200
Backbone angle equilibrium	θ	rad	π
Excluded volume strength	ϵ	kJ/(mol·nm^2)	200
Excluded volume radius	r_{ev}	nm	0.1
Excluded volume shift	δ	nm	0.01
Excluded volume cut-off distance	r_{cut}	nm	0.2
LEF bond strength	k_{le}	kJ/mol	5×10^4
LEF bond length	ℓ	nm	0.1
Replication bond strength	k_{rep}	kJ/(mol·nm^2)	5×10^4
Replication bond strength (replicated)	k_{rep}	kJ/(mol·nm^2)	10^3
Replication bond length	d_{rep}	nm	0.0
Replication bond length (replicated)	d_{rep}	nm	$r_{\text{rw}}/4$
Pulling force strength	f	kJ/(mol·nm)	10^3
Compartment interaction range	σ	nm	$r_{\text{rw}}/2$
Compartment equilibrium distance	ℓ_b	nm	0.2
Compartment energy (state $s = 2$)	E_2	kJ/mol	-0.2
Compartment energy (state $s = 1$)	E_1	kJ/mol	-0.4
Compartment energy (state $s = 0$)	E_0	kJ/mol	-0.6
Compartment energy (state $s = -1$)	E_{-1}	kJ/mol	-0.8
Compartment energy (state $s = -2$)	E_{-2}	kJ/mol	-1.0

cial for the disentanglement and subsequent full separation of sister chromatids following the completion of the S phase.

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

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

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 $k_{\text{le}} = 5 \times 10^4$ kJ/(mol · nm^2).

The replication potential, $U_{\text{rep}}(t)$, maintains strong harmonic bonds between paired monomers k and $k + N_{\text{beads}}$ across the two chains prior to replication; thus $\tau_i = |\vec{r}_i - \vec{r}_{i+N_{\text{beads}}}|$. As replication progresses, these bonds weaken, allowing chain separation:

$$U_{\text{rep}} = \sum_{k=1}^{N_{\text{beads}}} k_{\text{rep},k} \left(\tau_k - d_{\text{rep},k}\right)^2 , \tag{20}$$

where both the bond strength and equilibrium distance between the two polymer chains depend on the replication state of each monomer:

$$k_{\text{rep},k} = \begin{cases} 5 \times 10^4 \text{ kJ/(mol} \cdot \text{nm}^2), & \text{if not replicated,} \\ 10^3 \text{ kJ/(mol} \cdot \text{nm}^2), & \text{if replicated,} \end{cases} \tag{21}$$

and,

$$d_{\text{rep},k} = \begin{cases} 0, & \text{if not replicated,} \\ r_{\text{rw}}/4, & \text{if replicated.} \end{cases} \tag{22}$$

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 detachment process. The factor $r_{\text{rw}} = d_{\text{bk}}\sqrt{N_{\text{beads}}}$ 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 beads.

At the end of replication, upon entering the G2/M phase of the simulation, the replication-related potential U_{rep} is set to zero, and a new force is introduced to promote the spatial separation of sister chromatids. At this stage, an external force is introduced to drive the spatial separation of the two polymer chains. This force is implemented via the linear potential:

$$U_{\text{pull}} = -f\,x, \tag{23}$$

where x denotes the coordinate along the fixed pulling direction (1,0,0). The force magnitude is set to $f = \pm 10^3$ kJ/(mol · nm), with opposite signs applied to the two polymer chains, thereby inducing their separation along this direction. This pulling force acts in competition with excluded volume interactions, facilitating chromatid disentanglement while preserving steric constraints.

The block-copolymer potential U_{block} is modeled as a Gaussian-shaped non-bonded interaction [111–113], promoting attraction between monomers belonging to the same epigenetic state. The potential acts *over all pairs of monomers* and is written as:

$$U_{\text{block}} = \sum_{k<l} E \exp\left(-\frac{(r_{kl} - \ell_b)^2}{2\sigma^2}\right), \tag{24}$$

where the interaction strength E is defined as:

$$E = \sum_{s \in \{0, \pm 1, \pm 2\}} E_s\, \delta(s'_k - s)\, \delta(s'_l - s), \tag{25}$$

with E_s taking values specified in Table 1 when both interacting monomers belong to the same compartment, i.e. $s'_k = s'_l$. Here, r_{kl} denotes the distance between monomers k and l, ℓ_b is the preferred interaction distance, and σ sets the interaction range. The interaction energy E_s is stronger for the B compartment, which correlates with late RT. The range of excluded volume interactions is defined as $\sigma = r_{\text{rw}}/2$, meaning that the interaction range is half the average distance of a random walk. 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 [9, 107].

The 3D dynamics of the polymer chains were simulated using Langevin dynamics, which incorporates conservative forces, viscous damping, and thermal noise. Each bead is assigned a mass of $m = 16{,}427.889$ amu, chosen to approximate a small segment of chromatin represented in the coarse-grained model (roughly 1/10 of a nucleosome). All beads are identical, so the precise mass does not affect the relative structural organization or qualitative dynamics of the system. The equation of motion reads:

$$m\frac{d^2\vec{r}}{dt^2} = -\nabla U - \gamma\frac{d\vec{r}}{dt} + \sqrt{2\gamma k_B T}\,\vec{\eta}(t), \tag{26}$$

where U is the total potential energy of the system, γ is the friction coefficient, k_B is the Boltzmann constant, and $T = 310$ K is the simulation temperature (this is different parameter from T_{mc}). The stochastic term $\vec{\eta}(t)$ represents Gaussian white noise with zero mean and unit variance, satisfying $\langle \eta_i(t)\eta_j(t')\rangle = \delta_{ij}\delta(t-t')$. The friction coefficient was set to $\gamma = 0.1$ fs^{-1}, and the integration time step was $\Delta t = 100$ fs. The first term on the right hand side describes conservative forces derived from the total potential U, which includes bonded and non-bonded interactions defined previously. The second term accounts for viscous damping due to the implicit solvent, while the third term enforces the fluctuation dissipation theorem, ensuring correct thermal sampling at 310 K. The Langevin integrator is used by default unless otherwise stated. Energy minimization was performed prior to dynamics with a tolerance of 1.0 kJ/mol. This Langevin framework allows the polymer to explore thermally accessible conformations while remaining consistent with the underlying physical potentials encoded in the OpenMM force field [90, 91].

Parallelization

Monte Carlo simulations, particularly those based on the Metropolis algorithm, are computationally demanding due to the large number of stochastic updates required to explore the energy landscape thoroughly. Nevertheless, their key advantage lies in generating structural ensembles rather than relying on single averaged configurations, allowing more accurate representation of biological variability. To improve computational efficiency, RepliSage incorporates just-in-time (JIT) compilation using the `numba` Python library [114], which translates critical parts of the Monte Carlo code into low-level machine instructions. This optimization results in substantial speedups, with performance gains of up to 100-fold as reported in the library's documentation. The same approach has also been applied to develop an optimized version of LoopSage. For the molecular dynamics component, RepliSage employs OpenMM [90, 91], which supports parallel execution on multi-core CPUs and enables GPU acceleration via OpenCL or CUDA backend. The complete RepliSage pipeline has been validated across a range of computational environments, including personal laptops, desktop workstations, and high-performance computing (HPC) clusters.

Results

The primary aim of the RepliSage analysis is to investigate how 3D chromatin architecture changes across the cell cycle and how RS modulates these structures. In particular, we use the model to test mechanistic assumptions regarding the interaction between replication forks and loop extrusion, as well as the potential role of forks in organizing epigenetic states. Although such assumptions could, in principle, be explored using purely synthetic inputs, we incorporate experimental data to ensure biological relevance. The RepliSage framework integrates three interconnected modules (replication simulation, stochastic loop extrusion, and 3D physical modeling) linking 1D DNA stochastic processes to 3D chromatin structure. To assess whether these mechanisms capture meaningful features of genome organization, we compare RepliSage simulations with orthogonal experimental datasets and with NucDynamics simulations [93], providing a basis for model validation.

RepliSage integrates two main input data types: (i) RT profiles and (ii) chromatin loop data from 3C-type assays, which are used to define CTCF-mediated interactions. By default, RepliSage internally loads single-cell RT data [115] generated for the GM12878 human LCL. However, alternative RT datasets can be easily substituted as required. RepliSage also supports bulk Repli-seq data provided in BigWig format. In this study, we used CTCF ChIA-PET data from ENCODE as 3C-based input for two cell lines: GM12878 and K562 human erythroleukemia cell line to leverage available validation data. Single-cell Hi-C data is available for the K562 [92], whereas in-house cohesin HiChIP data were generated for GM12878. For modeling purposes, we used GM12878 single-cell RT data [115] as a baseline reference for both cell lines, as single-cell RT data for K562 cells is not available. The strong correlation (over 80%) between pseudo-balk single-cell RT in GM12878 and the K562 bulk Repli-seq data from ENCODE indicates substantial conservation of global RT structure across both cell types (Supplementary Fig. S2A and B). To assess potential bias arising from the use of different cell lines for single-cell RT and ChIA-PET input data, we performed additional simulations using matched datasets, i.e. GM12878 Repli-seq with GM12878 ChIA-PET and K562 Repli-seq with K562 ChIA-PET.

Modelling of cell cycle and replication stress

To model the cell cycle, we divided the simulation into three distinct phases (G1, S, and G2/M) each requiring specific energy modifications and implementation strategies. In the G1 phase, chromatin dynamics are driven by loop extrusion, where BEs act as static barriers to LEFs, and epigenetic states evolve according to the discrete-state interaction. The number of LEFs in this phase was set to $N_{\mathrm{lef}} = 1000$, with a baseline folding coefficient $f_1 = 1$. In the S phase, replication forks, imported from the *Replikator* module, are introduced and act as dynamic, transient obstacles to LEF progression. Once replication is complete, the two daughter DNA chains remain closely associated. The G2/M phase is modeled by (i) introducing a second population of LEFs ($N_{\mathrm{lef},2} = 1000$) with a higher folding coefficient $f_2 = 2$, representing condensins, and (ii) enabling local relaxation of excluded volume interactions through random Monte Carlo moves, mimicking topoisomerase activity with probability $p_{\mathrm{top}} = 0.01$. This reflects experimental findings that condensins operate 2–5 times faster than cohesins and can bypass them without displacement [116], and that topoisomerases act throughout the cell cycle. Both LEF populations operate independently and continuously extrude loops. Chromosome-wide simulations were performed using a coarse-grained polymer model of two chains,

each consisting of $N_{beads} = 10,000$ beads. The Monte Carlo order parameter was fixed at $T_{mc} = 1.8$ to balance structural order and fluctuation.The total simulation length was 2×10^5 MCs. The three cell-cycle phases were assigned the following temporal windows: G1: 0–5×10^4 MCs, S phase: 5×10^4–10^5 MCs, and G2/M: 10^5–2×10^5 MCs. For RS simulations, we used twice as many steps in S phase, reflecting the slower progression of replication under stress conditions. This choice was applied consistently across all simulations presented in this work. It is important to note that Monte Carlo time represents a simulation timescale (pseudotime) and does not correspond directly to physical time. One Monte Carlo step corresponds to $N_{sweep} = 10,000$ attempted Metropolis updates, ensuring sufficient LEF and epigenetic updates for proper mixing, particularly given the strong autocorrelations inherent to cell-cycle dynamics. Configurations were sampled every $f_{mc} = 200$ steps after a burn-in period of $\tau_{burn\text{-}in} = 10^3$ MCs. Execution of both the stochastic Monte Carlo simulation and associated molecular dynamics via OpenMM required slightly less than one day on a standard computational infrastructure. To model RT under RS conditions, we increased the initiation rate while reducing replication fork speed (further details on the specific parameter choices are provided in the Supplementary Information). Under normal conditions, fewer than 10% of origins fire, whereas under stress this proportion increases to 20%–40% (for more details, see Supplementary Information).

We simulated the 3D organization of human chromosome 14 throughout the cell cycle (Fig. 3A). This analysis was performed for both GM12878 and K562 cell lines (Fig. 3, and Supplementary Figs S3 and S4), focusing on a subtelomeric-excluded region of chromosome 14 spanning positions 23 535 000 to 99 174 700 with the cell-cycle specific parameters discussed earlier. The chromosome was simulated with two polymer replicates, each consisting of 10 000 beads. To analyze chromatin folding dynamics, we monitored the evolution of the average loop length over Monte Carlo pseudo-time (Fig. 3B). A slight but statistically significant decrease in loop length is observed during the S phase. A comparable decrease in the average loop length was also observed when simulations were performed using Repli-seq data (Supplementary Fig. S2C and D). In the G2/M phase, condensin activity restricts the formation of long, mobile loops, thereby rapidly guiding the system toward structural equilibrium.

The increased replication activity under stress leads to a statistical significant reduction in average loop size, highlighting that RS alters loop organization (Fig. 3B and C). However, this reduction remains moderate because the "*moving barrier*" effect exerted by replication forks on LEFs is limited in both time and space. Replication forks act only within regions that are actively being replicated at a given time, and their influence does not accumulate uniformly across the genome as replication progresses and forks merge. Furthermore, under RS conditions, the elevated initiation rate leads to faster synchronization of links connecting loci with identical epigenetic states, as evidenced by the earlier rise to higher signal intensities in the plot (Fig. 3D). This phenomenon arises because a greater number of replication origins fire earlier. According to the model, these replication origins act as epigenetic organisers that locally promote the alignment of chromatin regions from B into the A compartment. RepliSage simulations driven by Repli-seq data show the same qualitative behavior (Supplementary Fig. S2E). Despite these differences in synchronization dynamics, the overall compartmental organization remains largely consistent between normal and stress conditions. These findings are further supported by HiChIP data, which reveal a highly consistent compartmentalization pattern between normal and RS conditions, exhibiting Pearson correlation coefficients exceeding 94% (Supplementary Fig. S5). Nonetheless, when we enable the model assumption that replication forks are able to reorganize epigenetic marks (Eq. 9), we observe an +7.6% increase in B→A switching of the node epigenetic states (Fig. 3E). These epigenetic state changes are correlated with the compartment state.

In Fig. 4, gyration radius and convex hull volume of the single polymer chain are calculated for each structure of the ensemble. These metrics reveal that during the G1 phase, the chromosome is in a compact state. In the S phase, decompaction occurs due to the separation of the two polymer chains and the initiation of replication origins. Subsequently, in the G2/M phase, the chromatin re-compacts, driven by the activity of condensins and enhanced by block copolymer interactions.

Validation of the model

Validating 3D chromatin models remains challenging due to the shortage of ground truth structural data defining the exact spatial positions of atoms in the 3D structure. Consequently, validation commonly relies on comparisons between simulated distance maps and experimental contact matrices, as well as on the detection of characteristic architectural features such as loops, stripes, and TADs [6, 117]. Replication-associated dynamics have been linked to fountain-like interaction patterns and to TAD-border enrichment in G2/M, reflecting sister chromatid interactions [118, 119]. However, direct map-to-map correlations can be misleading due to diagonal dominance, normalization procedures, and differences in resolution and smoothness, all of which introduce systematic biases [120–122]. Moreover, robust metrics for detecting specific structural features remain limited. These challenges motivate the combined quantitative and qualitative validation strategies employed in this study.

For validation, we compared RepliSage outputs with orthogonal Hi-C datasets in GM12878 and K562 (ENCODE), and qualitatively assessed key biophysical features including loop sizes and chromatin compaction using single-cell Hi-C [92]. Importantly, to directly investigate changes in 3D chromatin organization upon RS, we performed cohesin HiChIP [94] under control conditions and following RS induction.

RepliSage integrates biophysical modeling across multiple chromatin scales. At the TAD level, it incorporates a modified version of LoopSage [86], which we previously validated through comparisons of loop-specific features between simulations and experimental datasets. Although RepliSage focuses on chromosome-scale modeling rather than loop extrusion itself, accurate representation of loop extrusion is essential because it directly interacts with DNA replication and the propagation of epigenetic states. At the compartment scale, RepliSage employs an epigenetic field derived from RT data. Since RT strongly correlates with chromatin compartmentalization [40], the resulting epigenetic state patterns closely recapitulate Hi-C derived A/B compartments. Consistent with this, the ensemble-averaged epigenetic states reach over 80% Pearson correlation with K562 Hi-C compartment profiles

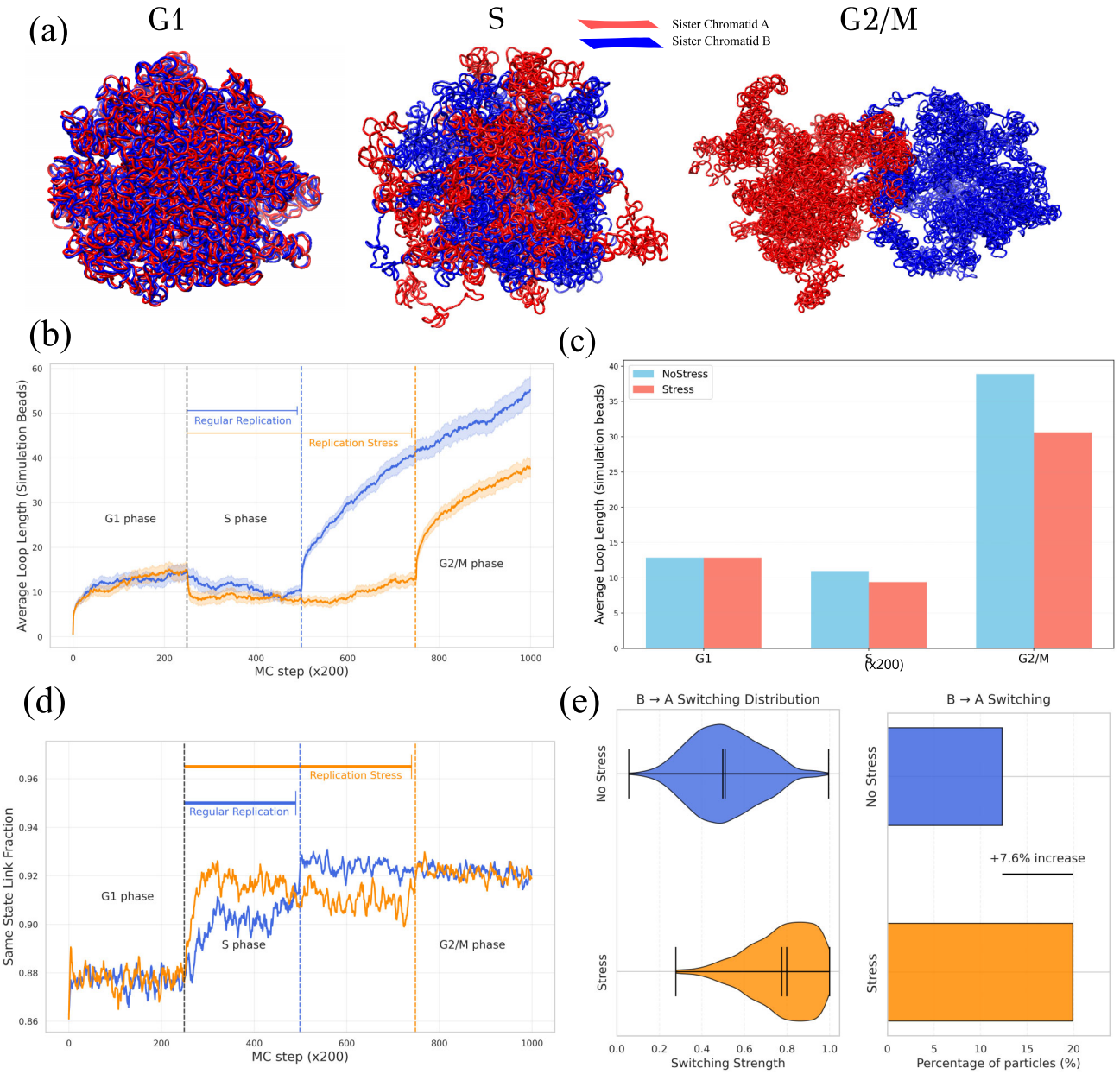


Figure 3. Modeling of the representative region [23535000, 99174700] bp of chromosome 14 in K562 cells using RepliSage. (**a**) Representative 3D structures of chromosome 14 spanning positions 23 535 000 to 99 174 700 bp during the G1, S, and G2/M phases of the cell cycle. Each polymer bead represents approximately 7.56 *kb* of genomic DNA. (**b**) Evolution of the average loop length as a function of MCs, including comparison with RS conditions. The shaded area represents the 95% confidence interval of the average loop length, highlighting the clear separation between control and stress conditions. For regular replication, the mean relative decrease compared to G1 is approximately 25%, with a paired *t*-test confirming significance ($t = 3.82$, $p \approx 1.37 \times 10^{-4}$). Under RS, the decrease remains highly significant ($t = 7.18$, $p \approx 9.71 \times 10^{-13}$). (**c**) Average loop size across cell cycle phases in normal and RS conditions. (**d**) Fraction of LEF-mediated links connecting nodes of identical state (i.e. links between nodes with the same sign; for example, nodes with states −2 and −1 are considered equivalent). (**e**) Violin plot showing the distribution across simulation beads of the change in average node state during S phase, calculated as the difference between each bead's average state in G1 and in G2/M (left panel). The average percentage of B→A switching relative to normal replication conditions (right panel).

(Fig. 5A and B). To further validate large-scale organization, we compared the first eigenvectors of inverse distance matrices from simulated ensembles (Fig. 5C) with those from Hi-C data (Fig. 5D), observing correlations above 60%. In contrast, random walk ensembles, lacking compartmental structure, exhibit near-zero correlation, further supporting the biological relevance of RepliSage output.

As RepliSage is primarily designed to model DNA replication dynamics, we performed qualitative validation using single-cell Hi-C datasets for K562 cell line [92] which provides insight into the distribution of loop sizes across distinct cell cycle phases. We constructed phase-specific interaction sets by merging overlapping anchors and retaining statistically significant interactions supported by at least three contacts. Interactions exceeding 2000 kb were excluded, as such long-range contacts are unlikely to arise from canonical loop extrusion mechanisms. This analysis revealed a consistent trend across cell cycle: during S phase, average loop

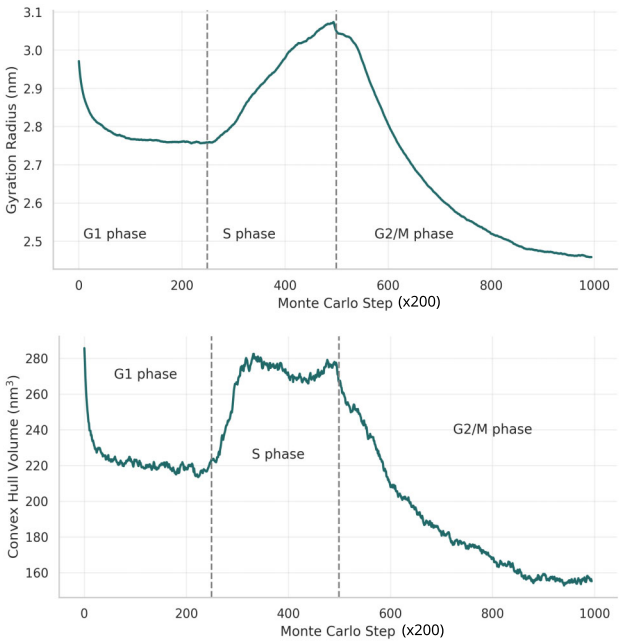


Figure 4. Structural metrics of the obtained 3D chromatin configurations derived using the K562 ChIA-PET dataset. During the G1 phase, the chromatin exhibits compaction. In the S phase, the activity of replication forks leads to a transient de-compaction of the structure. In the G2/M phase, increased compaction is observed again, driven by block–copolymer interactions and condensin-mediated binding.

lengths are shorter than in G1, while in G2/M, loops tend to be longer, likely reflecting condensin-mediated chromosome compaction (Fig. 5E and F) in agreement with Fig. 3B and C.

Finally, we experimentally validated whether RS influences chromatin loop architecture, as observed in the RepliSage simulations (Fig. 3B and C). To this end, we performed a cohesin HiChIP experiment to capture cohesin-mediated chromatin contacts at high resolution. We analyzed loop length distributions in the GM12878 LCL under mild RS conditions induced by two different chemical compounds that interfere with DNA replication: (i) aphidicolin, a DNA polymerase inhibitor, and (ii) hydroxyurea, which inhibits ribonucleotide reductase and leads to the depletion of the deoxyribonucleotide pools. As control conditions, we used DMSO or water, respectively. A statistically significant reduction in chromatin loop sizes was observed under both stress conditions, indicating that RS leads to a diminishing of the length of chromatin contacts (Fig. 5G). These findings are consistent with the simulation results and support the mechanistic predictions of the RepliSage framework (Fig. 3B and C).

Low resolution models of chromosomes

To gain deeper insights into chromatin structural dynamics throughout the cell cycle, we generated low-resolution (100 kb) 3D models for the K562 (Fig. 6) and GM12878 (Supplementary Fig. S6) cell lines. Violin plots (Fig. 6A) confirm that S phase exhibits the shortest loop sizes. The results were consistent with previous high-resolution analyses (Fig. 3, and Supplementary Figs S2 and S4), showing a robust decrease in average loop size during S phase under RS compared to normal conditions (Fig. 6B), likely due to increased replication fork activity acting as "moving barriers" to LEFs.

Furthermore, to quantitatively assess the structural properties of chromatin across the cell cycle, we computed a range of geometric descriptors, as detailed in the Supplementary Information (see Definition of structural metrics section). These descriptors were used to perform principal component analysis (PCA) as shown in Fig. 6C. The PCA results revealed clear clustering of chromatin conformations according to cell cycle phase, with each phase forming distinct groups in the reduced-dimensional space. Furthermore, analysis of the radius of gyration R_g across chromosomes (Fig. 6D) demonstrated a consistent trend of chromatin decompaction during the S phase, followed by increased compaction during the G2/M phase relative to G1. These findings underscore the phase-specific reorganization of chromatin structure during cell cycle progression.

Comparison with models from single-cell Hi-C data

To assess the consistency of RepliSage predictions, we compared them with single-cell chromatin structure models reconstructed from single-cell Hi-C data using the NucDynamics framework [93], applied to the K562 single-cell Hi-C dataset of Chai *et al.* (2025) [92]. In that study, cells were classified into G1, S, and G2/M phases based on the expression of phase-specific genes. PCA was performed, and the contact data were projected onto the first two principal components to group cells into pseudo-bulk metacells, representing distinct stages along the cell cycle pseudo-time. Following their methodology, we employed the NucDynamics software [93] to reconstruct 3D chromatin structures for each meta-cell. This allowed us to investigate the overall patterns of chromatin structural changes and compaction across the cell cycle, which we then compared to analogous results generated using RepliSage for the same K562 cell line.

To quantitatively assess chromatin compaction throughout the cell cycle, we calculated the radius of gyration (R_g) and the convex hull volume for each chromosomal structure. The results are presented in Supplementary Figs S7–10. In the NucDynamics-based analysis, both R_g and convex hull volume gradually increased from G1-to-S phase, indicating a relaxation of chromosome, and then decreased upon entry into the G2/M phase, suggesting a re-compaction of chromatin. A comparable pattern was observed with the RepliSage results (see Fig. 6D), thereby supporting the validity of our approach. Notably, although RepliSage does not incorporate single-cell contact data in its modeling, the chromatin dynamics it infers align closely with known chromatin compaction trends from other studies [35] as well as with our own 3D structural reconstructions. This consistency highlights the robustness of RepliSage in capturing broad chromatin behavior across the cell cycle, despite relying on an orthogonal data modality.

Parameter study

To explore how model parameters influence the relationship between DNA replication and chromatin structure, we performed a targeted parameter study on a representative region of human chromosome 14, spanning genomic coordinates [70835000, 98674700] bp. This specific segment was selected to balance biological relevance and computational efficiency, as simulating an entire chromosome at high resolution would be computationally prohibitive.

Simulations were carried out using two polymer chains, each comprising $N_{\text{beads}} = 2000$ beads. The total simulation

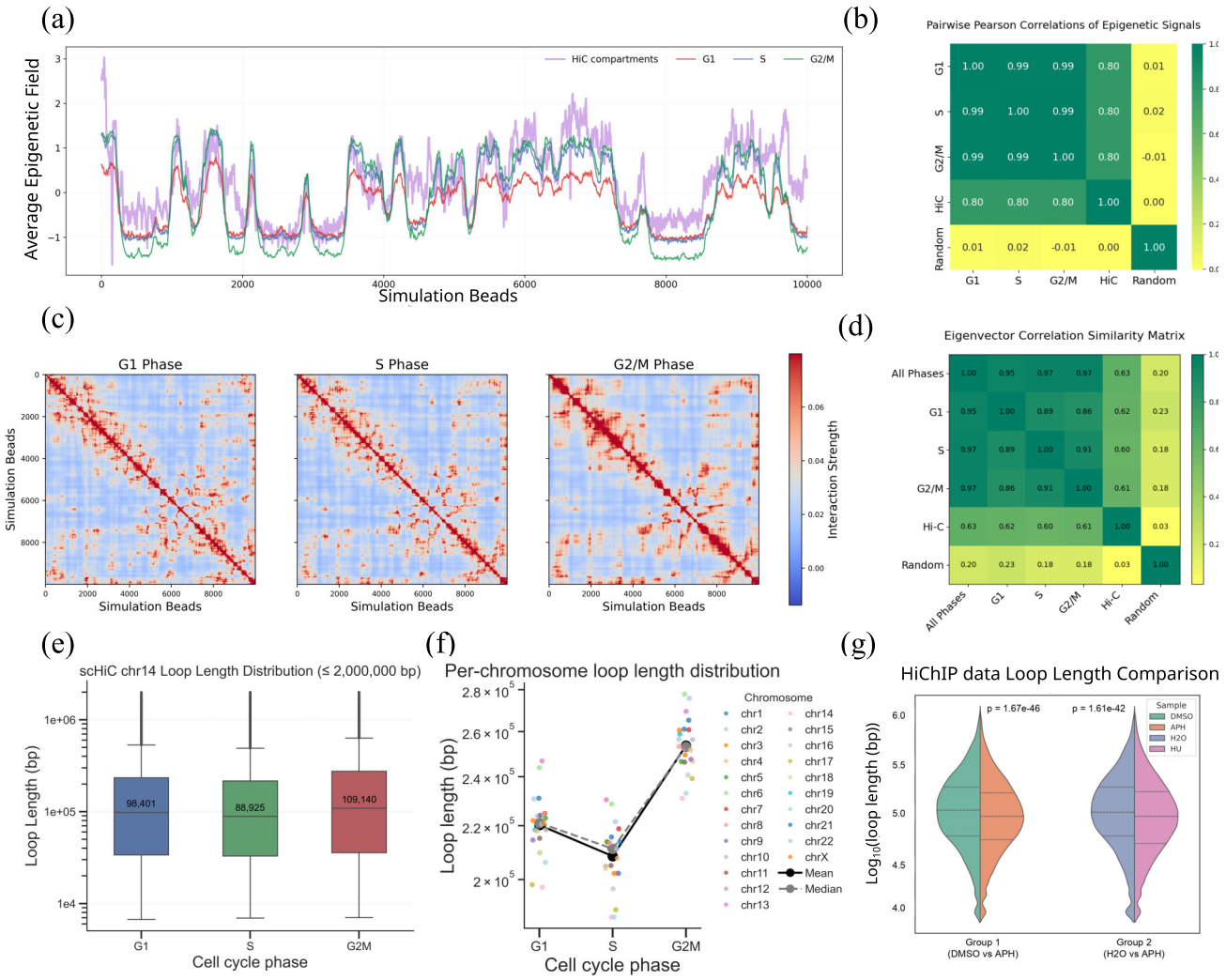


Figure 5. Validation of the RepliSage model. (**a**) Ensemble-averaged epigenetic field compared to K562 Hi-C (ENCODE) compartmentalization profiles across different cell cycle phases. (**b**) Pearson correlation between simulated and experimental compartment signals. (**c**) Averaged inverse all-versus-all distance heatmaps from RepliSage across cell cycle phases. (**d**) Pearson correlation between the first eigenvector of simulated distance maps and Hi-C compartment eigenvectors. (**e**) Distribution of interaction distances derived from single-cell K562 Hi-C data [92]. (**f**) Per-chromosome distribution of interaction lengths across cell-cycle phases from single-cell K562 Hi-C data. Each point represents a chromosome-specific mean interaction length, with global phase-wise mean and median trends shown as solid and dashed lines, respectively. (**g**) Comparison of loop length distribution derived from GM12878 SMC1 HiChIP experiments in two groups: DMSO and APH (Group 1), and H_2O and HU (Group 2). Violin plots show the distribution of log-transformed loop lengths ($\log_{10}$ scale) with quartile lines overlaid. In both groups the difference upon treatment is statistically significant ($p < 0.001$, Mann–Whitney *U* test). APH, aphidicolin; HU, hydroxyurea.

time was set to $N_{\text{steps}} = 2 \times 10^5$ MCs. The system was first equilibrated during the G1 phase for $\tau_{\text{rep}} = 5 \times 10^4$ MCs. Replication was then simulated over $T_r = 5 \times 10^4$ MCs (S phase), followed by additional MCs to allow the system to reach equilibrium post-replication (G2/M phase). In this framework, one Monte Carlo step corresponds to $N_{\text{sweep}} = 2000$ individual Monte Carlo moves. Each move consists of either LEF dynamics or an epigenetic state update, ensuring efficient configurational sampling and proper stochastic reshuffling of the system. We mainly focused on the behavior of the system under the change of the order parameter taking the following values $T_{\text{mc}} \in [1, 1.4, 1.8, 2.2, 2.6, 3.0]$. This controlled setup enabled systematic investigation of parameter-dependent chromatin reorganization across the cell cycle. For each temperature, RepliSage was independently executed 10 times, and the statistical averages from these runs were presented in Fig. 7.

RepliSage exhibits strong sensitivity to its parameter settings. Therefore, gaining at least a qualitative understanding of its behavior under perturbations of key parameters is essential. Figure 7A shows how the average loop length varies as a function of the order parameter T_{mc}. Lower values of T_{mc} promote longer and more stable loops, whereas higher values favor shorter loops. This reflects the role of T_{mc} as an effective biological temperature, where increased kinetic noise enhances LEF unbinding and limits loop growth. Reducing T_{mc} stabilizes LEF binding, favoring energy-minimizing configurations and enabling clearer distinctions in loop size across cell-cycle phases.

By varying replication fork speed and initiation rate (Supplementary Fig. S11A) for a fixed $\hat{T}_{\text{mc}} = 1.4$, we find that slow forks combined with higher initiation rates produce smaller average loop sizes, whereas fast fork progression leads to larger loops. This behavior is consistent with earlier

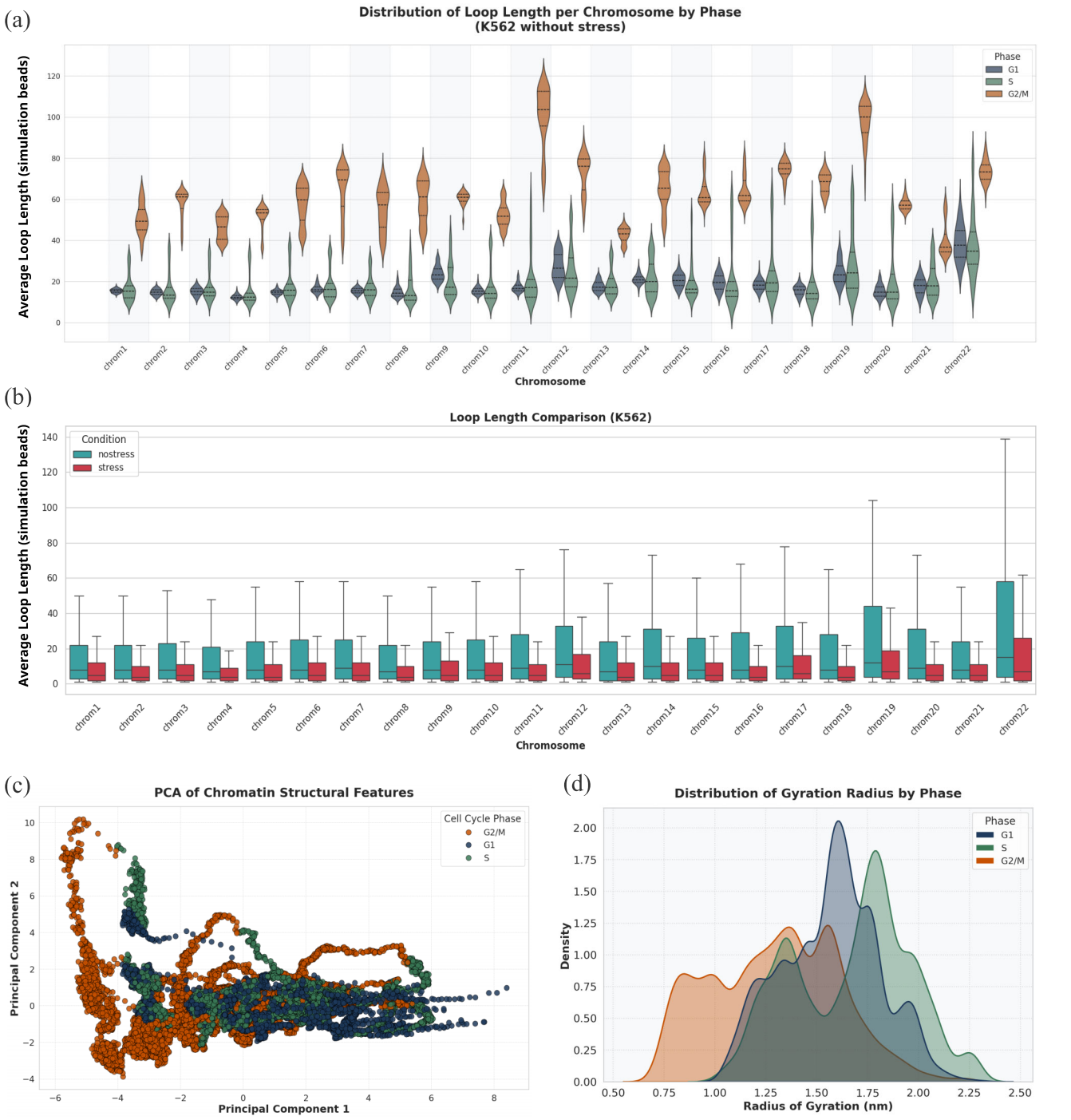


Figure 6. Chromosome-scale structural changes across cell cycle phases and upon RS as predicted by RepliSage using CTCF ChIA-PET input data in K562 cells. Each polymer bead represents 100 kb of genomic DNA, providing a coarse-grained view that emphasizes large-scale organization. As a result, the apparent units and structures may differ slightly from those in higher-resolution simulations, reflecting the model's focus on long-range chromosomal architecture. (**a**) Distribution of loop lengths across cell cycle phases. (**b**) Comparison of loop length distributions under normal and RS conditions. (**c**) PCA of structural descriptors (as defined in the Definition of structural metrics section) for each polymer conformation. Each structure is encoded as a high-dimensional vector capturing geometric attributes. The unsupervised 2D projection shows weak clustering by cell cycle phase, indicating subtle phase-specific structural signatures. Each point represents a single 3D structure. (**d**) Distribution of the radius of gyration R_g across chromosomes and cell cycle phases represented by the KDE estimation of the histograms. It shows global chromatin decompaction during S phase (shift toward higher R_g) and increased compaction in G2/M phase (shift toward lower R_g).

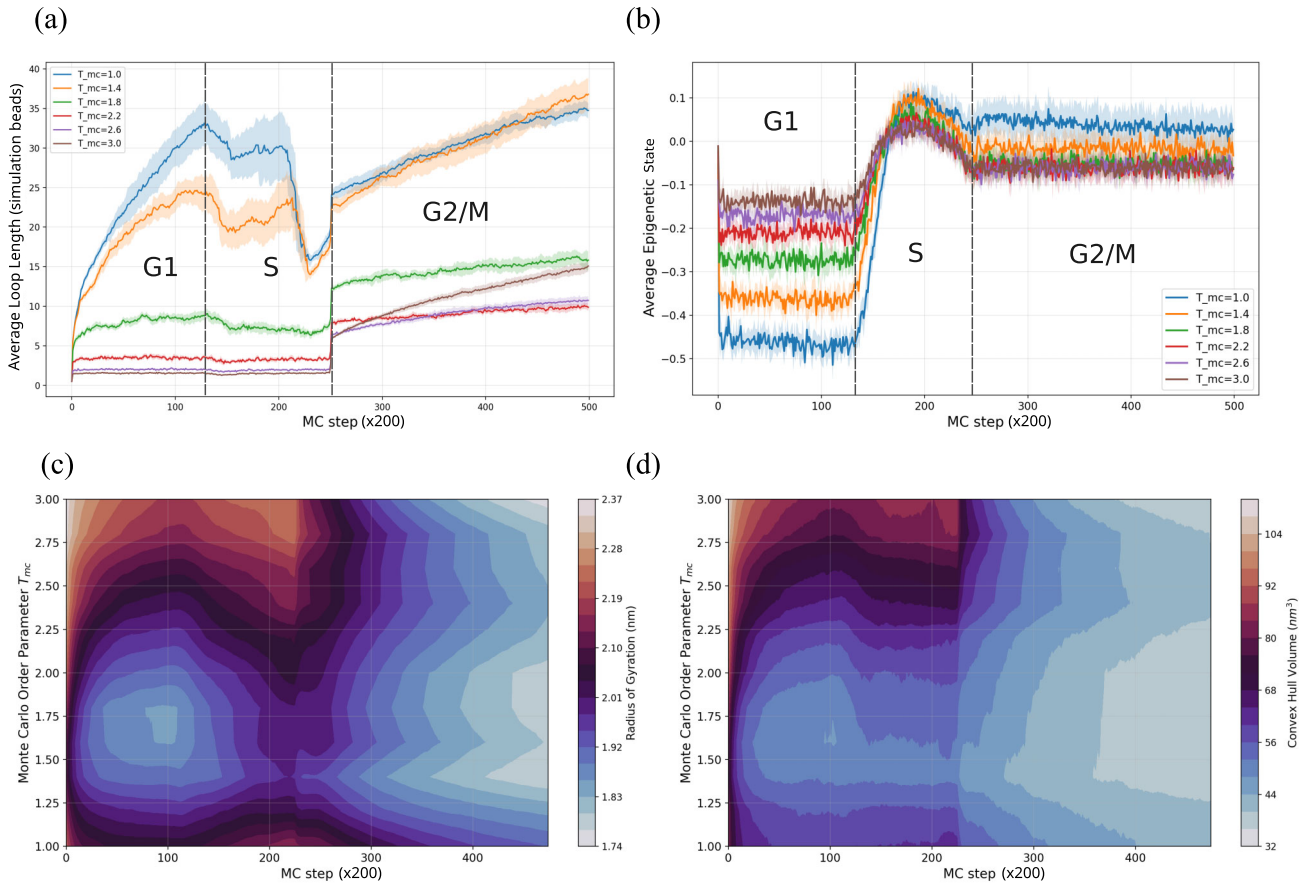


Figure 7. Parameter study of RepliSage. Results are based on 10 independent simulations using CTCF ChIA-PET data from the GM12878 cell line, focused on the illustrative region [70 835 000, 98 674 700] of chromosome 14. The polymer model uses 2000 beads, with each bead representing approximately 14 kb of genomic DNA. For each value of T_{mc}, 10 independent runs of the RepliSage algorithm were performed and subsequently averaged. (**a**) Average loop length as a function of the order parameter T_{mc}. Elevated values of T_{mc} correspond to shorter loops, whereas lower values reproduce previously reported trends: moderate compaction during S phase followed by enhanced compaction in G2/M phase. (**b**) Average epigenetic state, quantified as the mean value $\langle s \rangle$ across all nodes in the chromatin graph, plotted as a function of T_{mc}. A discrete jump in $\langle s \rangle$ reflects replication-induced epigenetic reconfiguration driven by replication fork progression. (**c**) Contour plot depicting the evolution of the radius of gyration as a function of both T_{mc} and Monte Carlo time. A range of critical T_{mc} values is observed at which chromatin compaction is maximized. (**d**) Contour plot of the convex hull volume as a function of T_{mc} and Monte Carlo time.

observations from whole-chromosome simulations under both normal and RS conditions (Fig. 3, and Supplementary Figs S2 and S4). Another factor influencing replication dynamics is the blocking effect of replication forks on LEFs. While previous simulations assumed complete LEF blockage at replication forks, experimental evidence indicates that cohesin can bypass a fork [41–44]. We tested this scenario by varying the fork permeability score (c_{rep}) (see Eq. 6 and Eq. 15). Accordingly, in Supplementary Fig. S11B, we quantify the mean number of LEF–replisome violations over the course of S phase as a function of the permeability score c_{rep}. For small values of c_{rep} (e.g. $c_{\text{rep}} \lesssim 10^{-6}$), rare violations are observed, indicating that LEFs can occasionally traverse replication forks even under strong barrier conditions. This effect becomes more pronounced at lower values of the order parameter T_{mc}, where reduced stochasticity favors configurations more consistent with the assumed energy, but still allows occasional crossing events. Taken together, this behavior indicates that c_{rep} acts as a continuous regulator of fork permeability, determining whether LEF crossing is suppressed or allowed.

An additional point of interest relates to the recent work of Liu *et al.* [118], which demonstrated that replication fork progression can generate striking fountain-like patterns in specially designed Hi-C experiments. Although a direct comparison with our simulations is technically difficult, given the different nature of simulated and experimental contact maps and the lack of any robust quantitative metric for identifying such features, we nevertheless observe qualitatively similar structures near the G1–S transition (Supplementary Fig. S11A). In particular, we detect a pronounced fountain orthogonal to the main diagonal (Supplementary Fig. S11B) at the end of S phase, located at the center of the heatmap precisely at the genomic position where replication terminates in the region of interest (Supplementary Fig. S12A). This raises the intriguing possibility that fountain-like patterns may reflect transient chain separation induced by fork progression, though further work would be required to test this connection rigorously. A second relevant discussion concerns the observations of [119], who reported enhanced TAD borders in G2/M, attributed to interactions between sister chromatids. Our simulations also exhibit strengthened TAD boundaries during G2/M (Supplementary Fig. S11C); however, in our framework this arises primarily from condensin-driven compaction rather than explicit sister-chromatid interactions.

The epigenetic node state model, includes two coefficients: one controlling chromatin–chromatin interactions and another representing the epigenetic field (Eq. 8). As expected, at high Monte Carlo temperatures the epigenetic field fluctuates randomly, while at low temperatures transitions are suppressed, producing ordered configurations (Fig. 7B). Replication forks shift the epigenetic state energy minimum, enhancing domain organization. In simulations, although the global epigenetic landscape is largely preserved (Supplementary Fig. S12B), local rearrangements appear when the interaction term dominates, consistent with statistical physics results where long-range interactions can drive phase transitions [106, 123–126]. In our chromatin model, loop extrusion operates over length scales that are short compared to chromosomal and compartmental dimensions, keeping the system below the transition threshold and maintaining compartment stability across the cell cycle and stress conditions [127, 128]. According to our model, two conditions would be required for loop extrusion alone to drive compartmental organization: (i) it must play a central role in establishing compartments, and (ii) it must generate long-range loops spanning compartment-scale distances. Current evidence indicates that neither condition holds [129, 130], and the local, short-ranged nature of loop extrusion [5, 52, 131] explains the remarkable stability of compartment patterns even under stress [108–110, 127, 128].

An important observation concerns how the 3D chromatin structure is influenced by variations in the Monte Carlo parameters. This connection is far from obvious, as the internal graph states of the model, defined in 1D, just ultimately give rise to a spatial structure in 3D, where the system has significantly more degrees of freedom. Figure 7C and D present contour plots of the radius of gyration and convex hull volume, respectively, as functions of both T_{mc} and Monte Carlo time. As expected from prior results, the chromatin structure is more compact in the G1 and G2/M phases. However, a particularly interesting result is the emergence of a critical temperature region around $T^{*}_{\mathrm{mc}} \approx 1.7 \pm 0.3$, where compaction reaches a maximum. This optimal folding does not occur at the lowest temperatures, but rather within an intermediate range, suggesting that a certain level of stochasticity is necessary to achieve biologically realistic 3D organization. Notably, this behavior is consistent across all cell cycle phases, indicating a robust temperature-dependent folding regime within the model.

Discussion

Modeling of DNA replication has progressed from early 1D stochastic and kinetic frameworks, in which origin activation and fork progression were simulated explicitly [71–81], to more recent 3D approaches that incorporate spatial aspects of replication, including origin clustering, differential RT in open versus closed chromatin, and explicit modeling of fork propagation within the 3D nuclear environment [82–85]. Parallel developments in loop-extrusion modeling [86, 131–136] have integrated stochastic LEF dynamics with molecular dynamics, connecting chromatin organization to cell-cycle processes. Yet, most existing approaches focus on isolated mechanistic hypotheses, neglecting the interplay between replication forks, LEFs, and epigenetic or compartmental organization (Supplementary Information and Supplementary Table S1).

RepliSage provides the first proof-of-concept framework capable of simulating the entire cell cycle within a single model, integrating the three key biophysical processes of chromatin: loop extrusion, compartmentalization (epigenetic mark organization), and DNA replication, while modeling their combined dynamics. This model not only enables testing mechanistic assumptions, as other replication models do, but also allows users to incorporate their own 3C-type interaction and RT data and to generate ensembles for each phase of the cell cycle. In addition, RepliSage can simulate RS by allowing adjustments to origin initiation rates and replication fork speed parameters. Finally, G2/M phase dynamics are captured by modeling the progressive separation of replicated chromatids driven by topoisomerase activity and condensin-mediated loop extrusion. The model is fully documented and available through an easy-to-install GitHub and PyPI repository, supporting reproducible and flexible research applications.

Biological support of RepliSage is divided into three distinct parts: (i) incorporating biologically realistic rules derived from experimental observations, (ii) quantitative validation through direct correlation analyses, and (ii) qualitative comparisons with experimental data. The first category includes the fundamental biophysical rules, such as cohesin-mediated loop extrusion with sliding or binding/unbinding moves, replication kinetics, and the pushing effect of replication forks on LEFs. In RepliSage, all of these rules are implemented ad hoc; so, their validity can be assessed directly. Similarly, for the mitosis phase, the topoisomerase activity and condensin-driven loop extrusion are modeled. For quantitative validation, we compared the contact heatmaps generated by *RepliSage* with orthogonal publicly available Hi-C dataset, benchmarked as well as with heatmaps produced by random walk models, and consistently observing stronger correlations with our physically grounded model. Additionally, the average epigenetic state predicted by the node state model aligns with the first eigenvector of the Hi-C contact matrix, indicating that RT curves alone can reconstruct key structural features observed experimentally.

Qualitative validation further demonstrates RepliSage's ability to capture advanced aspects of cell cycle dynamics. Comparisons with NucDynamics models derived from single-cell Hi-C data [92] show chromosome decompaction during S phase with shorter loops, followed by compaction in G2/M and an increase in loop size due to condensin activity. These patterns are captured in RepliSage through the "moving barrier" action of replication forks, offering a potential explanation for the TAD weakening during S phase observed experimentally [35]. Another important layer of qualitative validation came from the cohesin HiChIP datasets generated for this study. To our knowledge, these are the first genome-wide 3C maps to demonstrate RS-induced alterations to 3D chromatin architecture in mammalian cells. Importantly, RepliSage successfully captures the RS-induced loop shortening observed in these experiments. This is in line with previous microscopic observations using the fluorescent DNA halo technique showing diminishment of length of chromatin loops following aphidicolin treatment [137]. Furthermore, stalled replication forks have been shown to hinder cohesin-dependent loop extrusion [138], which could contribute to the reduction in loop size observed under RS. At the level of compartments, chromatin architecture remains largely stable under RS. Consistent with this, the RT profiles, which are known to correlate strongly with compartments [40], are generally preserved upon mild RS, [110, 139]. In line with this studies

examining the impact of cellular perturbations, such as heat shock, ionizing irradiation, and DNA double-strand break induction, on chromatin architecture, show that overall chromosomal compartmentalization is largely preserved [109, 127, 128, 140–142]. The most significant changes are observed following DSB induction, with approximately 15% of genomic regions transitioning from the B to the A compartment [109, 110]. A similar trend is observed with RepliSage modeling, in which replication forks are assumed to organize epigenetic states. This leads to B-to-A switching due to increased origin activation and earlier replication of some regions under RS compared to normal conditions.

RepliSage relies on simplifications that improve computational efficiency but also introduce limitations. Fork velocities are sampled from a Gaussian distribution and remain constant throughout S phase, whereas in reality replication speed varies across the genome and changes over time, sometimes differently for left- and right-moving forks [47, 48]. Future improvements could include deeper integration of replication dynamics with Hi-C–derived structural features, such as fountain-like interaction patterns associated with fork progression [118] and G2-phase TAD structures linked to sister-chromatid organization [119]. Studying the interactions between the replication and loop extrusion in eukaryotes is challenging due to the dynamic and stochastic nature of both processes. This is further complicated by the existence of distinct functional states of cohesin (loop-extruding and cohesion-establishing) which differ in their modes of DNA engagement [143–145] and likely in their interactions with the replisome. Further experimental studies targeting interactions between the replisome and extrusive cohesin are essential to elucidate the underlying mechanisms. RepliSage is not designed to model cohesion establishment. Incorporating this complex biophysical process is an important area for future development. Extensions to the RepliSage could include modeling the effects of topological bonds between sister chromatids and transitions between the extruding- and cohesion-competent states of cohesin. Notably, cohesive cohesin can act as a barrier to loop-extruding cohesin [24, 146], highlighting the importance of integrating these distinct functional modes within a unified framework.

Overall, *RepliSage* aims to reconstruct qualitative cell cycle behavior and explore 3D chromatin dynamics under varying assumptions and parameters. Despite simplifications, it captures key biophysical features, including loop extrusion and replication-driven remodeling, using minimal input data. Its predictive capacity and modular design make it a versatile tool for investigating genome spatial organization and testing mechanistic hypotheses in DNA replication and chromatin dynamics.

Acknowledgements

Finally, we also gratefully acknowledge the Koren Lab for providing access to single-cell replication timing data whenever needed.

Author contributions: Sevastianos Korsak (Conceptualization [lead], Investigation [lead], Methodology [lead], Software [lead], Validation [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [lead]), Krzysztof H. Banecki (Conceptualization [supporting], Investigation [supporting], Methodology [supporting], Software [supporting], Validation [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), Abhishek Agarwal (Data curation [equal], Investigation [equal], Validation [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), Joanna Borkowska (Data curation [supporting], Resources [equal], Validation [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Piotr J. Górski (Conceptualization [supporting], Methodology [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Haoxi Chai (Data curation [equal], Resources [equal], Validation [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), Yijun Ruan (Resources [equal], Funding acquisition [equal], Supervision [equal]), Karolina Buka (Conceptualization [supporting], Data curation [lead], Investigation [supporting], Methodology [supporting], Project administration [lead], Resources [equal], Supervision [lead], Validation [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), and Dariusz Plewczynski (Conceptualization [lead], Data curation [supporting], Formal analysis [equal], Funding acquisition [lead], Investigation [equal], Methodology [supporting], Project administration [lead], Resources [supporting], Software [supporting], Supervision [lead], Validation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

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. P.G. and D.P. were funded by the European Union under the Horizon Europe grant OMINO (grant no. 101086321) and co-financed with funds from the Ministry of Science and Higher Education, Poland under the program entitled International Co-Financed Projects. 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." N.G.S. was performed thanks to Genomics Core Facility CeNT UW (RRID:SCR_022718), using NovaSeq 6000 platform financed by Polish Ministry of Science and Higher Education (decision no. 6817/IA/SP/2018 of 2018-04-10). Funding to pay the Open Access publication charges for this article was provided by University of Warsaw within OPUS grant from Polish National Science Centre (2020/37/B/NZ2/03757).

Data availability

The RepliSage model is available as an open-source project in GitHub https://github.com/SFGLab/RepliSage (https://doi.org/10.5281/zenodo.15915842). The updated LoopSage implementation is available at https://github.com/SFGLab/pyLoopSage [86]. Cohesin HiChIP data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE302711. Publicly available CTCF ChIA-PET, Hi-C, and Repli-seq datasets were obtained from ENCODE with the following accession numbers: CTCF ChIA-PET for GM12878 (ENCSR184YZV), CTCF ChIA-PET in K562 (ENCSR597AKG), Hi-C in GM128782 (ENCSR968KAY), Hi-C in K562 (ENCSR545YBD), Repli-seq in GM12878 (ENCSR000CXJ), and Repli-seq in K562 (ENCSR958IUL). Single-cell RT data in GM12878 were obtained from the Koren laboratory and K562 single-cell Hi-C–based chromatin structure benchmarks were taken from Chai *et al.* [92]. All simulation outputs and processed data supporting the findings of this study are available from the corresponding authors upon reasonable request.

References

1. Hansen AS. CTCF as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism. *Nucleus* 2020;11:132–48. https://doi.org/10.1080/19491034.2020.1782024.
2. Agarwal A, Korsak S, Choudhury A *et al.* The dynamic role of cohesin in maintaining human genome architecture. *BioEssays* 2023;45:2200240. https://doi.org/10.1002/bies.202200240
3. Lazniewski M, Dawson WK, Rusek AM *et al.* One protein to rule them all: the role of CCCTC-binding factor in shaping human genome in health and disease. *Semin Cell Dev Biol* 2019;90:114–27. https://doi.org/10.1016/j.semcdb.2018.08.003
4. Davidson IF, Peters JM. Genome folding through loop extrusion by SMC complexes. *Nat Rev Mol Cell Biol* 2021;22:445–64. https://doi.org/10.1038/s41580-021-00349-7
5. Lieberman-Aiden E, Van Berkum NL, Williams L *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009;326:289–93. https://doi.org/10.1126/science.1181369
6. Rao SS, Huntley MH, Durand NC *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159:1665–80. https://doi.org/10.1016/j.cell.2014.11.021
7. Rippe K. Liquid–liquid phase separation in chromatin. *CSH Perspect Biol* 2022;14:a040683.
8. Erdel F, Rippe K. Formation of chromatin subcompartments by phase separation. *Biophys J* 2018;114:2262–70. https://doi.org/10.1016/j.bpj.2018.03.011
9. Fortin JP, Hansen KD. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome biol* 2015;16:1–23. https://doi.org/10.1186/s13059-015-0741-y
10. Harris HL, Gu H, Olshansky M *et al.* Chromatin alternates between A and B compartments at kilobase scale for subgenic organization. *Nat Commun* 2023;14:3303. https://doi.org/10.1038/s41467-023-38429-1
11. Cremer T, Cremer M. Chromosome territories. *CSH Perspect Biol* 2010;2:a003889.
12. Ozer G, Luque A, Schlick T. The chromatin fiber: multiscale problems and approaches. *Curr Opin Struct Biol* 2015;31:124–39. https://doi.org/10.1016/j.sbi.2015.04.002
13. Weinreb C, Raphael BJ. Identification of hierarchical chromatin domains. *Bioinformatics* 2016;32:1601–9. https://doi.org/10.1093/bioinformatics/btv485
14. Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev* 2011;21:175–86. https://doi.org/10.1016/j.gde.2011.01.022
15. Woodcock CL, Ghosh RP. Chromatin higher-order structure and dynamics. *CSH Perspect Biol* 2010;2:a000596.
16. Korsak S, Banecki K, Plewczynski D. Multiscale molecular modeling of chromatin with MultiMM: from nucleosomes to the whole genome. *Comput Struct Biotechnol J* 2024;23:3537–48. https://doi.org/10.1016/j.csbj.2024.09.025
17. Fierz B, Poirier MG. Biophysics of chromatin dynamics. *Annu Rev Biophys* 2019;48:321–45. https://doi.org/10.1146/annurev-biophys-070317-032847
18. Hübner MR, Spector DL. Chromatin dynamics. *Annu Rev Biophys* 2010;39:471–89. https://doi.org/10.1146/annurev.biophys.093008.131348
19. Prieto EI, Maeshima K. Dynamic chromatin organization in the cell. *Essays Biochem* 2019;63:133–45. https://doi.org/10.1042/EBC20180054
20. Haering CH, Löwe J, Hochwagen A *et al.* Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell* 2002;9:773–88. https://doi.org/10.1016/S1097-2765(02)00515-4
21. Haering CH, Farcas AM, Arumugam P *et al.* The cohesin ring concatenates sister DNA molecules. *Nature* 2008;454:297–301. https://doi.org/10.1038/nature07098
22. Srinivasan M, Fumasoni M, Petela NJ *et al.* Cohesion is established during DNA replication utilising chromosome associated cohesin rings as well as those loaded de novo onto nascent DNAs. *Elife* 2020;9:e56611. https://doi.org/10.7554/eLife.56611
23. Ochs F, Green C, Szczurek AT *et al.* Sister chromatid cohesion is mediated by individual cohesin complexes. *Science* 2024;383:1122–30. https://doi.org/10.1126/science.adl4606
24. Zhao H, Shu L, Qin S *et al.* Extensive mutual influences of SMC complexes shape 3D genome folding. *Nature* 2025;640:1–11.
25. Di Pierro M, Cheng RR, Lieberman Aiden E *et al. De novo* prediction of human chromosome structures: epigenetic marking patterns encode genome architecture. *Proc Natl Acad Sci* 2017;114:12126–31. https://doi.org/10.1073/pnas.1714980114
26. Zhang H, Tian XJ, Mukhopadhyay A *et al.* Statistical mechanics model for the dynamics of collective epigenetic histone modification. *Phys Rev Lett* 2014;112:068101. https://doi.org/10.1103/PhysRevLett.112.068101
27. Mo Q, Liang F. A hidden Ising model for ChIP-chip data analysis. *Bioinformatics* 2010;26:777–83. https://doi.org/10.1093/bioinformatics/btq032
28. Mo Q, Liang F. Bayesian modeling of ChIP-chip data through a high-order Ising Model. *Biometrics* 2010;66:1284–94. https://doi.org/10.1111/j.1541-0420.2009.01379.x
29. Michieletto D, Orlandini E, Marenduzzo D. Polymer model with epigenetic recoloring reveals a pathway for the *de novo* establishment and 3D organization of chromatin domains. *Phys Rev X* 2016;6:041047.
30. Adachi K, Kawaguchi K. Chromatin state switching in a polymer model with mark-conformation coupling. *Phys Rev E* 2019;100:060401. https://doi.org/10.1103/PhysRevE.100.060401
31. McCoy BM, Wu TT. *The Two-Dimensional Ising Model.* Cambridge, MA and London, England: Harvard University Press, 1973. https://doi.org/10.4159/harvard.9780674180758
32. Beaudin L. A review of the Potts model. *Rose-Hulman Undergraduate Mathematics Journal* 2007;8:13.
33. Pelham-Webb B, Murphy D, Apostolou E. Dynamic 3D chromatin reorganization during establishment and maintenance of pluripotency. *Stem Cell Rep* 2020;15:1176–95. https://doi.org/10.1016/j.stemcr.2020.10.012

34. Alabert C, Groth A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* 2012;13:153–67. https://doi.org/10.1038/nrm3288
35. Nagano T, Lubling Y, Várnai C *et al.* Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* 2017;547:61–7. https://doi.org/10.1038/nature23001
36. Costa A, Diffley JF. The initiation of eukaryotic DNA replication. *Annu Rev Biochem* 2022;91:107–31. https://doi.org/10.1146/annurev-biochem-072321-110228
37. Hu Y, Stillman B. Origins of DNA replication in eukaryotes. *Mol Cell* 2023;83:352–72. https://doi.org/10.1016/j.molcel.2022.12.024
38. Ge XQ, Jackson DA, Blow JJ. Dormant origins licensed by excess Mcm2–7 are required for human cells to survive replicative stress. *Genes Dev* 2007;21:3331–41. https://doi.org/10.1101/gad.457807
39. Ibarra A, Schwob E, Méndez J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc Natl Acad Sci* 2008;105:8956–61. https://doi.org/10.1073/pnas.0803978105
40. Marchal C, Sima J, Gilbert DM. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol* 2019;20:721–37. https://doi.org/10.1038/s41580-019-0162-y
41. Kanke M, Tahara E, Huis in't Veld PJ *et al.* Cohesin acetylation and Wapl–Pds5 oppositely regulate translocation of cohesin along DNA. *EMBO J* 2016;35:2686–98. https://doi.org/10.15252/embj.201695756
42. Cameron G, Gruszka DT, Gruar R *et al.* Sister chromatid cohesion establishment during DNA replication termination. *Science* 2024;384:119–24. https://doi.org/10.1126/science.adf0224
43. Glaser S, Molodtsov MI, Diffley JF *et al.* Replisome passage through the cohesin ring. *Cell* 2025;188:5618–31. https://doi.org/10.1016/j.cell.2025.08.028
44. Liao Q, Brandão HB, Ren Z *et al.* Replisomes restrict SMC translocation *in vivo*. *Nat Commun* 2025;16:7151. https://doi.org/10.1038/s41467-025-62596-y
45. Dequeker BJ, Scherr MJ, Brandão HB *et al.* MCM complexes are barriers that restrict cohesin-mediated loop extrusion. *Nature* 2022;606:197–203. https://doi.org/10.1038/s41586-022-04730-0
46. Jackson DA, Pombo A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J Cell Biol* 1998;140:1285–95. https://doi.org/10.1083/jcb.140.6.1285
47. van den Berg J, van Batenburg V, Geisenberger C *et al.* Quantifying DNA replication speeds in single cells by scEdU-seq. *Nat Methods* 2024;21:1175–84. https://doi.org/10.1038/s41592-024-02308-4
48. Jones MJ, Rai SK, Pfuderer PL *et al.* A high-resolution, nanopore-based artificial intelligence assay for DNA replication stress in human cancer cells. *Nat Commun* 2025;16:7732. https://doi.org/10.1038/s41467-025-63168-w
49. Conti C, Saccà B, Herrick J *et al.* Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Mol Biol Cell* 2007;18:3059–67. https://doi.org/10.1091/mbc.e06-08-0689
50. Milo R, Phillips R. *Cell biology by the numbers*. New York, NY: Garland Science, 2015.https://doi.org/10.1201/9780429258770
51. Corsi F, Rusch E, Goloborodko A. Loop extrusion rules: the next generation. *Curr Opin Genet Dev* 2023;81:102061. https://doi.org/10.1016/j.gde.2023.102061
52. Ganji M, Shaltiel IA, Bisht S *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science* 2018;360:102–5. https://doi.org/10.1126/science.aar7831
53. Zeman MK, Cimprich KA. Causes and consequences of replication stress. *Nat Cell Biol* 2014;16:2–9. https://doi.org/10.1038/ncb2897
54. Lecona E, Fernández-Capetillo O. Replication stress and cancer: it takes two to tango. *Exp Cell Res* 2014;329:26–34. https://doi.org/10.1016/j.yexcr.2014.09.019
55. Gaillard H, García-Muse T, Aguilera A. Replication stress and cancer. *Nat Rev Cancer* 2015;15:276–89. https://doi.org/10.1038/nrc3916
56. Macheret M, Halazonetis TD. DNA replication stress as a hallmark of cancer. *Annu Rev Pathol Mech Dis* 2015;10:425–48. https://doi.org/10.1146/annurev-pathol-012414-040424
57. Gaggioli V, Lo CS, Reverón-Gómez N *et al.* Dynamic *de novo* heterochromatin assembly and disassembly at replication forks ensures fork stability. *Nat Cell Biol* 2023;25:1017–32. https://doi.org/10.1038/s41556-023-01167-z
58. Jasencakova Z, Scharf AN, Ask K *et al.* Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell* 2010;37:736–43. https://doi.org/10.1016/j.molcel.2010.01.033
59. Nasmyth K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet* 2001;35:673–745. https://doi.org/10.1146/annurev.genet.35.102401.091334
60. Hirano T. Mitotic genome folding. *J Cell Biol* 2025;224:e202504075. https://doi.org/10.1083/jcb.202504075
61. Green LC, Kalitsis P, Chang TM *et al.* Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *J Cell Sci* 2012;125:1591–604.
62. Nagasaka K, Hossain MJ, Roberti MJ *et al.* Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat Cell Biol* 2016;18:692–9. https://doi.org/10.1038/ncb3353
63. Gibcus JH, Samejima K, Goloborodko A *et al.* A pathway for mitotic chromosome formation. *Science* 2018;359:eaao6135. https://doi.org/10.1126/science.aao6135
64. Samejima K, Samejima I, Vagnarelli P *et al.* Mitotic chromosomes are compacted laterally by KIF4 and condensin and axially by topoisomerase IIα. *J Cell Biol* 2012;199:755–70. https://doi.org/10.1083/jcb.201202155
65. Hildebrand EM, Polovnikov K, Dekker B *et al.* Mitotic chromosomes are self-entangled and disentangle through a topoisomerase-II-dependent two-stage exit from mitosis. *Mol Cell* 2024;84:1422–41. https://doi.org/10.1016/j.molcel.2024.02.025
66. Earnshaw WC, Heck M. Localization of topoisomerase II in mitotic chromosomes. *J Cell Biol* 1985;100:1716–25. https://doi.org/10.1083/jcb.100.5.1716
67. Nitiss JL. DNA topoisomerase II and its growing repertoire of biological functions. *Nat Rev Cancer* 2009;9:327–37. https://doi.org/10.1038/nrc2608
68. Rose D, Thomas W, Holm C. Segregation of recombined chromosomes in meiosis I requires DNA topoisomerase II. *Cell* 1990;60:1009–17. https://doi.org/10.1016/0092-8674(90)90349-J
69. Holm C, Goto T, Wang JC *et al.* DNA topoisomerase II is required at the time of mitosis in yeast. *Cell* 1985;41:553–63. https://doi.org/10.1016/S0092-8674(85)80028-3
70. McGuinness BE, Hirota T, Kudo NR *et al.* Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. *PLoS Biol* 2005;3:e86. https://doi.org/10.1371/journal.pbio.0030086
71. Retkute R, Nieduszynski CA, De Moura A. Mathematical modeling of genome replication. *Phys Rev E* 2012;86:031916. https://doi.org/10.1103/PhysRevE.86.031916
72. Yousefi R, Rowicka M. Stochasticity of replication forks' speeds plays a key role in the dynamics of DNA replication. *PLoS Comput Biol* 2019;15:e1007519. https://doi.org/10.1371/journal.pcbi.1007519
73. Baker A, Bechhoefer J. Inferring the spatiotemporal DNA replication program from noisy data. *Phys Rev E* 2014;89:032703. https://doi.org/10.1103/PhysRevE.89.032703

74. Herrick J, Jun S, Bechhoefer J *et al.* Kinetic model of DNA replication in eukaryotic organisms. *J Mol Biol* 2002;320:741–50. https://doi.org/10.1016/S0022-2836(02)00522-3
75. de Moura AP, Retkute R, Hawkins M *et al.* Mathematical modelling of whole chromosome replication. *Nucleic Acids Res* 2010;38:5623–33. https://doi.org/10.1093/nar/gkq343
76. Gauthier MG, Norio P, Bechhoefer J. Modeling inhomogeneous DNA replication kinetics. *PLoS One* 2012;7:e32053. https://doi.org/10.1371/journal.pone.0032053
77. Yang SCH, Gauthier MG, Bechhoefer J. Computational methods to study kinetics of DNA replication. *DNA Repl* 2009;521:555–73.
78. Koutroumpas K. *Stochastic hybrid systems for DNA replication modeling*. ETH Zurich, 2010.
79. Koutroumpas K, Lygeros J. Modeling and analysis of DNA replication. *Automatica* 2011;47:1156–64. https://doi.org/10.1016/j.automatica.2011.02.007
80. Lygeros J, Koutroumpas K, Dimopoulos S *et al.* Stochastic hybrid modeling of DNA replication across a complete genome. *Proc Natl Acad Sci* 2008;105:12295–300. https://doi.org/10.1073/pnas.0805549105
81. Berkemeier F, Cook PR, Boemo MA. DNA replication timing reveals genome-wide features of transcription and fragility. *Nat Commun* 2025;16:4658. https://doi.org/10.1038/s41467-025-59991-w
82. D'Asaro D, Tortora MM, Vaillant C *et al.* DNA replication and polymer chain duplication reshape the genome in space and time. *Phys Rev X* 2024;14:041020.
83. D'Asaro D, Arbona JM, Piveteau V *et al.* Genome-wide modeling of DNA replication in space and time confirms the emergence of replication specific patterns *in vivo* in eukaryotes. *Genome Biol* 2025;26:431. https://doi.org/10.1186/s13059-025-03872-4
84. Forte G, Buonomo S, Cook P *et al.* Modeling the 3D spatiotemporal organization of chromatin replication. *PRX Life* 2024;2:033014. https://doi.org/10.1103/PRXLife.2.033014
85. Löb D, Lengert N, Chagin V *et al.* 3D replicon distributions arise from stochastic initiation and domino-like DNA replication progression. *Nat Commun* 2016;7:11207. https://doi.org/10.1038/ncomms11207
86. Korsak S, Plewczynski D. LoopSage: an energy-based Monte Carlo approach for the loop extrusion modelling of chromatin. *Methods* 2024; 223:106–17. https://doi.org/10.1016/j.ymeth.2024.01.015
87. Gross T, Blasius B. Adaptive coevolutionary networks: a review. *J R Soc Interf* 2008;5:259–71. https://doi.org/10.1098/rsif.2007.1229
88. Toruniewska J, Suchecki K, Hołyst JA. Unstable network fragmentation in co-evolution of Potts spins and system topology. *Physica A: Stat Mech Appl* 2016;460:1–15. https://doi.org/10.1016/j.physa.2016.04.029
89. Fronczak P, Fronczak A, Hołyst JA. Self-organized criticality and coevolution of network structure and dynamics. *Phys Rev E* 2006;73:046117. https://doi.org/10.1103/PhysRevE.73.046117
90. Eastman P, Swails J, Chodera JD *et al.* OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 2017;13:e1005659. https://doi.org/10.1371/journal.pcbi.1005659
91. Eastman P, Galvelis R, Peláez RP *et al.* OpenMM 8: molecular dynamics simulation with machine learning potentials. *J Phys Chem B* 2023;128:109–16. https://doi.org/10.1021/acs.jpcb.3c06662
92. Chai H, Huang X, Xiong G *et al.* Tri-omic single-cell mapping of the 3D epigenome and transcriptome in whole mouse brains throughout the lifespan. *Nat Methods* 2025;22:1–14.
93. Stevens TJ, Lando D, Basu S *et al.* 3D structure of individual mammalian genomes studied by single cell Hi-C. *Nature* 2017;544:59–64. https://doi.org/10.1038/nature21429.3D
94. Buka K, Parteka-Tojek Z, Agarwal A *et al.* Improved cohesin HiChIP protocol and bioinformatic analysis for robust detection of chromatin loops and stripes. *Commun Biol* 2025;8:437. https://doi.org/10.1038/s42003-025-07847-w
95. Bechhoefer J, Rhind N. Replication timing and its emergence from stochastic processes. *Trends Genet* 2012;28:374–81. https://doi.org/10.1016/j.tig.2012.03.011
96. Finn JM. *Classical mechanics*. Sudbury, MA, USA: Jones & Bartlett Publishers, 2009.
97. Newman ME, Barkema GT. *Monte Carlo methods in statistical physics*. Oxford, United Kingdom: Clarendon Press, 1999.
98. Hildebrand EM, Dekker J. Mechanisms and functions of chromosome compartmentalization. *Trends Biochem Sci* 2020;45:385–96. https://doi.org/10.1016/j.tibs.2020.01.002
99. Singh PB, Newman AG. On the relations of phase separation and Hi-C maps to epigenetics. *Roy Soc Open Sci* 2020;7:191976. https://doi.org/10.1098/rsos.191976
100. Kardar M. *Statistical physics of particles*. Cambridge: Cambridge University Press, 2007.
101. Hansen AS, Pustova I, Cattoglio C *et al.* CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *elife* 2017;6:e25776. https://doi.org/10.7554/eLife.25776
102. Cooper G, Adams KW. *The cell: a molecular approach*. Oxford: Oxford University Press, 2022. https://doi.org/10.1093/hesc/9780197583746.001.0001
103. Gindin Y, Valenzuela MS, Aladjem MI *et al.* A chromatin structure-based model accurately predicts DNA replication timing in human cells. *Mol Syst Biol* 2014;10:722. https://doi.org/10.1002/msb.134859
104. Villa-Hernández S, Bermejo R. Replisome-cohesin interfacing: a molecular perspective. *BioEssays* 2018;40:1800109. https://doi.org/10.1002/bies.201800109
105. Meyer-Nava S, Rivera-Mulia JC. You shall not pass! Unveiling the barriers for cohesin-mediated loop extrusion. *Mol Cell* 2022;82:2541–3. https://doi.org/10.1016/j.molcel.2022.06.028
106. Bayong E, Diep H, Dotsenko V. Potts model with long-range interactions in one dimension. *Phys Rev Lett* 1999;83:14. https://doi.org/10.1103/PhysRevLett.83.14
107. Oji A, Choubani L, Miura H *et al.* Structure and dynamics of nuclear A/B compartments and subcompartments. *Curr Opin Cell Biol* 2024;90:102406. https://doi.org/10.1016/j.ceb.2024.102406
108. Giles KA, Taberlay PC, Cesare AJ *et al.* Roles for the 3D genome in the cell cycle, DNA replication, and double strand break repair. *Front Cell Dev Biol* 2025;13:1548946. https://doi.org/10.3389/fcell.2025.1548946
109. Zagelbaum J, Schooley A, Zhao J *et al.* Multiscale reorganization of the genome following DNA damage facilitates chromosome translocations via nuclear actin polymerization. *Nat Struct Mol Biol* 2023;30:99–106. https://doi.org/10.1038/s41594-022-00893-6
110. Courtot L, Bournique E, Maric C *et al.* Low replicative stress triggers cell-type specific inheritable advanced replication timing. *Int J Mol Sci* 2021;22:4959. https://doi.org/10.3390/ijms22094959
111. Jost D, Carrivain P, Cavalli G *et al.* Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res* 2014;42:9553–61. https://doi.org/10.1093/nar/gku698
112. Bates FS, Fredrickson GH. Block copolymer thermodynamics: theory and experiment. *Annu Rev Phys Chem* 1990;41:525–57. https://doi.org/10.1146/annurev.pc.41.100190.002521
113. Zhou R, Gao YQ. Polymer models for the mechanisms of chromatin 3D folding: review and perspective. *Phys Chem Chem Phys* 2020;22:20189–201. https://doi.org/10.1039/D0CP01877E
114. Lam SK, Pitrou A, Seibert S. Numba: A llvm-based python jit compiler. In: *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*. Berkeley, CA, USA: USENIX Association, 2015, 1–6.

115. Massey DJ, Koren A. High-throughput analysis of single human cells reveals the complex nature of DNA replication timing control. *Nat Commun* 2022;13:2402. https://doi.org/10.1038/s41467-022-30212-y
116. Samejima K, Gibcus JH, Abraham S *et al.* Rules of engagement for condensins and cohesins guide mitotic chromosome formation. *Science* 2025;388:eadq1709. https://doi.org/10.1126/science.adq1709
117. Vian L, Pękowska A, Rao SS *et al.* The energetics and physiological impact of cohesin extrusion. *Cell* 2018;173:1165–78. https://doi.org/10.1016/j.cell.2018.03.072
118. Liu Y, Zhangding Z, Liu X *et al.* Fork coupling directs DNA replication elongation and termination. *Science* 2024;383:1215–22. https://doi.org/10.1126/science.adj7606
119. Mitter M, Gasser C, Takacs Z *et al.* Conformation of sister chromatids in the replicated human genome. *Nature* 2020;586:139–44. https://doi.org/10.1038/s41586-020-2744-4
120. Imakaev MV, Fudenberg G, Mirny LA. Modeling chromosomes: beyond pretty pictures. *FEBS Lett* 2015;589:3031–6. https://doi.org/10.1016/j.febslet.2015.09.004
121. Kubica J, Korsak S, Banecki KH *et al.* The challenge of chromatin model comparison and validation: a project from the first international 4D Nucleome Hackathon. *PLOS Comput Biol* 2025;21:e1013358. https://doi.org/10.1371/journal.pcbi.1013358
122. Banecki K, Korsak S, Plewczynski D. Advancements and future directions in single-cell Hi-C based 3D chromatin modeling. *Comput Struct Biotechnol J* 2024;23:3549–58. https://doi.org/10.1016/j.csbj.2024.09.026
123. Affonso L, Bissacot R, Endo EO *et al.* Long-range Ising models: contours, phase transitions and decaying fields. *J Eur Math Soc* 2024;27:1679–714. https://doi.org/10.4171/jems/1529
124. Anglès d'Auriac JC, Iglói F. Phase transitions of the random-bond Potts chain with long-range interactions. *Phys Rev E* 2016;94:062126. https://doi.org/10.1103/PhysRevE.94.062126
125. Lecce I, Picco M, Santachiara R. Magnetic exponent for the long-range bond-disordered Potts model. *Phys Rev E* 2024;110:064154. https://doi.org/10.1103/PhysRevE.110.064154
126. Biskup M, Chayes L, Crawford N. Mean-field driven first-order phase transitions in systems with long-range interactions. *J Stat Phys* 2006;122:1139–93. https://doi.org/10.1007/s10955-005-8072-0
127. Ray J, Munn PR, Vihervaara A *et al.* Chromatin conformation remains stable upon extensive transcriptional changes driven by heat shock. *Proc Natl Acad Sci* 2019;116:19431–9. https://doi.org/10.1073/pnas.1901244116
128. Li L, Lyu X, Hou C *et al.* Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* 2015;58:216–31. https://doi.org/10.1016/j.molcel.2015.02.023
129. Schwarzer W, Abdennur N, Goloborodko A *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 2017;551:51–6. https://doi.org/10.1038/nature24281
130. Nuebler J, Fudenberg G, Imakaev M *et al.* Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Biophys J* 2018;114:30a. https://doi.org/10.1016/j.bpj.2017.11.211
131. Goloborodko A, Marko JF, Mirny LA. Chromosome compaction by active loop extrusion. *Biophys J* 2016;110:2162–8. https://doi.org/10.1016/j.bpj.2016.02.041
132. Crippa M, Zhan Y, Tiana G. Effective model of loop extrusion predicts chromosomal domains. *Phys Rev E* 2020;102:032414. https://doi.org/10.1103/PhysRevE.102.032414
133. Banigan EJ, van den Berg AA, Brandão HB *et al.* Chromosome organization by one-sided and two-sided loop extrusion. *eLife* 2020;9:e53558. https://doi.org/10.7554/eLife.53558
134. Banigan EJ, Mirny LA. The interplay between asymmetric and symmetric DNA loop extrusion. *eLife* 2020;9:e63528. https://doi.org/10.7554/eLife.63528
135. Fudenberg G, Abdennur N, Imakaev M *et al.* Emerging evidence of chromosome folding by loop extrusion. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press, 2018. https://doi.org/10.1101/264648
136. Le Treut G, Képès F, Orland H. A polymer model for the quantitative reconstruction of chromosome architecture from HiC and GAM data. *Biophys J* 2018;115:2286–94.
137. Courbet S, Gay S, Arnoult N *et al.* Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature* 2008;455:557–60.
138. Jeppsson K, Sakata T, Nakato R *et al.* Cohesin-dependent chromosome loop extrusion is limited by transcription and stalled replication forks. *Sci Adv* 2022;8:eabn7063.
139. Sarni D, Kerem B. Oncogene-induced replication stress drives genome instability and tumorigenesis. *Int J Mol Sci* 2017;18:1339.
140. Lyu X, Rowley MJ, Corces VG. Architectural proteins and pluripotency factors cooperate to orchestrate the transcriptional response of hESCs to temperature stress. *Mol Cell* 2018;71:940–55.
141. Sanders JT, Freeman TF, Xu Y *et al.* Radiation-induced DNA damage and repair effects on 3D genome organization. *Nat Commun* 2020;11:6178. https://doi.org/10.1038/s41467-020-20047-w
142. Arnould C, Rocher V, Finoux AL *et al.* Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature* 2021;590:660–5. https://doi.org/10.1038/s41586-021-03193-z
143. Srinivasan M, Scheinost JC, Petela NJ *et al.* The cohesin ring uses its hinge to organize DNA using non-topological as well as topological mechanisms. *Cell* 2018;173:1508–19. https://doi.org/10.1016/j.cell.2018.04.015
144. Davidson IF, Bauer B, Goetz D *et al.* DNA loop extrusion by human cohesin. *Science* 2019;366:1338–45. https://doi.org/10.1126/science.aaz3418
145. Kim Y, Shi Z, Zhang H *et al.* Human cohesin compacts DNA by loop extrusion. *Science* 2019;366:1345–9. https://doi.org/10.1126/science.aaz4475
146. Bastié N, Chapard C, Cournac A *et al.* Sister chromatid cohesion halts DNA loop expansion. *Mol Cell* 2024;84:1139–48. https://doi.org/10.1016/j.molcel.2024.02.004

Received: August 25, 2025. **Revised:** May 26, 2026. **Accepted:** May 31, 2026