Biophysical Journal, Volume *122*

Supplemental information

Compartmentalization with nuclear landmarks yields random, yet precise, genome organization

Kartik Kamat, Zhuohan Lao, Yifeng Qi, Yuchuan Wang, Jian Ma, and Bin Zhang

Manuscript submitted to **Biophysical** *Journal*

Article

Supporting Information: Compartmentalization with Nuclear Landmarks Yields Random yet Precise Genome Organization

Kartik Kamat[1], Zhuohan Lao[1], Yifeng Qi[1], Yuchuan Wang[2], Jian Ma[2], and Bin Zhang[1,*]

[1]Departments of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
[*]Correspondence: binz@mit.edu

SYSTEM SETUP

The nuclear lamina as a particle-based mesh

The nuclear envelope is a highly complex structure with two lipid membrane and a large number of protein molecules (1). To account for chromatin-lamina interactions while keeping the model simple, we used discrete particles uniformly placed on a sphere to approximate the nuclear lamina. No explicit dynamics was included to account for nuclear deformation, though it could be implemented rather straightforwardly.

The positions of the coarse-grained particles were determined as follows. A straightforward way of generating lamina is by drawing random points uniformly distributed on the surface of the sphere with radius R. Clearly, the process may lead to overlapping lamina particles (or huge gaps in the lamina) and a large number of lamina particles may be required to form a caged network that can contain the chromatin inside the nucleus. Additionally, the generated network is not a strictly reproducible starting point because the lamina configuration depends on the state of the random number generator.

We circumvent these issues by using the Fibonacci grid to initialize the lamina particles that allows a uniform and an almost equidistant network of lamina particles on the surface of the nucleus.(2, 3) The angular coordinates (θ, ϕ) associated with the lamina particles are

$$\begin{aligned}\theta_i &= \frac{\pi}{2} + \sin^{-1}\frac{2i}{2N+1}\\ \phi_i &= 2\pi i \Phi^{-1}\end{aligned} \tag{S1}$$

where $i \in \{-N, -N+1, ..., 0, ..., N-1, N\}$ such that $N_L = 2N+1$ and $\Phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. The spherical coordinates (R, θ_i, ϕ_i) for an i^{th} lamina particle are finally converted to Cartesian coordinates. We set the radius $R = 13.0\sigma = 5\mu$m. As mentioned in the text, σ is the length unit and is related to the diameter of a chromatin bead ($1\sigma = 385$ nm).

The number of lamina particles (N_L) were chosen such that for every lamina particle there is at least one neighboring lamina particle within a distance of 1σ. The *query_radius* function under the KDTree method available in the scikit library (for Python programming language) was used to identify the neighbors. We determined that 2087 lamina particles are required to meet the aforementioned constraint for a nucleus with radius 13σ.

Nucleoli and speckles as phase-separated droplets

Nucleoli

We fixed the number of nucleolus particles ($N_n = 300$) based on the experimentally reported values of nuclear protein NPM1 concentration as done in a recent study (4).

The size of nucleolar particles (σ_n) was estimated as follows. We accounted for the multi-droplet state of the nucleolus where the typical number of nucleoli experimentally observed within a cell ranges from 2-5. Accordingly, we used a three droplet state as a reference and a space filling approximation within a individual nucleoli droplet (with $N_n/3$ particles)

$$\frac{(4\pi/3)(2^{1/6}\sigma_n/2)^3(N_n/3)}{(4\pi/3)R_N^3} = \left(\frac{R_n}{R_N}\right)^3 \tag{S2}$$

where the subscript N denotes the nucleus and $2^{1/6}\sigma_n/2$ is the effective radius of a nucleoli particle that interacts via a Lennard-Jones potential. For a typical size of the nucleoli of $R_n = 0.5\ \mu$m and $R_N = 13.0\sigma = 5\ \mu$m, the above equation yields $\sigma_n = 0.5$.

Speckles

We estimated the number of speckle particles per cluster using the protein densities calculated from refractive index measurements (5). The relative mass densities of the protein concentrations in the speckle and nucleolus droplet is given by

$$\frac{N \times m/0.3^3}{100 \times m/0.5^3} = \frac{170}{203}, \tag{S3}$$

where we used a speckle radius of 0.3 μm that yields $N \approx 20$. We target 30 speckles in the model to be commensurate with the experimentally reported range of 20-50 speckles. Accordingly, we use a total of 600 speckle particles. Using the speckle radius of 0.3 μm with 20 particles in the droplet in equation S2 yields a $\sigma = 0.5$ for the size of speckle particles.

ENERGY FUNCTION OF THE NUCLEUS MODEL

Hi-C inspired interactions for the diploid human genome

The energy function of the genome model is defined as

$$U_{\mathrm{Genome}}(\mathbf{r}) = U(\mathbf{r}) + U_{\mathrm{sc}}(\mathbf{r}) + U_{\mathrm{ideal}}(\mathbf{r}) + U_{\mathrm{compt}}(\mathbf{r}). \tag{S4}$$

$U(\mathbf{r})$ represents a generic potential applied to each chromosome to ensure the polymeric topology of chromosomes:

$$U(\mathbf{r}) = \sum_i \left[u_{\mathrm{bond}}(r_{i,i+1}) + u_{\mathrm{angle}}(\vec{r}_{i,i+1}, \vec{r}_{i+1,i+2})\right], \tag{S5}$$

where $u_{\mathrm{bond}}(r_{i,i+1})$ and $u_{\mathrm{angle}}(r_{i,i+1}, r_{i+1,i+2})$ are the bonding and angular potential applied for neighboring beads to ensure the connectivity of the chromatin chain.

$$\begin{aligned} u_{\mathrm{bond}}(r_{i,i+1}) &= -\frac{1}{2}K_b R_0^2 \ln\left[1 - \left(\frac{r_{i,i+1}}{R_0}\right)^2\right], K_b = 30\epsilon, R_0 = 1.5\sigma \\ u_{\mathrm{angle}}(\vec{r}_{i,i+1}, \vec{r}_{i+1,i+2}) &= K_a\left[1 - \cos(\theta - \pi)\right], K_a = 2\epsilon, \cos\theta = \frac{\vec{r}_{i,i+1} \cdot \vec{r}_{i+1,i+2}}{\left|\vec{r}_{i,i+1}\right| \cdot \left|\vec{r}_{i+1,i+2}\right|}. \end{aligned} \tag{S6}$$

The soft-core potential provides excluded volume effects for pairs of beads from the same or different chromosomes

$$U_{\mathrm{sc}}(\mathbf{r}) = \sum_{j>i} u_{\mathrm{sc}}(r_{ij}). \tag{S7}$$

$u_{\mathrm{sc}}(r_{ij})$ is a soft-core potential added to each pair formed by beads index i and j to account for the excluded volume effect while allowing finite probability of cross-over of polymer chains.

$$u_{\mathrm{sc}}(r_{ij}) = \begin{cases} 0.5E_{\mathrm{cut}}\left(1 + \tanh\left[\frac{2U_{\mathrm{LJ}}(r_{ij})}{E_{\mathrm{cut}}} - 1\right]\right), & r_{ij} \le r_{\mathrm{cut}} \\ U_{\mathrm{LJ}}(r_{ij}), & r_{\mathrm{cut}} < r_{ij} \le 2^{1/6}\sigma \\ 0, & r_{ij} > 2^{1/6}\sigma \end{cases} \tag{S8}$$

which corresponds to the Lennard-Jones potential capped off at a finite volume within a replusive core to allow for chain crossing at finite energy cost. $E_{\mathrm{cut}} = 4\epsilon$ and r_{cut} is chosen as the distance at which $U_{\mathrm{LJ}}(r) = 0.5E_{\mathrm{cut}}$.

$U_{\mathrm{ideal}}(\mathbf{r})$ is the intra-chromosomal potential applied to genomic loci within the same chromosome, while $U_{\mathrm{compt}}(\mathbf{r})$ is the compartment-specific interaction potential. The ideal potential adopts the following form:

$$U_{\mathrm{ideal}}(\mathbf{r}) = \sum_I \sum_{i,j \in I} \alpha_{\mathrm{ideal}}(|i-j|) f(r_{ij}) \tag{S9}$$

where I indexes over each chromosome and i and j index over pair of beads on that chromosome. $\alpha_{\mathrm{ideal}}(|i-j|)$ depends only on the sequence separation between two beads i and j. $f(r_{ij})$ measures the probability of contact formation for two loci separated by a distance of r_{ij}, and its ensemble average corresponds to the contact probability measured in Hi-C experiments. $f(r_{ij})$ adopts the form:

$$f(r_{ij}) = \begin{cases} \frac{1}{2}\left[1+\tanh\left[\eta(r_c - r_{ij})\right]\right], & r_{ij} \le r_c \\ \frac{1}{2}\left(r_c/r\right)^4, & r_{ij} > r_c \end{cases} \tag{S10}$$

Numerical values of r_c and η were determined from imaging data as detailed in the next section. Additionally, it is to be noted that we truncate the ideal potential to be applicable for a sequence separation less than or equal to 100 MB and set the alpha parameters for the higher sequence separations to be zero. This is done because the number of occurrences of bead pairs with sequence separation greater than 100 MB drastically reduces. Incorporating these interactions for improving the model further is straightforward but the resulting model would have higher number of parameters than used in this study.

Similar to the ideal potential discussed above we have

$$U_{\text{compt}}(\mathbf{r}) = \sum_{i,j} \alpha_{\text{compt}}(T_i, T_j) f(r_{ij}) \tag{S11}$$

T_i and T_j denote the compartment types of beads i and j which can be A, B or C.

Mathematical expressions for the various energy terms in $U_{\text{Genome}}(\mathbf{r})$ were designed such that their ensemble averages can be mapped onto combinations of contact frequencies measured in Hi-C. The correspondence between the energy functions and Hi-C measurements allows model parameterization with an efficient maximum entropy optimization algorithm. Specifically, $\alpha_{\text{ideal}}(|i-j|)$ and $\alpha_{\text{compt}}(T_i, T_j)$ were tuned to satisfy the following constraints:

$$\begin{aligned} \left\langle \sum_I \sum_{i,j \in I} f(r_{ij}) \delta_{|i-j|,s} \right\rangle &= \sum_I \sum_{i,j \in I} f_{ij}^{\text{exp}} \delta_{|i-j|,s}, \quad \text{for } s = 1, \cdots, n-1 \\ \left\langle \sum_{i,j} f(r_{ij}) \delta_{T_i,T_1} \delta_{T_j,T_2} \right\rangle &= \sum_{i,j} f_{ij}^{\text{exp}} \delta_{T_i,T_1} \delta_{T_j,T_2}, \quad \text{for } T_1, T_2 \in \{A, B, C\} \end{aligned} \tag{S12}$$

where δ_{T_i,T_1} is the Kronecker delta function with the following definition:

$$\delta_{T_i,T_1} = \begin{cases} 1, \text{if } T_i = T_1 \\ 0, \text{otherwise} \end{cases} \tag{S13}$$

The angular bracket represents the ensemble average and f_{ij}^{exp} is the corresponding experimental contact frequency.

We applied an iterative algorithm to derive the values for $\alpha_{\text{ideal}}(|i-j|)$ and $\alpha_{\text{compt}}(T_i, T_j)$ that enforce the constraints defined in Eq. S12. We chose $n = 101$, leading to a total of 100 parameters for $\alpha_{\text{ideal}}(|i-j|)$. Together with the six parameters in $\alpha_{\text{compt}}(T_i, T_j)$, the genome model consists of 106 parameters that are all derived from Hi-C data.

Parameterizing the contact function

The function $f(r)$ defined in Eq. S10 was used to determine the chromatin contact probabilities. These contact probabilities form the basis for optimization as the maximum entropy algorithm iteratively modifies the force field parameters to Hi-C data. Accordingly, it is important to select the parameters r_c and η to be commensurate with experimental data.

A recent study conducted by Su et al. reported the spatial positions of thousands of loci from individual cells of human lung fibroblasts (IMR90) using DNA-MERFISH imaging (6). The availability of both spatial positions and Hi-C data makes possible the definition of a contact function, $f(r)$, that converts distances into contact probabilities. The raw Hi-C contacts were first converted to probabilities by normalizing with the total number of contacts. We apply a Savitzky-Golay filter to these contact probabilities to convert the noisy data to a smooth contact probability function. We re-normalize the smoothed contact probability function such that the it starts from a value of 1.0. We then calculated r_c by identifying the distance where the smoothed contact probability function approaches the value 0.5. Finally, we curve fitted the processed data to the left of r_c with the tanh contact function to define the parameter η. We determined that $r_c = 334$ nm (see figure S23) and $\eta = 0.0117$ nm^{-1}. As discussed previously we used reduced units for our simulations with the conversion $1\sigma = 385$ nm. Accordingly, the r_c and η in reduced units are 0.868 and 4.503 respectively.

SPIN state renormalized chromosome-nuclear landmark interactions

The three nuclear landmarks incorporated in this work are lamina, nucleoli and speckles. Interactions between chromatin and nuclear landmarks were described with the Lennard Jones potential.

We used the following procedure to rescale the interactions between the genome and nuclear landmarks and incorporate specificity. SPIN combines TSA-Seq, DamID data with Hi-C data to partition the chromatin into nine unique states that occupy distinct nuclear regions.(7) It further assigns each chromatin region (25 KB in size) with a probability for being in a particular state. We used these probabilities to denote the likelihood of each chromatin chromatin bead i for interacting with the three nuclear landmarks as P_i^{L}, P_i^{N} and P_i^{S}. The lamina associating probability, P_i^{L}, was determined by accumulating SPIN probabilities for states Lamina, Near_Lm1, Near_Lm2, Interior_Repr1 and Interior_Repr2 SPIN states. All these states were found with significant Lamin-B DamID scores. The nucleolus and speckle associating probabilities, P_i^{N} and P_i^{S}, were based on the Int_Repr2 and Speckle state respectively. The interaction strength ϵ in the Lennard Jones potential was then rescaled by P_i. The base value for ϵ is 0.75, 1.5, and 1.5 for lamina, nucleolus and speckle interactions respectively.

Parameter interpretation and sensitivity analysis

We employed a data-driven mechanistic modeling approach to derive parameters from experimental data. These parameters ensure the biological relevance of simulated nuclear structures. The parameters describe effective interactions among chromatin segments, and between chromatin and nuclear landmarks. For example, intra-chromosomal interactions were approximated with an ideal potential instead of explicitly building in the loop-extrusion model. Similarly, we did not explicitly include chromatin regulators that may promote the contacts among chromatin of similar type and their micro phase separation from other types of chromatin.

A significant advantage of our approach is its simplicity and efficiency. For example, explicitly modeling loop extrusion would require non-equilibrium simulations of Cohesin molecule along the DNA sequence coupled with molecular simulations of the chromosome conformational dynamics (8, 9). Additionally, to ensure the biological relevance of simulated chromosome structures, cohesin extrusion rate, concentration, dissociation rate, and the CTCF permeability, will need to be tuned in a cell type specific manner. An exhaustive search of the parameter space can be computationally demanding. Simulating compartmentalization with molecular details can be equally challenging since the exact set of protein molecules mediating chromatin-chromatin interactions is not fully known. For the known proteins that are involved in chromatin organization, parameters such as protein concentration, protein-protein interaction energy, protein-chromatin interaction energy, remain to be determined. The computational cost for explicitly modeling these additional protein molecules will be high as well. On the other hand, under the theoretical framework of maximum entropy optimization, all parameters for chromatin interactions in our model can be derived uniquely from the Hi-C data. Because they are fixed by experimental measurements, the energy function is deemed effective and accounts for both the physico-chemical interactions between chromatin segments and the contribution of non-equilibrium processes.

Furthermore, parameters in our model are indeed interpretable and consistent with proposed mechanisms of genome organization. A total of 106 tunable parameters were used to describe both intra and inter-chromosomal contacts. This number is rather minimal compared to the millions of contacts detected in Hi-C experiments. Consistent with the knowledge that heterochromatin is more condensed (10), the interactions among B compartments are indeed stronger than that for A compartments (see Table S1). Similarly, values in the ideal potential are mostly negative (see the Excel file), consistent with the formation of cross links by cohesin molecules to favor chromatin segments in close contacts.

Interaction parameters among nuclear body particles. Interactions among nuclear body particles were designed to simplify the resulting model. Parameters were tuned to produce nuclear body numbers comparable to experimental values.

For example, a cut and shifted Lennard-Jones (LJ) potential was used to model the interactions among nucleolus particles, i.e.,

$$U_{\mathrm{LJ}}(r_{ij}) = \begin{cases} 4\epsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right) - E_{\mathrm{cut}} & \text{for } r <= r_{\mathrm{cut}} \\ 0 & \text{for } r > r_{\mathrm{cut}} \end{cases}, \tag{S14}$$

with $E_{\mathrm{cut}} = 4\epsilon\left(\left(\frac{\sigma}{r_{\mathrm{cut}}}\right)^{12} - \left(\frac{\sigma}{r_{\mathrm{cut}}}\right)^{6}\right)$. The two parameters σ and ϵ in this potential were determined based on the average size of nucleoli and an average of two nucleoli per cell. In a previous study (4), we carried out systematic explorations of the parameter space and found that weakening the interactions results in the dissolution of nucleoli, while increasing the strength can cause the formation of more nucleoli.

A simple LJ potential cannot produce the desired number of speckles over a large range of parameter values (Figure S3). This observation motivated our use of a linear super-position of the LJ (Eq.S14) and the Yukawa potential (11) given by

$$U_{\mathrm{Sp}}(r_{ij}) = U_{\mathrm{LJ}}(r_{ij}) + \left(A\frac{e^{-\kappa r_{ij}}}{r_{ij}} - U_{\mathrm{cut}}^{\mathrm{Yuk}}\right), \tag{S15}$$

where $U_{\mathrm{cut}}^{\mathrm{Yuk}} = Ae^{-\kappa r_{\mathrm{cut}}^{\mathrm{Y}}}/r_{\mathrm{cut}}^{\mathrm{Y}}$. This Yukawa potential can account for long-range electrostatics that may arise due to the charges present in RNA and protein molecules. As discussed in the *Section: Non-equilibrium model for speckle formation*, it may further approximate non-equilibrium mechanisms that suppress the coarsening. The three parameters ϵ, χ, and A were tuned to produce speckle numbers on the order of 30-40 (12, 13). As shown in Figure 5 of the main text, varying the value of χ causes dramatic changes in the speckle number.

Interaction parameters between chromatin and nuclear landmarks. For simplicity, we set the parameter in the LJ potential between chromatin and nucleoli the same as that between nucleolus particles. In a previous study (4), we found that the interaction strength between nucleoli and chromatin also affects the number of nuclear bodies and stronger interactions can lead to more nucleoli. The same parameter was also adopted for the LJ potential between chromatin and speckle particles. Varying this parameter within a reasonable range does not significantly impact the agreement between simulated and experimental TSA-Seq profiles(Figure S2B).

When using the same parameter for chromatin-lamina interactions, we found a poor agreement between simulated and experimental DamID profiles. Therefore, we varied ϵ in the LJ potential to find the optimal value that produces the best agreement between simulation and experiment DamID profiles (See Figure S2).

Perturbing chromatin-nuclear landmark interactions. As mentioned in the *Section: SPIN state renormalized chromosome-nuclear landmark interactions*, we normalized the interaction strength with probabilities estimated from SPIN states. These probabilities account for the presence of multiple different types of small chromatin domains that coexist in a one MB bead. To test the role of the probabilistic assignment on reproducing chromatin-nuclear body contacts, we explored a model that adopts a binary classification for the three states. This binarization was done by converting the SPIN annotation probabilities into 0 or 1 (with 0.5 threshold) such that each chromatin bead is now either completely interacting or non-interacting with speckles, lamina, or nucleoli. As shown in Figure S1, removing the probabilistic assignment significantly deteriorated the agreement between simulated and experimental DamID profiles. However, the agreement between TSA-Seq data was relatively unaltered. The differential behavior arises since most of the speckle associating domains are comparable to the size of the coarse grained bead. However, the size of lamina associating domains is much more heterogenous and the probabilistic assignment is crucial for capturing the difference among them.

Chromatin-lamina interactions. Nuclear lamina plays important role in genome function, and mutations found in lamin proteins are frequently implicated in diseases. To study how chromatin-lamina interactions affect genome organization, we performed additional simulations that removed the interactions between the chromatin and the lamina, i.e., by $\epsilon_{\text{chr-lamina}} = 0$ (Table S2). As shown in Figure S19 A and B, the ensemble of structures from such a model completely loses its capacity to reproduce LaminB1 DamID experimental data. Additionally, the simulated structures do not agree with experimental radial positions of chromosomes (see Figure S19C). Our simulation results are consistent with several recent experimental studies that observed significant LAD dissociation from the nuclear envelope in cells with depleted lamin proteins (14–16). We note that, in addition to passive interactions, non-equilibrium mechanisms could contribute to the spatial organization of chromatin, as shown in Ref. (17). It is possible that disrupting lamin proteins alone is not sufficient to completely abolish the preferential localization of certain chromatin at the nuclear periphery (18, 19).

We predicted that the chromatin dynamics is increased in the perturbed model. We quantified the increase by calculating the mean squared displacement of the chromatin beads and computing the resulting change in diffusion coefficients (see Figure S20). Our model predicts a 63% slower diffusion coefficient for the normal cell than the depleted lamina cell. A slow down of similar magnitude was observed by Chang et al. when comparing chromatin dynamics of normal cells with the LaminB1 protein depleted cell (16).

Chromatin-speckle interactions. We ran simulations with removed interactions between chromatin and speckles/nucleoli by setting $\epsilon_{\text{chr-speckle/nucleoli}} = 0$ (Table S2). Not surprisingly, the resulting genome structures fail to produce TSA-seq profiles that match experimental values (see Figure S21). However, this perturbation does not significantly alter chromatin-lamina interactions. Together with results shown in Figure S19 A and B, these simulations support a decoupling between two nuclear landmarks, lamina and speckles. Their interactions with chromatin are relatively independent of each other.

Surprisingly, removing chromatin-speckle interactions does not impact the radial positions of chromosomes. The agreement between simulated and experimental chromosome positions is preserved even after removing the compartmentalization force, i.e., by setting α_{compt} defined in Eq. S10 to zero. Therefore, chromatin lamina interactions largely dictate the radial positions of individual chromosomes.

Chromatin speckle interactions are also key for giving rise to heterogeneity of genome organization and viscoelastic behaviors of chromatin. As shown in Figure S16, upon removing these interactions, the heterogeneity among A1-A1 contacts across trajectories almost vanished, and the system can equilibrate these contacts on simulation timescales.

We further computed the complex modulus of chromatin with and without the presence of speckle interactions, and observed a significant difference in the relaxation timescales. As shown in Figure S18, chromatin exhibits multiple timescales and behaves as a viscoelastic network. The crossover at longtimescales between $G'(t)$ and $G''(t)$ signifies the transition from gel to liquid. Upon removing chromatin speckle interactions, we found that all the four timescales τ_n decreased (see Eq. S26), supporting the role of speckles in chromatin dynamics.

To our surprise, the longest timescale does not completely disappear upon removing speckles. We further computed the shear relaxation modulus for simulations that abolished chromatin-lamina interactions. As shown in Figure S18 a significant decrease in the longest relaxation timescale is observed.

Our simulation results are consistent with the experimental observations of significantly enhanced chromatin dynamics upon depleting lamina A (20). Perturbing lamina A has been shown to both alter chromatin speckle (21) and chromatin lamina interactions. Future experiments that more specifically address chromatin-nuclear landmark interactions could further elucidate contributions of various components to chromatin dynamics.

MOLECULAR DYNAMICS SIMULATION DETAILS

Initializing simulations with randomly places chromosomes

The results presented in the main text were produced from simulations initialized with chromosome configurations obtained from a separate sampling of a genome model introduced in a previous study (22). To explore the sensitivity of these results with respect to initial configurations, we ran 12 independent simulations with configurations of the chromosomes initialized randomly. We briefly highlight the procedure we used for the randomization. The first bead of the chromosome is initialized at the origin (center of the nucleus model). This is followed by generating a vector that is uniformly distributed on the surface of a unit sphere, with the center of the sphere corresponding to the first bead. Scaling this vector with the expected bond distance between the beads (1σ) yields the Cartesian coordinates of the next bead. We repeat the procedure for all the subsequent beads within the chromosome, and the Cartesian coordinate of the subsequent bead is only accepted if the radial distance from the origin is less than 4σ (so as to randomly initialize them as globules). The procedure is similar across all the chromosomes. Lastly, we shift the chromosomes such that their center of masses are at random locations inside the nucleus.

Perturbing speckle numbers

We modified the number of speckles in the system by perturbing the interaction parameters of speckles to study the resulting effect on A compartment organization. As discussed earlier, the 36 speckle cluster system discussed in the main text is stabilised through the Yukawa repulsion potential with $\chi_0 = 0.95$. To perturb the number of speckles we ran simulations with χ equal to $0.8\chi_0$ and $0.9\chi_0$ that resulted in a 10 and 20 speckle system respectively.

Modeling the genome at the 100 KB resolution

We further carried out simulations of the 100 KB resolution model to study nuclear organization. The finer resolution results in an increase of particles by a factor of 10 and the model is computationally more costly. To avoid long relaxation that's required to reach equilibrium, we initialized simulations using configurations predicted by the 1MB resolution model. In particular, we increased the resolution of the model by linear interpolation between every two 1MB particles. The number of speckle, nucleus, and lamina particles was kept the same. We then performed a brief minimization before launching equilibrium simulations.

We adjusted the energy function for improved simulation stability and to account for the increased resolution. The bond potential in Eq. S6 was modified to a less stiff potential as

$$u_{\text{bond}}(r_{i,i+1}) = K_2(r - r_0)^2 + K_3(r - r_0)^3 + K_4(r - r_0)^4 \tag{S16}$$

with $r_0 = 0.5$ and $K_2 = K_3 = K_4 = 20.0$. The soft-core potential in Eq. S7 was changed to

$$u_{\text{sc}}(r_i) = A[1 + \cos(\frac{\pi r}{r_c})] \tag{S17}$$

with $A = 10.0$, $r_c = 0.56$. We also rescaled the pair-wise interactions $U_{\text{ideal}}(\mathbf{r})$ and $U_{\text{inter}}(\mathbf{r})$ by a factor of 0.01 to account for almost 100 times more contacts when increasing the model resolution from 1MB to 100KB. To prevent non-lamina beads from moving outside the nuclear envelope, we added a WCA potential with the form

$$U_{\text{wall}}(r) = \begin{cases} 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6] + E_{\text{cut}}, & r < r_c \\ 0, & \text{otherwise,} \end{cases} \tag{S18}$$

where E_{rm} is the energy of the LJ potential at the cutoff distance $r_c = 2^{1/6}\sigma$.

We estimated the size of 100 KB beads as 192 nm based on super-resolution imaging data, which was converted to 0.5σ in the reduced unit.

We performed 23 independent 800000-step-long simulations to compute simulated DamID and TSA-Seq profiles. As shown in Figure S7, we found that the 100kb model provides a better representation of the various chromatin domains. In particular, for DamID profiles, simulation results now better captures the heterogeneity of enrichment in lamina-associated domains, as evidenced by the more pronounced fluctuations compared to that shown in Figure 3 of the main text. To our surprise, however, the genome-wide average Pearson correlation coefficient between simulated and experimental DamID/TSA-Seq profiles is not significantly improved in the 100kb model. Further improving the agreement potentially require going beyond SPIN state based modeling strategy and would be an interesting direction for future exploration.

Non-equilibrium model for speckle formation

When modeling speckles, we introduced the Yukawa potential with a long-range repulsion between the coarse-grained speckle particles. This potential approximates the electrostatic repulsion among RNA and protein molecules and the impact of chemical modifications on proteins, both of which can suppress droplet coarsening (23). This approximation suffices for stabilizing the multidroplet state, but cannot fully capture the fusion kinetics of speckles. However, the central aim of this study was to incorporate the effects of nuclear landmarks such as speckles on chromatin and not to capture the complete physico-chemical behavior of speckles. We believe that conclusions on the nuclear organization should be insensitive to the scheme used to model the details of speckle clusters.

To examine whether the effective equilibrium model for speckles introduces any biases on our conclusions of genome organization, we tested a kinetic scheme of speckle formation introduced by de Vries and coworkers (23). In this scheme, we consider two types of speckle molecules: phosphorylated (P) and de-phosphorylated (dP). The two types can inter-convert via chemical reactions with rates ρ and σ, and only dP particles share attractive interactions with each other. For simplicity, we set $\rho = \sigma = 0.0694\ \mathrm{s}^{-1}$ which compare well with experimental values.(24) dP particles interact with chromatin segments through SPIN states as discussed previously. The *fix atom/swap semi-grand* functionality within the LAMMPS simulation package was used to swap P and dP particles on the fly, while keeping the total speckle particles constant. We increased the total number of speckle particles to 1512 to produce approximately 30-50 droplets upon reaching the steady state.

As shown in Figure S4, this model succeeds in suppressing droplet coarsening, producing an average of 45 speckles in individual simulations. The simulated TSA-seq profiles also agree well with experimental profiles (Figure S4C). Notably, the non-equilibrium model for speckles reproduced the two important features of genome organization, disordered and precise contacts. As shown in Figure S4D, there are significant fluctuations among “contacting” $A1$ segments that are bound to the same speckle across simulation trajectories. On the other hand, contacts among $A1$ evaluated at different time points of the same trajectory are relatively robust with minimal changes. These results are remarkably similar to the ones presented in Figure 6 of the main text. Therefore, we conclude that the effective model presented in the main text indeed provides a good approximation to the role of speckles on genome organization.

DETAILS ON ANALYSIS AND VALIDATION

Computing simulated contact maps

In this section, we discuss the details of calculating the simulated Hi-C maps. For two chromatin segments (say i and j) we use the contact function defined in Eq. S10 to enumerate a contact between them. The contacts between these segments are monitored over different frames and trivially ensemble averaged. Thus, one trajectory yields one Hi-C contact matrix. We run several independent trajectories (starting from distinct initial configurations) and average the contact matrices across trajectories to yield an ensemble averaged Hi-C contact matrix (for e.g. as reported in Figure 3A of the main text).

Distances from nuclear bodies and association frequencies

In this section we discuss in detail the procedure used to calculate distances and associations with the nuclear landmarks. The center of the spherical nucleus is located at the origin and hence the Cartesian coordinates of a bead i readily yieds the radial position r_i. As the lamina is located exactly at $R = 13.0\sigma$, the distance from the lamina ($d_{i,\mathrm{lam}}$) is defined as $R - r_i$.

For calculating distances from the nucleoi and speckles, we first identify clusters of the nuclear bodies. Accordingly, we use the density-based spatial clustering algorithm DBSCAN(25) as implemented in the scikit library for Python. The procedure is same for nucleoi and speckles and we describe the procedure for speckles in the following lines. The algorithm is used to return cluster labels in order to identify all the speckle particles belonging to the same cluster. This information is used to calculate the center of mass of cluster and the radius of gyration of the cluster which is treated as a proxy for the radius of the cluster. The distance of a bead i from a speckle is

$$d_{i,s} = \|\vec{r}_i - \vec{r}_s^{\,com}\| - R_s \tag{S19}$$

where the subscript s denotes the speckle, and the radius of the speckle cluster is subtracted from the distance of the bead to the cluster center of mass. Clearly, bead i will have an array of speckle distances where the length of the array is the number of speckle clusters in the nucleus. We use the minimum distance in this array for calculations shown in Figure 4 of the main text. The procedure is the same for nucleolus clusters.

For a particular frame of a simulation trajectory we define associations with lamina using a heaviside contact function. Thus, the lamina association frequency of a bead i is

$$\mathrm{LAF}_i = \frac{1}{N_t}\sum_t \Theta(d_c - d_{i,\mathrm{lam}}) \tag{S20}$$

where $d_c = 1.5\sigma$ and the subscript t and N_t account for averaging over all the frames of the trajectory. Additionally, we convert the diploid simulation data into haploid by averaging over the homologs. Thus, the model DamID value for the bead i is given by

$$\mathrm{DamID}_i = \log_2\Big(\frac{\mathrm{LAF}_i}{\overline{\mathrm{LAF}}}\Big) \tag{S21}$$

where $\overline{\mathrm{LAF}}$ is the genome-wide average of the lamina association frequencies. We calculated the genome-wide DamID correlation between model prediction and experiment as reported in Figure 3 of the main text.

The model prediction of TSA-Seq uses the exponential decay function associated with speckle distances. As described previously, a bead i has an array of distances from speckles each element denoted by $d_{i,s}$. Thus, the TSA-Seq signal for the genomic bead i is

$$\mathrm{signal}_i = \frac{1}{N_t}\sum_t\sum_s e^{-l_D d_{i,s}} \tag{S22}$$

where we sum over all the speckles and frames in the trajectory and normalize by the number of frames. The parameter l_D denotes the decay length and is set to 4.0 μm^{-1} as determined from the previously reported fit to microscopy data of TSA-Seq experiments.(26) We average the signal for homologs and the model value of TSA-Seq for bead i is

$$\mathrm{TSA_i} = \log_2\Big(\frac{\mathrm{signal_i}}{\overline{\mathrm{signal}}}\Big) \tag{S23}$$

where the denominator is the genome wide average of the TSA-Seq signals.

Shear relaxation modulus of chromatin

To characterize the viscoelastic properties of chromatin, we computed the shear relaxation modulus $G(t)$ using the Green-Kubo formula

$$G(t) = \frac{V}{k_BT}\left\langle\sigma_{\alpha,\beta}(t)\sigma_{\alpha,\beta}(0)\right\rangle, \quad \alpha \neq \beta, \tag{S24}$$

where $\sigma_{\alpha,\beta}$ is the stress tensor. V, k_B and T correspond to the volume, Boltzmann constant and the temperature of the system, respectively, and the Greek subscripts denote the Cartesian coordinates. We further averaged over different combinations for improved statistics with the following formula:

$$\begin{aligned} G(t) = {} & \frac{V}{5k_BT}\left[\left\langle\sigma_{x,y}(t)\sigma_{x,y}(0)\right\rangle + \left\langle\sigma_{x,z}(t)\sigma_{x,z}(0)\right\rangle + \left\langle\sigma_{y,z}(t)\sigma_{y,z}(0)\right\rangle\right] \\ & + \frac{V}{30k_BT}\left[\left\langle N_{x,y}(t)N_{x,y}(0)\right\rangle + \left\langle N_{x,z}(t)N_{x,z}(0)\right\rangle + \left\langle N_{y,z}(t)N_{y,z}(0)\right\rangle\right] \end{aligned} \tag{S25}$$

where $N_{\alpha,\beta}(t) = \sigma_{\alpha,\alpha} - \sigma_{\beta,\beta}$. We used 50 10-million-step-long simulation trajectories initialized from different configurations to compute the shear relaxation modulus. Configurations were saved at every 10 steps.

We used the Fourier transformation to determine the storage (G') and loss (G'') modulus from $G(t)$ as follows. We fitted $G(t)$ using the generalized Maxwell model as

$$G(t) = \sum_n G_n \exp(-\frac{t}{\tau_n}), \tag{S26}$$

where G_n and τ_n represent the amplitude and relaxation time of the n-th mode. We used $n = 4$, and larger n values do not improve goodness of fit. With the generalized Maxwell model, the storage and loss modulus were computed as

$$\begin{aligned} G' &= \sum_n G_n \frac{(\omega\tau_n)^2}{1 + (\omega\tau_n)^2} \\ G'' &= \sum_n G_n \frac{\omega\tau_n}{1 + (\omega\tau_n)^2} \end{aligned} \tag{S27}$$

Dynamical fluctuation of two different LADs.

Previous experimental studies have revealed two types of lamina-associated domains (LADs): constitutive and facultative. Constitutive LADs (cLADs) are cell-type invariant, while facultative LAD (fLADs) only interact with the NL in certain cell types. It has been speculated that cLADs collectively may form a structural “backbone”, tethering chromosomes to the NL at specific positions and thereby guiding the overall folding of interphase chromosomes. Falsifying this hypothesis has been challenging due to the lack of single-cell data until recently.

We attempted to resolve the differential behavior of cLADs and fLADs by examining the fluctuation in contact to the nuclear lamina using both simulated and imaging data. We collected the lamina distances of cLAD segments from imaging data, resulting in a $N_{cells} \times N_{cLADs}$ matrix. N_{cells} denotes the number of cells studied in the experiment, and N_{cLADs} denotes the number of cLAD segments. We calculated the standard deviation in the cLAD-lamina distances (i.e., fluctuations between cells) by computing the standard deviation over the N_{cells} axis. The resulting data is one-dimensional with length N_{cLADs}. We histogram this data, and cLADs data is shown by green distribution in the upper panel of Figure S10B. The procedure is repeated on the imaging data set to isolate the fluctuation in fLADs. The mean value of such distributions shown in Figure S10B yields the average size of LADs fluctuation, while the standard deviation of the distribution is a measure of the variability between the fluctuations of different LAD segments (fluctuation in fluctuations!). We repeat the process for the simulation data.

The mean value of the fluctuations for cLADs and fLADs computed from imaging data is $0.34 \pm 0.01\mu m$ and $0.35 \pm 0.02\mu m$, respectively. They agree with the mean fluctuation value of $0.55\mu m$ and $0.63\mu m$ predicted from simulations. Therefore, both data support that cLADs contacts with lamina are more stable, while fLADs can be more dynamic with larger fluctuations.

We note that a perfect agreement between simulated and experimental data is difficult to achieve at this point. Part of the discrepancy could be due to the approximate nature of our model. However, two additional aspects inherent to the experimental setup and data collection can also lead to the observed deviations. For example, the experimental nuclei adopt ellipsoidal shapes, with a height of 1.6 μm in the z axis compared to the 7.0 μm and 12.7 μm in the x and y axis respectively. Unlike the spherical nucleus used in simulations, this uneven shape will force most chromatin to be closer to the nuclear lamina. Furthermore, Zhuang et al. did not directly probe the nuclear lamina but approximated the boundary with a convex hull of the imaged chromatin data. This ignores the width of the nuclear envelope, leading to an underestimation of chromatin-lamina distances.

Subcompartment analysis of simulated contact maps.

Although the clustering method we used is the same as done by Rao et al., we discuss the process briefly for clarity.(27) From the simulated genome-wide contact map, we generate a subset contact matrix. In this contact matrix, the genomic segments on the odd-chromosomes appear as rows and the genomic segments on the even-chromosomes are the columns of the matrix. Chromosome X was left out from this analysis. Accordingly, to cluster A compartments on a particular odd chromosome we isolate complete rows that correspond to loci belonging to the chromosome and defined as type A in our model. The process is similar for even-chromosomes except that the subset contact matrix is transposed in this case before isolating A loci. The previous filtering yields a matrix, that is converted into z-scores by passing the rows through Python’s *scipy* z-score function. This is used as the input for a Gaussian hidden Markov model (GaussianHMM) to cluster the A compartment z-scores data into two clusters. Clearly, the clustering procedure returns one of two labels (say 0 and 1) for all the A loci on a particular chromosome without conveying if 0 is related to $A1$ or $A2$. Thus, the procedure returns an array of labels (0s and 1s) for the A loci on a chromosome. To compare with the experimental annotations we use the higher of $A1/A2$ experimental probabilities on a 1 MB A loci to classify the loci as either $A1/A2$ as inferred from the experiment. We set the experimental $A1$ as 0 and $A2$ as 1 to yield an array with 0s and 1s for A loci. This labels array from the experiment is compared to the labels array from the clustering and its binary complement (as label 0 in clustering may not relate to $A1$). The element-wise match between the two label arrays divided by the total number of elements is defined as the overlap between clustering and experimental annotations. The maximum value of the overlap is used to identify subcompartments in clustering and reported in this study.

α_{AA}	-1.441575
α_{AB}	-0.764335
α_{AC}	-0.690723
α_{BB}	-1.507070
α_{BC}	-0.939463
α_{CC}	-0.246267

Table S1: Summary of compartment-compartment interaction parameters, i.e. α_{compt} defined in Eq. S11, used in the study.

$\epsilon_{\text{chr-lamina}}$	0.75
$\epsilon_{\text{chr-nucleolus}}$	1.5
$\epsilon_{\text{chr-speckles}}$	1.5
$\sigma_{\text{chr-lamina}}$	1.0
$\sigma_{\text{chr-nucleolus}}$	0.75
$\sigma_{\text{chr-speckles}}$	0.75

Table S2: Summary of chromatin-nuclear landmark Lennard Jones interaction parameters, i.e. ϵ and σ, used in the study.

Potentials	Functional forms	Parameter values
Bonding potential	$u_{\text{bond}}(r_{i,i+1})$ in Eq. S6	Standard values in coarse-grained polymer models
Angular Potential	$u_{\text{angle}}(\vec{r}_{i,i+1}, \vec{r}_{i+1,i+2})$ in Eq. S6	Standard values in coarse-grained polymer models
Soft-core potential	$u_{\text{sc}}(r_{ij})$ in Eq. S8	Standard values in coarse-grained polymer models
Ideal potential	$U_{\text{ideal}}(\mathbf{r})$ in Eq. S9	Values for α_{ideal} were obtained from optimizations against Hi-C data (see Figure S24).
Compartment potential	$U_{\text{inter}}(\mathbf{r})$ in Eq. S11	Values for α_{compt} were obtained from optimizations against Hi-C data (see Table S1).

Table S3: Summary of the potentials applied to coarse-grained beads representing the genome at the one MB resolution. The bonded energy terms, including the bonding and angular potentials, were applied to consecutive beads on the same chromosome. The ideal potential was also limited to bead pairs from the same chromosome. The non-bonded soft-core potential, and compartment potential were applied to bead pairs both from the same and from different chromosomes. See also Section: Hi-C inspired interactions for the diploid human genome for details on the definition and parameterization of the potential energies.

Potentials	Function forms	Parameter values
Nucleolus-Nucleolus	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = 2.0, \sigma = 0.5, r_{\text{cut}} = 2.0$, and their values were chosen based on the average size of nucleoli and an average of two nucleoli per cell
Speckle-Speckle	$U_{\text{Sp}}(r_{ij})$ in Eq. S15	$\epsilon = 8.0, \chi = 0.95, \kappa = 1/\chi, A = 2.5\chi$ were chosen to reproduce speckle numbers observed in microscopic images. $r_{\text{cut}} = 2.0,\ r^{\text{Y}}_{\text{cut}} = 6.0\chi,\ \sigma = 0.5$
Nucleolus-Speckle	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = 1.0, \sigma = 1.0$ were chosen as standard values to provide excluded volume effect. $r_c = 2^{1/6}\sigma$
Nucleolus-Lamina	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = 1.0, \sigma = 1.0$ were chosen as standard values to provide excluded volume effect. $r_c = 2^{1/6}\sigma$
Speckle-Lamina	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = 1.0, \sigma = 1.0$ were chosen as standard values to provide excluded volume effect. $r_c = 2^{1/6}\sigma$

Table S4: Summary of the interaction potential among particles that make up the nuclear landmarks. As their names indicate, the Nucleoli-Nucleoli potential was applied to pairs of beads identified as nucleolar particles, and the Speckle-Speckle potential was applied to pairs of beads identified as Speckle particles. The remaining three potentials were applied to particles from two different nuclear landmarks as specified in their respective names.

Potentials	Functional forms	Parameter values
Chromatin-Nucleolus	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = P^{\text{N}}_i\epsilon_0, \sigma = 0.75$, with $\epsilon_0 = 1.5$. $r_{\text{cut}} = 4.0$.
Chromatin-Speckle	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	$\epsilon = P^{\text{S}}_i\epsilon_0, \sigma = 0.75$, with $\epsilon_0 = 1.5$. $r_{\text{cut}} = 4.0$.
Chromatin-Lamina	$U_{\text{LJ}}(r_{ij})$ in Eq. S14	For non-zero P^{L}_i values, $\epsilon = P^{\text{L}}_i\epsilon_0, \sigma = 1.0$, with $\epsilon_0 = 0.75$. $r_{\text{cut}} = 4.0$. If $P^{\text{L}}_i \equiv 0$, we have $\epsilon = 0.75$ and $r_c = 2^{1/6}\sigma$ to provide excluded volume effects that confine chromatin inside the nucleus.

Table S5: Summary of the non-bonded interactions between genomic beads at one MB resolution and various nuclear landmark particles. P^{N}_i, P^{S}_i, and P^{L}_i represents the probability for a chromatin bead to be in the nucleolus, speckle, and lamina state as assigned by the SPIN algorithm. As mentioned in the main text, SPIN annotates chromatin based on their relative position with respect to various nuclear structures through an integrative analysis of SON TSA-Seq, Lamin B DamID, and Hi-C data (7).

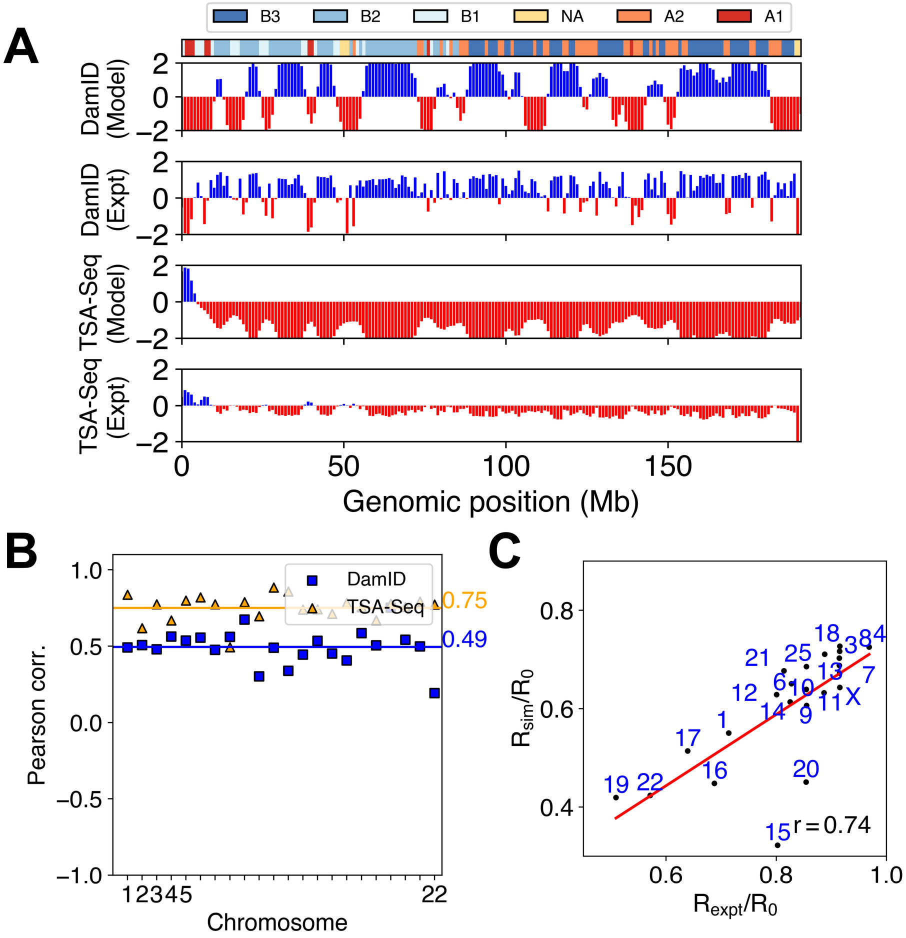


Figure S1: **Probabilistic assignment of chromatin domains helps to account for the heterogeneity of chromatin content.** See text *Perturbing chromatin-nuclear landmark interactions* for additional discussion. (A) Comparison (for chromosome 4) between the model predicted Lamin-B DamID and SON TSA-Seq signals and the experimental data at 1MB resolution. (B) Pearson correlation coefficients between simulated and experimental Lamin-B DamID (blue) and SON TSA-Seq (yellow) data for individual chromosomes. The genome-wide averages are shown as straight lines with the corresponding values on the side. (C) Comparison of the chromosome radial positions in experiment (28) and simulations. Here R_0 is the radius of the nucleus and $R_0 = 13\sigma = 5\mu$m.

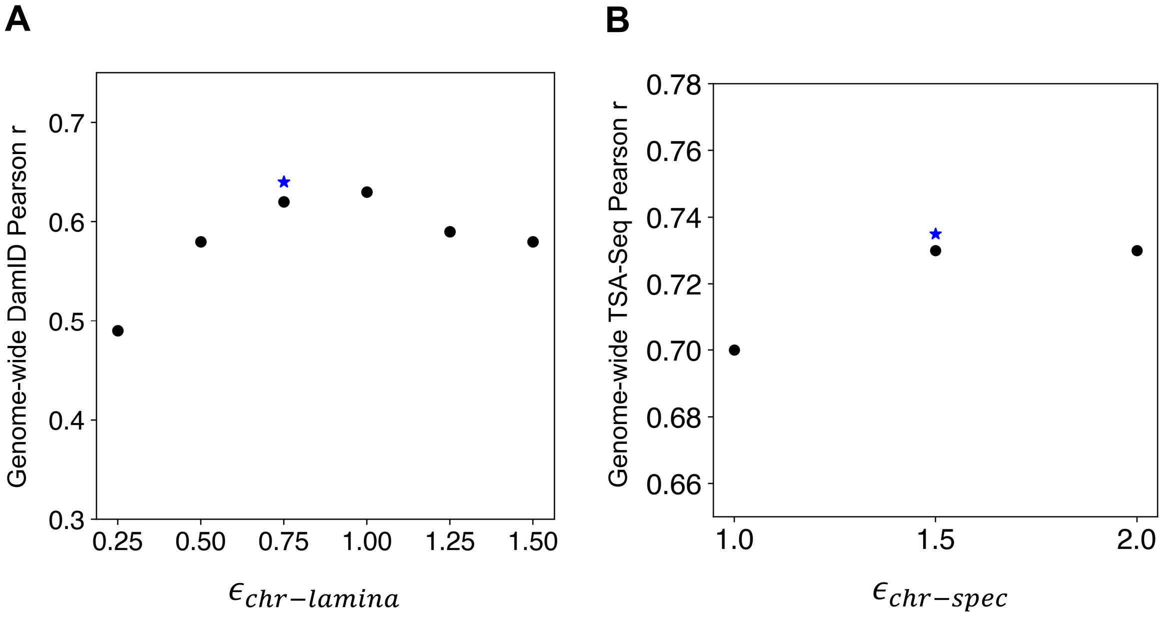
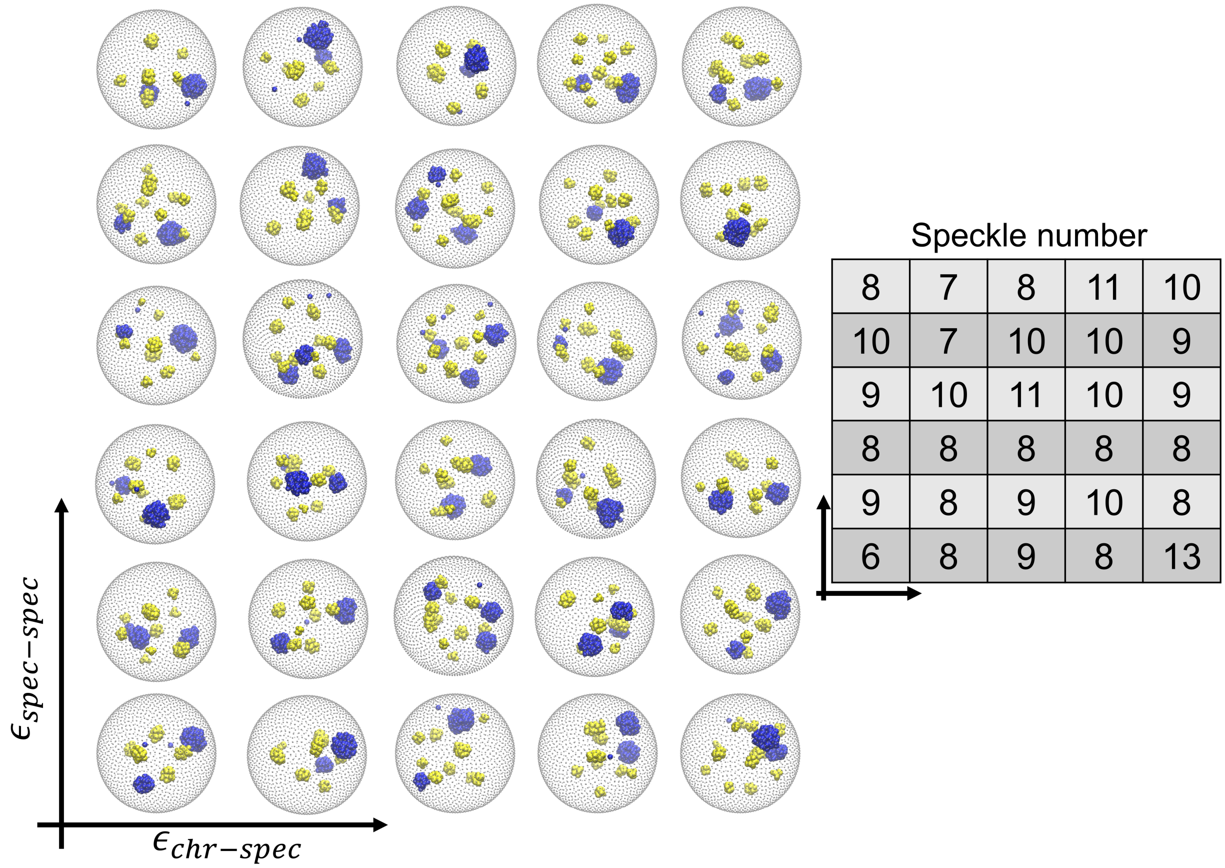


Figure S2: **Varying the interaction parameters in the model for optimal agreement with experimental measurements.** (A) Variation in the genome wide Pearson r correlation between model and experimental DamID profiles on changing the interaction strength of the chromatin-lamina interaction. We clearly observe an optimum in the correlation and the ⋆ symbol denotes the value we used. (B) A similar calculation for the agreement with TSA-Seq data on changing the interaction strength of the chromatin-speckle interaction.



Speckle number

8	7	8	11	10
10	7	10	10	9
9	10	11	10	9
8	8	8	8	8
9	8	9	10	8
6	8	9	8	13

Figure S3: **Dependence of speckle numbers on a model with only attractive interactions.** As discussed in detail in the main text, we observed that a simple model with attractive interactions between speckles (Eq. 2 of main text) is insufficient to yield the experimentally observed number of speckles (30-40). This figure shows the nuceloi (blue) and speckles (yellow) in the last frame of 30 simulations where we varied the epsilon of chromatin-speckle (3.5-5.5 with 0.5 increments) and speckle-speckle interaction (1.2-2.2 with 0.2 increments). Each trajectory is 6 million timesteps. The table summarizes the number of speckle clusters identified using the DBSCAN algorithm.

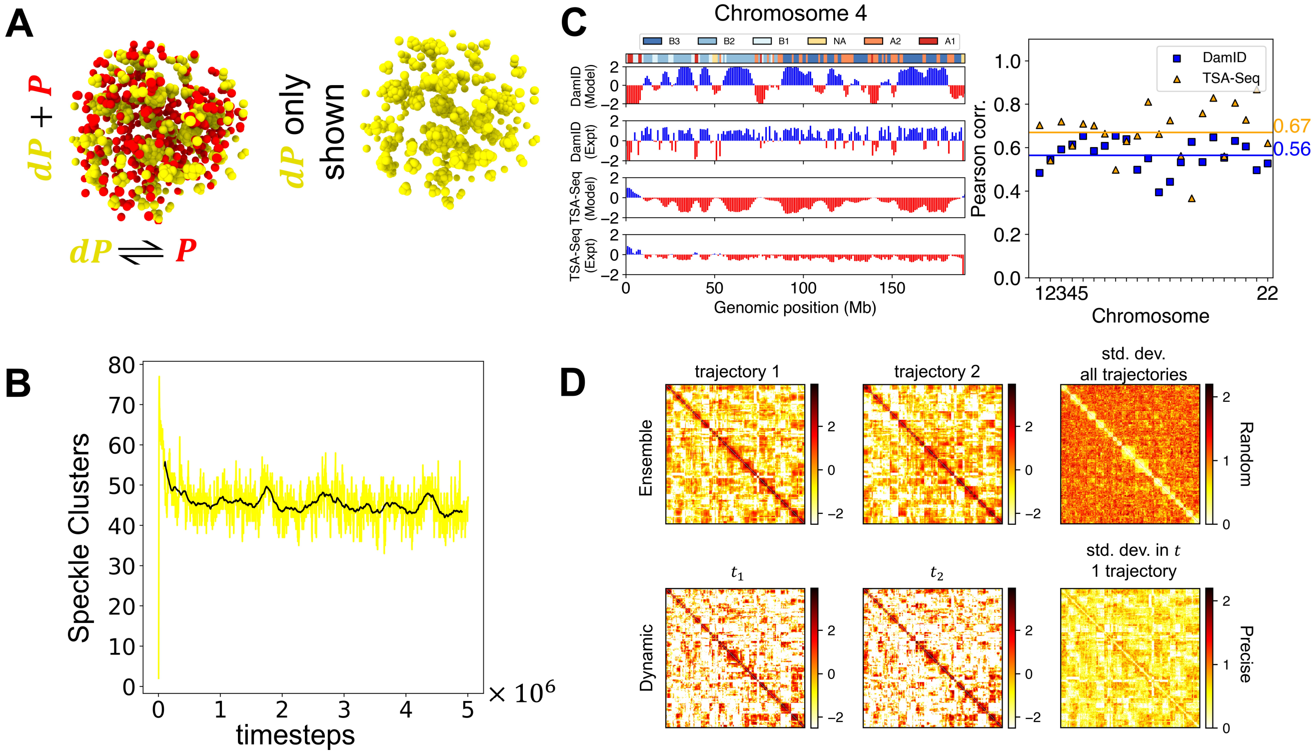


Figure S4: **Non-equilibrium speckle model.** (A) We developed an alternate model where we used semi-grand Monte Carlo to simulate chemical modifications within speckle particles. We ran simulations with two types of speckle particles (dephosphorylated and phosphorylated). The dP particles self assemble into speckle clusters while the constant exchange between dP and P particles stabilizes the individual clusters. (B) The number of speckle clusters as a function of the simulation timesteps. The dynamic fluctuations in the speckle number are achieved due to our kinetic scheme. (C) The sequencing predictions for Chromosome 4 are shown which are in good agreement with experimental data. Additionally the model predictions are similar to the one shown in Figure 3b of the main text. The genome-wide Pearson r correlation coefficient for agreement with sequencing experiments highlights the robustness of the model to the choice of speckle model parameters. (D) Contact maps were computed similarly as those presented in Figure 6 of the main text, but using genome structures from simulations of a non-equilibrium model for speckles. They demonstrate that incorporating a much more complex treatment of the speckle-speckle and speckle-chromatin interaction yields similar results as discussed in the main text.

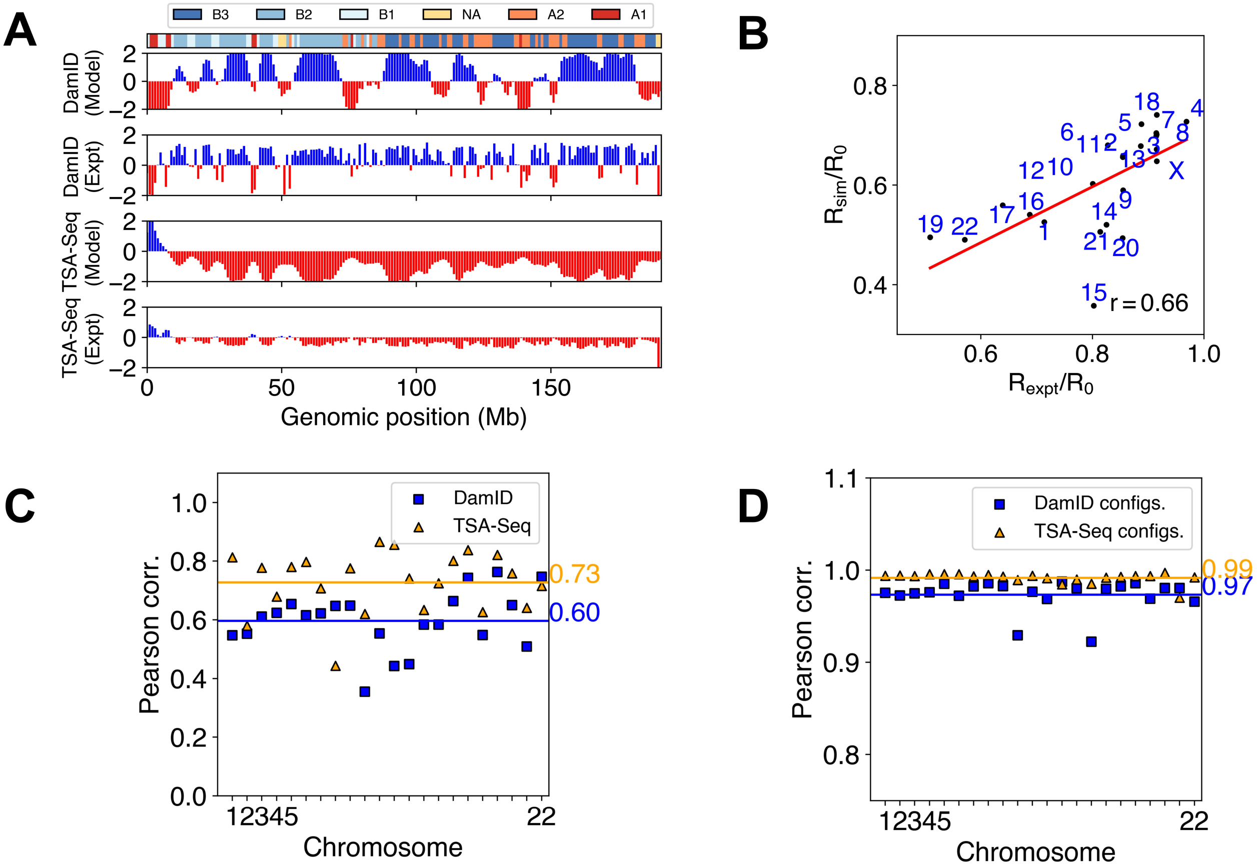


Figure S5: **Converged configurations are independent of initial conditions.** See text *Section: Initializing simulations with randomly places chromosomes* for simulation details. (A) The agreement with sequencing experiments (for chromosome 4) are shown for simulation started with random initial configurations. (B) The correlation between the simulated and experimental chromosome positions with random initial configurations is very similar to that obtained in Figure 4 of the main text. (C) The genome-wide agreement with sequencing experiments. (D) We calculated the Pearson r correlation between the DamID and TSA-Seq profiles for simulations with completely random configurations and with configurations initialised from a previous study.(22) Clearly, our simulations are well converged and our analysis is insensitive to the choice of the initial configurations.

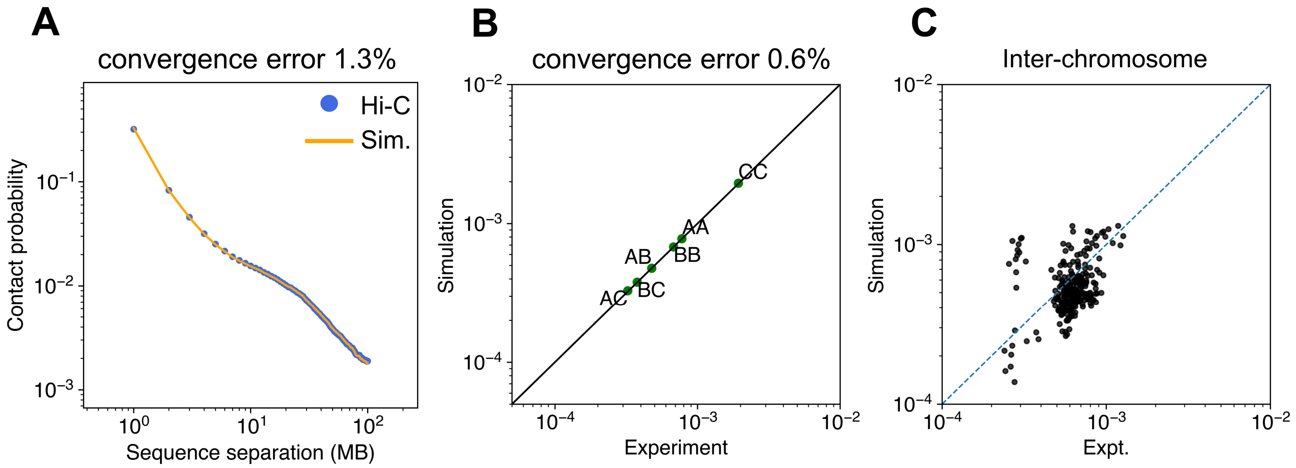


Figure S6: **Additional comparison of the model with Hi-C data.** (A) Power-law decay of the intra-chromosomal contact probability. The intra-chromosome contact probabilities in the model are optimized using the maximum entropy algorithm such that the model probabilities converge to the Hi-C contact probabilities. (B) The compartment-compartment interaction parameters in the model are optimized to match experimental contact probabilities and the resulting match is highlighted by the parity plot. (C) The model does not explicitly incorporate chromosome-specific interactions but yields a reasonable comparison with experimental inter-chromosome contact probabilities.

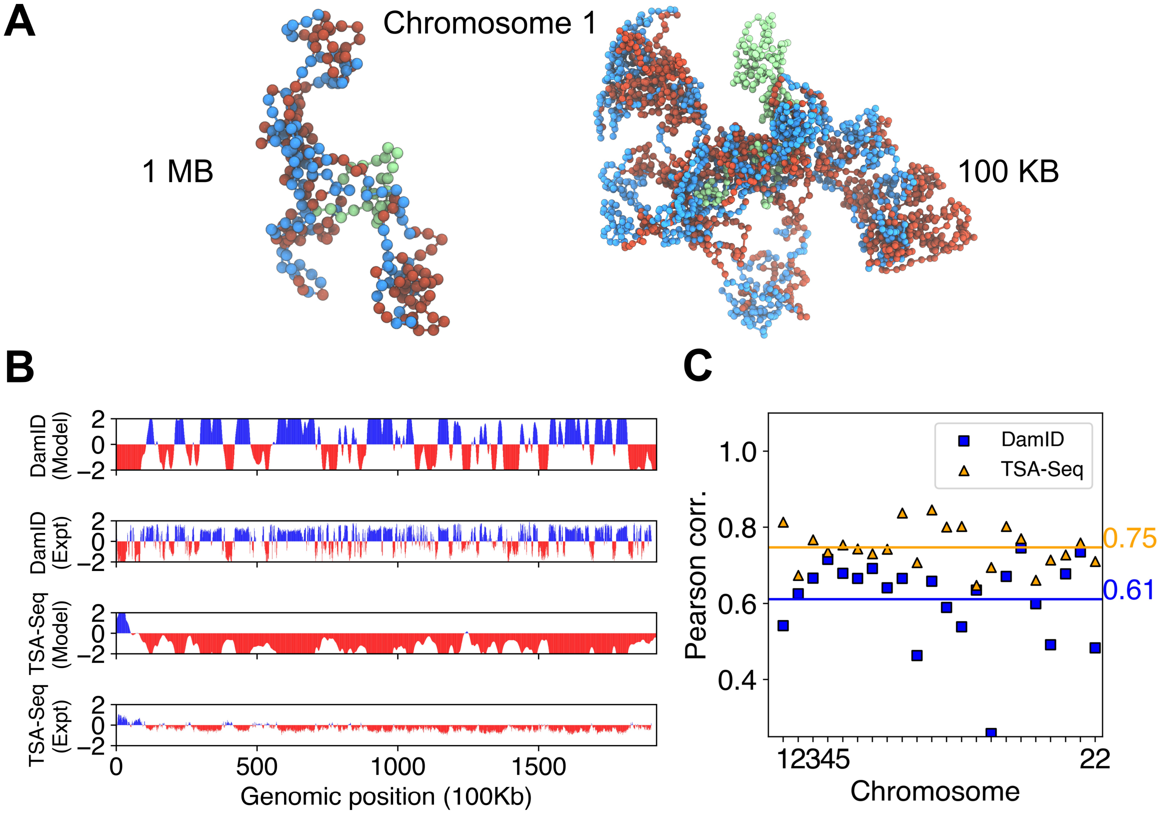


Figure S7: **Simulation results at the 100 kb resolution.** See text *Section: Modeling the genome at the 100 KB resolution* for simulation details and additional discussion. (A) Illustration of the increase in resolution of the model shown for Chromosome 1. (B) DamID and TSA-Seq profiles from the model compared against experiments at the 100kb resolution. (C) We observe a good genome wide agreement between model predictions and experimental sequencing profiles.

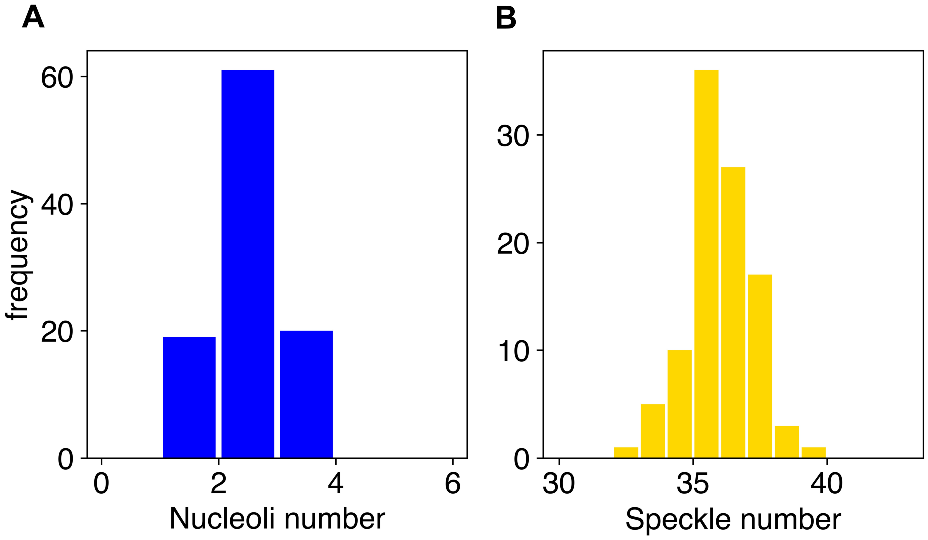


Figure S8: **Distribution of the number of nucleoli and speckles**. The average number (calculated from 100 independent trajectories) of nucleoli in our model is 2 and the average number of speckles in our model is 36.

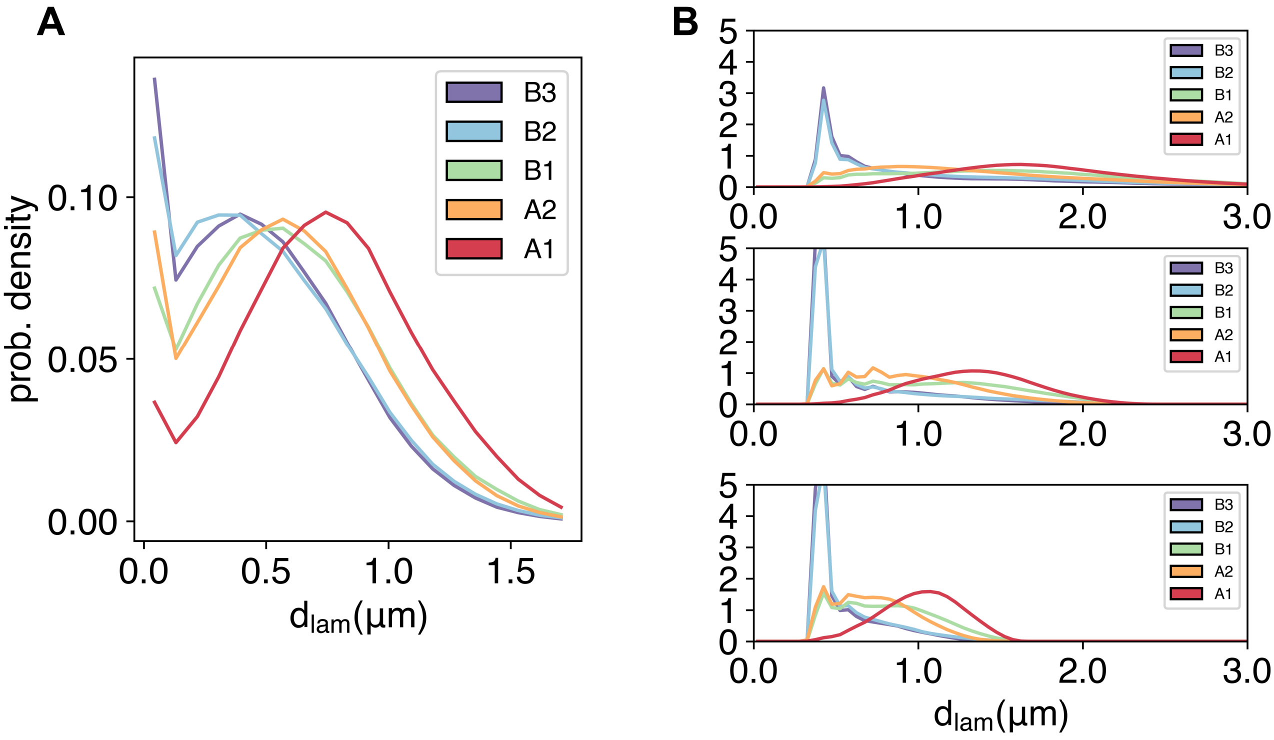


Figure S9: **Comparison between simulated and experimental distribution of genome subcompartments as a function of distance from nuclear lamina ($d_{\mathbf{lam}}$).**. (A) Sensitivity of the experimental probability distribution at small distances with respect to the definition of nuclear envelope. The nuclear lamina was not explicitly imaged in Ref. 6. Instead, Su et. al used a convex hull algorithm to identify the outermost probes in their imaging assays as lamina. The important distinction here is that chromatin probes in their study that are exactly at a $d_{lam} = 0$ are defined as lamina and left out from subsequent analysis. However, including all the points that form the outermost portion of their convex hull reveals a peak in the distribution of B compartments at the periphery. (B) Sensitivity of the simulated probability distribution with respect to the shape of nuclear envelope. We carried out additional simulations with ellipsoidal shapes be setting the three semi-axes of the ellipsoid set as (10, 5, 2.5) μm and (15, 5, 1.67) μm, respectively. The corresponding results are shown in the middle and bottom panel, Compared to those shown in the main text (top panel), the range of values for the distance to lamina from various genomic regions gradually decrease and eventually reach values comparable to that seen in the experiment.

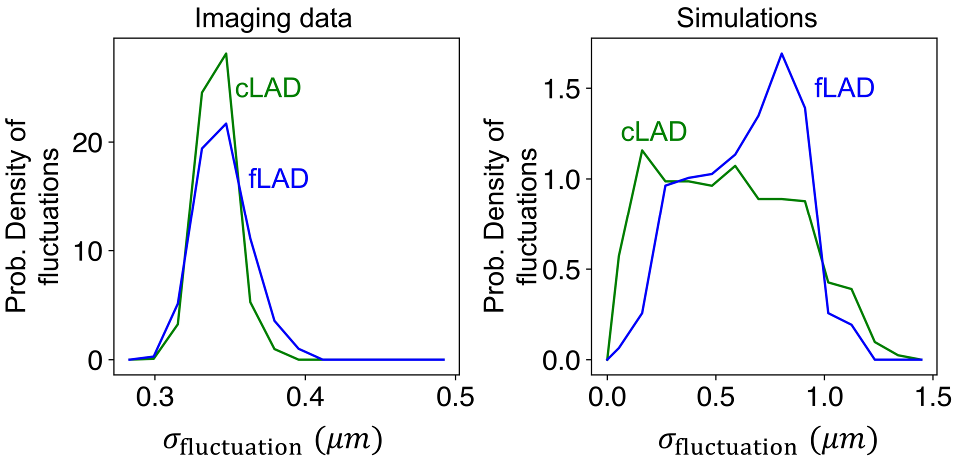


Figure S10: **The distribution of fluctuations in cLADs and fLADs in the imaging and simulation data.** $\sigma_{\text{fluctuation}}$ quantifies the variance of the distance to lamina among the two types of lamina-associated domains (LADs). The probability density plots were estimated using different cells from imaging (left) and different simulation trajectories (right). Consistent with the imaging data, our model predictions suggest that fLADs have a slightly higher fluctuation in distances from lamina compared to cLADs. See text *Section:Dynamical fluctuation of two different LADs* for additional discussion.

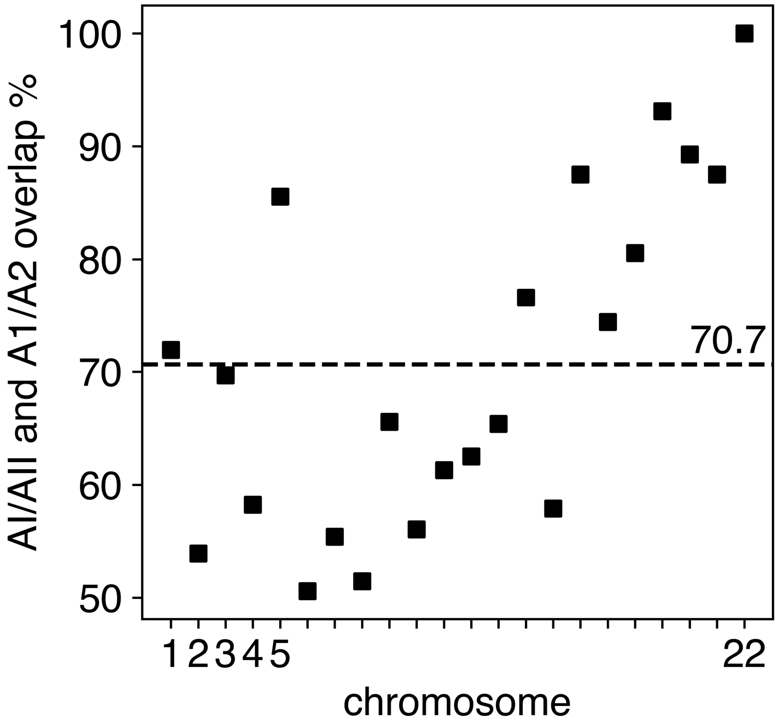


Figure S11: **Two A compartment sub-types obtained from clustering simulated genome-wide contact matrices compared with the experimental sub-compartment annotations.** The A subcompartments obtained from clustering the contact matrix are denoted as AI and AII (see Fig 5 of main text) and are in good agreement with the $A1$ and $A2$ inferred from subcompartment annotations. The genome-wide average overlap is 70.7%. See text *Section: Sub-compartment analysis of simulated contact maps* for more details on extracting AI and AII from simulated contact maps.

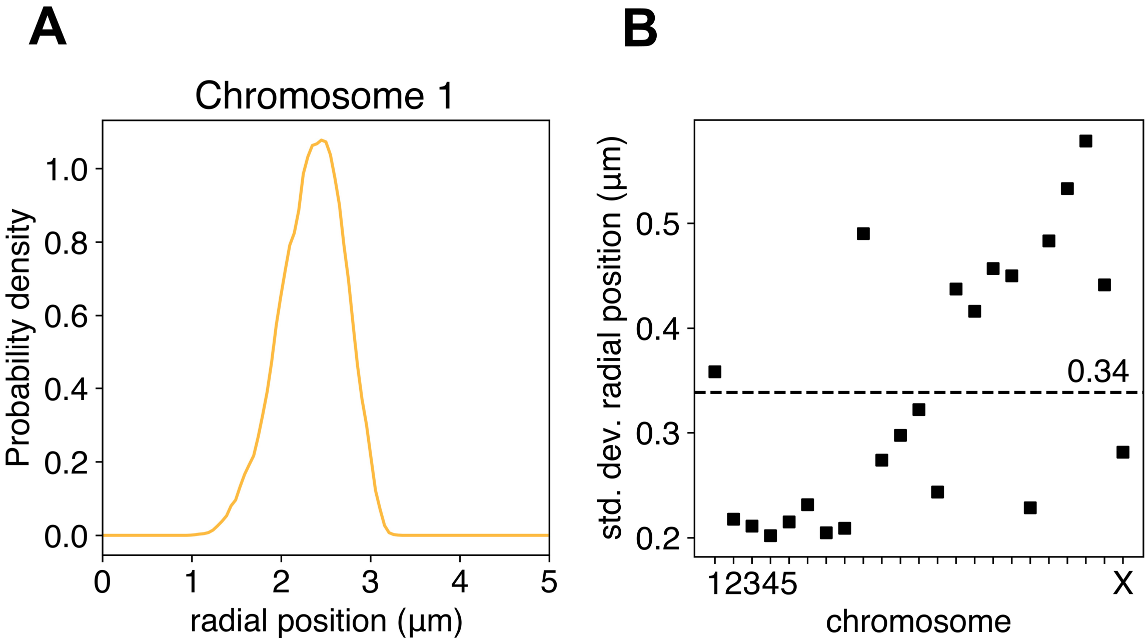


Figure S12: **Fluctuation in simulated chromosome positions.** (A) The distribution of chromosome 1 radial positions from simulations. (B) Standard deviations in radial position i.e. σ_r (in μm) of all chromosomes has a genome-wide average value of 0.34 μm. It is to be noted that on approximating the total spread of the distribution as $4\sigma_r$-$6\sigma_r$, the heterogeneity in radial positions 1.36-2.04 μm is fairly large.

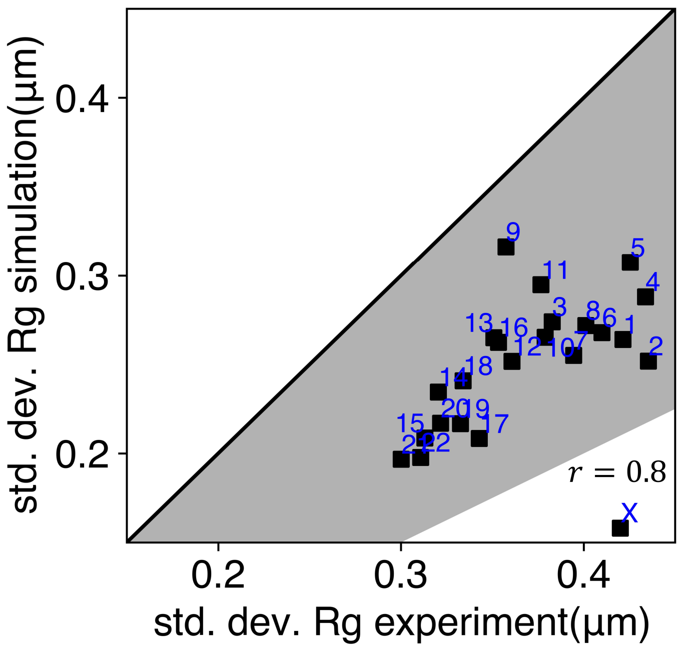


Figure S13: **Fluctuation of chromosome radius of gyration.** The standard deviations in radius of gyration distributions calculated from DNA-MERFISH experiments and our simulations is shown. We report the correlation coefficient on ignoring two outlier chromosomes (9 and X). The HFF cell is a male cell-type and we did not explicitly model the X and Y chromosomes differently which could be one of the reasons for chromosome X being a clear outlier. The black solid line is the parity line. Clearly, unlike the mean R_g (see Fig 4a of main text) the fluctuations in the simulations are not in perfect agreement with experiments. However, for almost all chromosomes the fluctuations are within a factor of two of the experiments as highlighted by the grey shaded region.

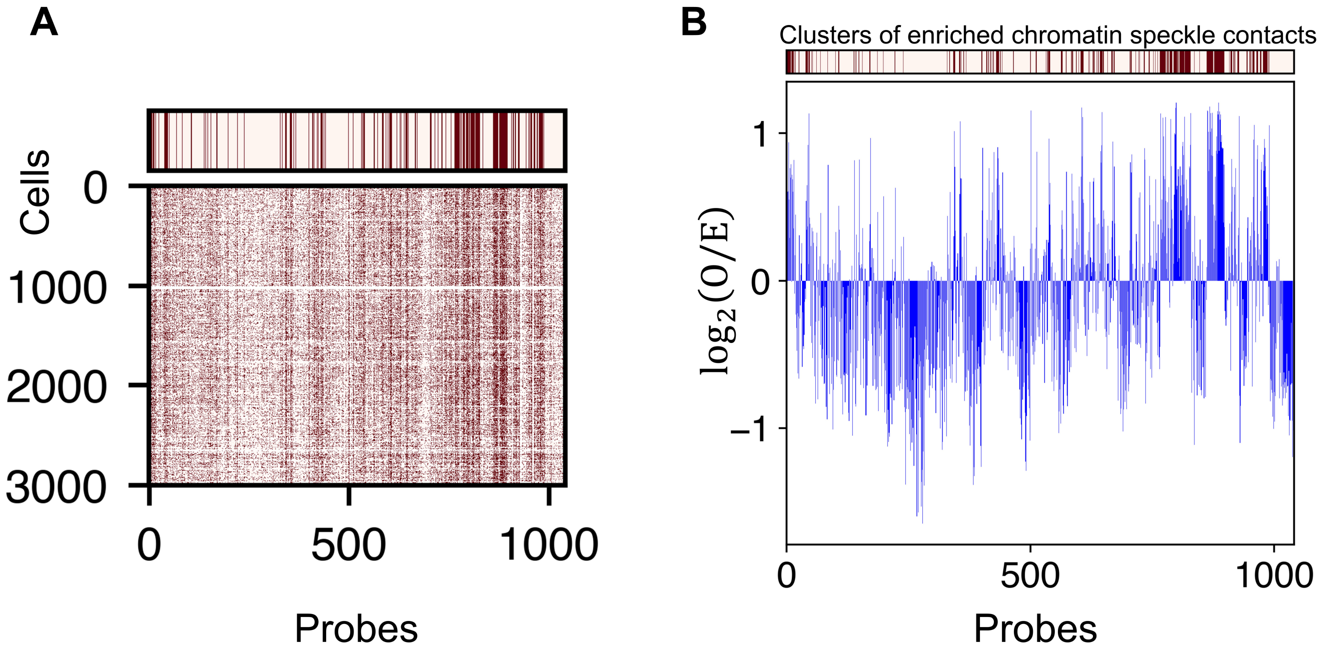


Figure S14: **Clustering chromatin-speckle contacts in the imaging data.** (A) Binary contact matrix for the imaging data with a speckle distance of probe less than $0.5\mu m$ defined as a contact. Clearly, some probes on the chromatin segments are more often in contact with speckles across different cells. We used K-means clustering to identify chromatin segments that are frequently in contact with the speckles. The clustered annotations are shown above the matrix. (B) The identified clusters line up well when compared against a segment-wise enrichment plot for speckle contacts. We use the cluster labels that are enriched in speckle contacts to compute the A1-A1 contact matrix discussed in Figure S15.

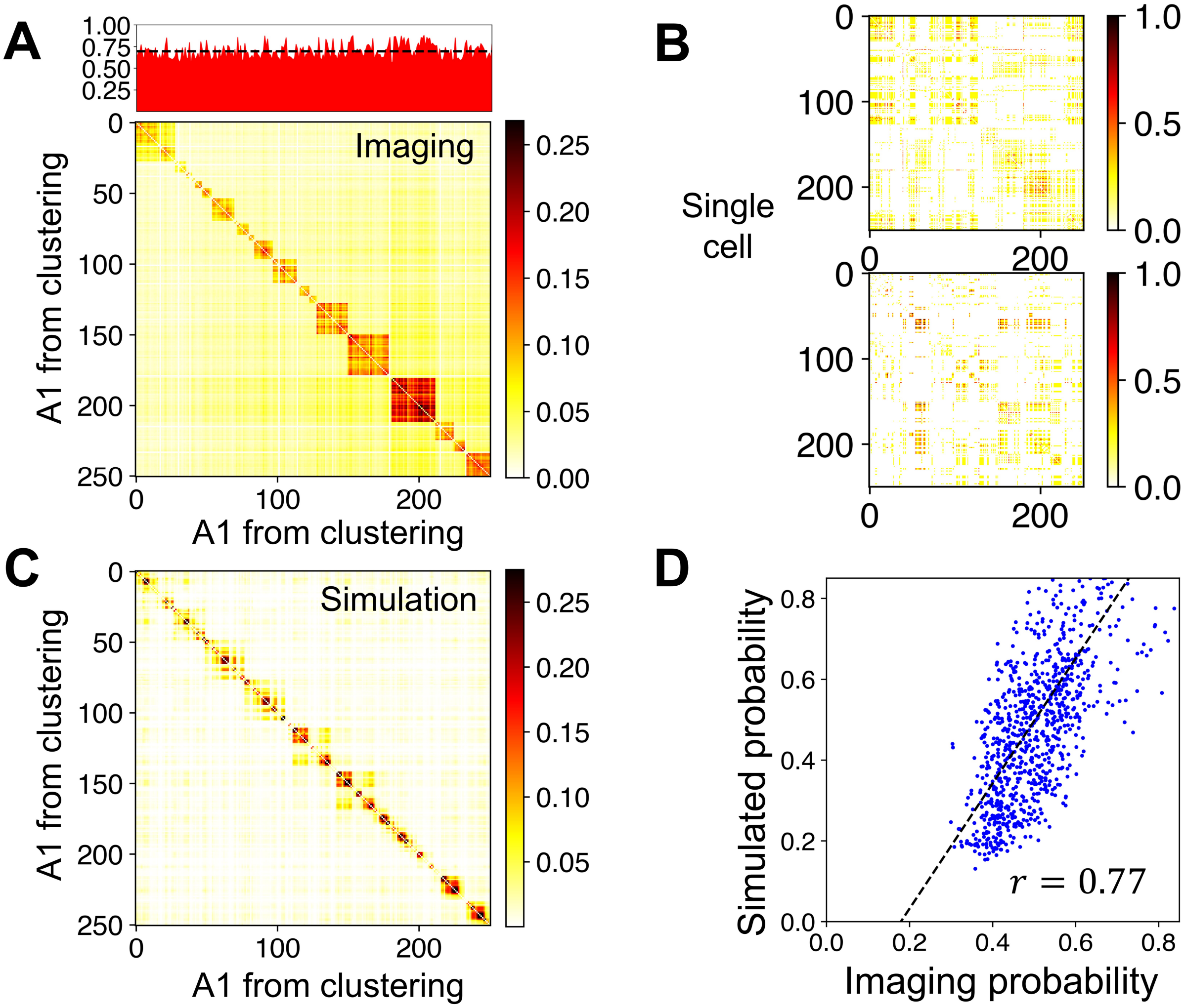


Figure S15: **Coexistence of heterogeneity and precision observed in imaging data** (A) The average A1-A1 contact matrix determined using DNA MERFISH imaging data. Contacts were defined only for pairs of chromatin segments bound to the same speckle, with the average probabilities across different cells ranges from 0-0.25. However, when we calculate the probability for the A1 segments to be in contact with speckles (see panel above the matrix) the contact probability is much higher as the segments are consistently in contact (with a mean value of 0.70) with speckles across different cells. (B) The contact matrix for individual single cells highlights the variability in contacts. (C) The average A1-A1 contact matrix from simulations for A1 segments identified using clustering strategy discussed in S14. We observed that this matrix is in good agreement with the matrix in Figure S15A with a Pearson correlation coefficient of 0.39. The less than perfect agreement between simulation and experimental results could arise from the difference in cell types, nucleus shape, and speckle numbers. Despite the complexity of cell nuclei, we found it encouraging that the average A1-A1 contact probabilities for the off diagonals (inter-chromosomal) is within a factor of 3.5 of the imaging data set. Therefore, the simulated genome ensemble provides a reliable representation of the structural heterogeneity of cell nuclei. (D) We compare the contact probability of all the chromatin segments (with speckles) in the DNA-MERFISH study with the chromatin-speckle contacts observed in our simulations. The high degree of correlation between the model and the experiment further validates the predictive capacity of the model.

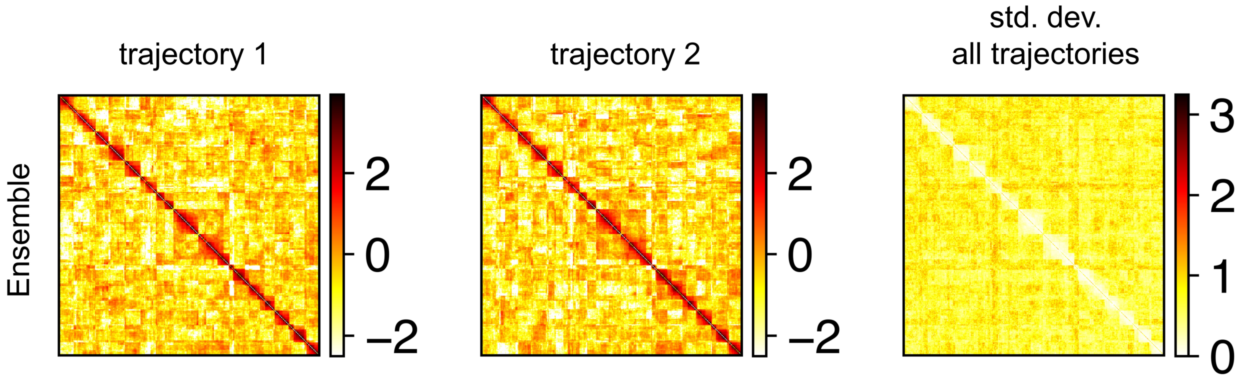


Figure S16: **Loss in heterogeneity between A1-A1 contacts upon removal of chromatin-speckle interactions**. See text *Perturbing chromatin-nuclear landmark interactions* for additional discussion. Contacts maps obtained from two independent trajectories (from different initial configurations) are shown in the left and middle panel. The standard deviation of contact maps obtained from 10 trajectories is shown in the right panel. Compared to Figure 6 of the main text, the difference in A1-A1 contact matrix (captured by the std. dev. matrix) is now much smaller.

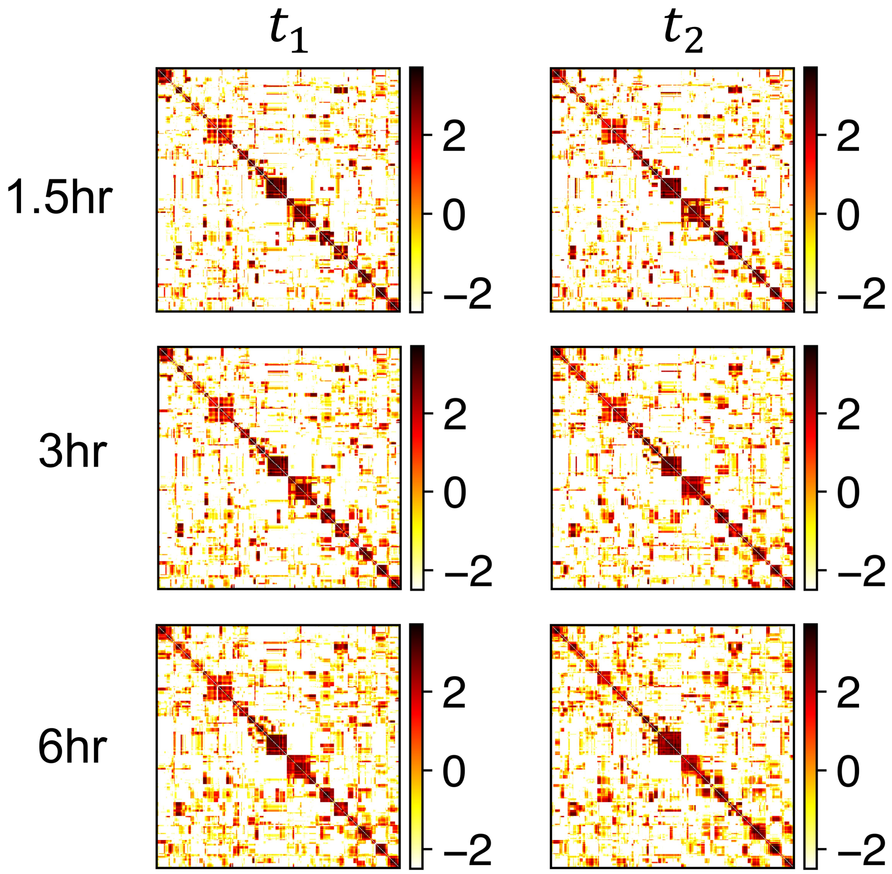
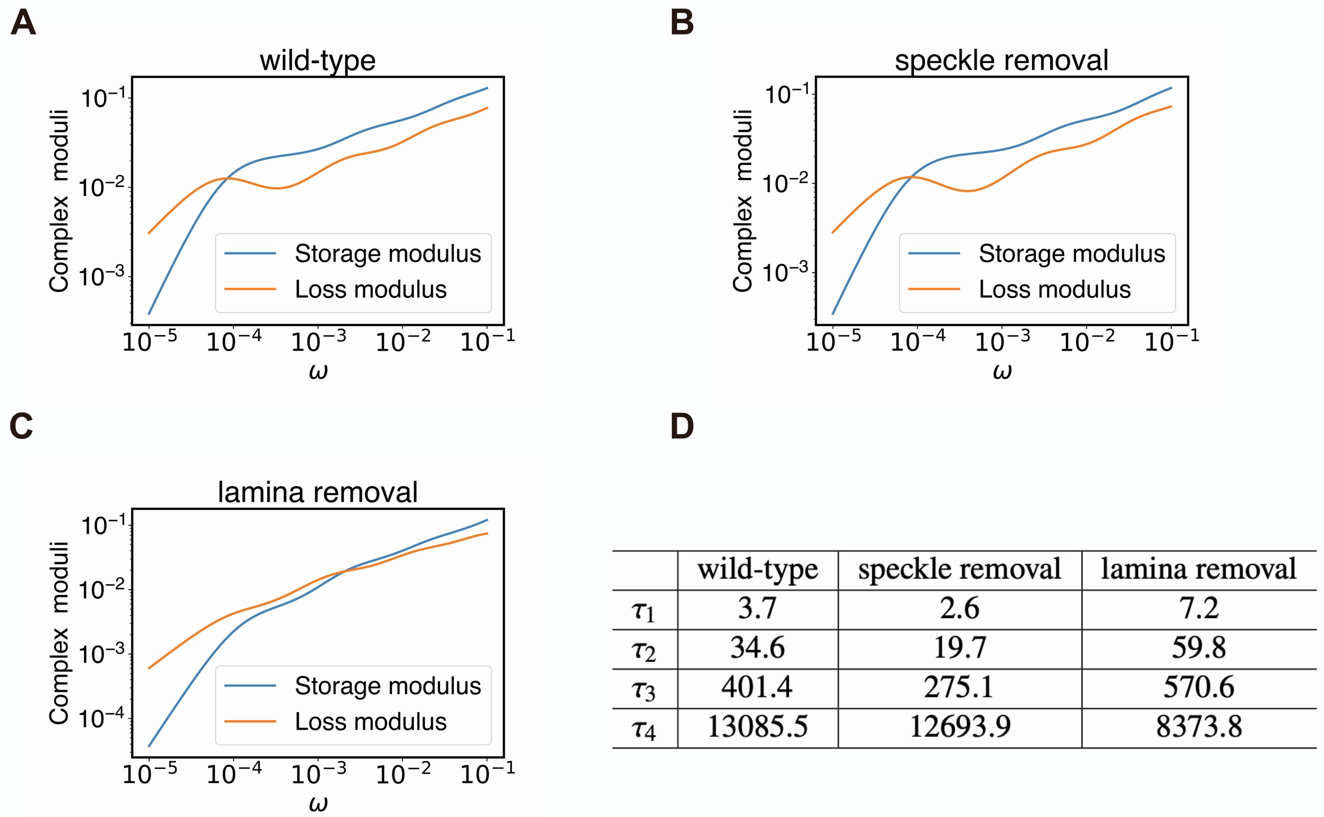


Figure S17: **A1-A1 contact matrix within a single trajectory.** The contacts between chromatin and speckles once formed have stable lifetimes across the time span of a single trajectory as observed in the structures generated from our model. We used the middle row for our results in Figure 6 of the main text. However, clearly the contact matrix is insensitive to the time difference over which the averages are computed because of the stability of the contacts within a trajectory.



	wild-type	speckle removal	lamina removal
τ_1	3.7	2.6	7.2
τ_2	34.6	19.7	59.8
τ_3	401.4	275.1	570.6
τ_4	13085.5	12693.9	8373.8

Figure S18: **Complex modulus for the viscoelastic nuclear network.** See text *Section: Shear relaxation modulus of chromatin* for calculation details and *Section: Perturbing chromatin-nuclear landmark interactions: Chromatin-speckle interactions* for additional discussion. (A) Complex moduli for the 1 MB model developed in this study (wild-type). (B) Moduli for the system with abolished chromatin-speckle interactions. (C) Moduli for system with abolished chromatin-lamina interactions. (D) Timescale τ_n obtained by fitting Eq. S26 to $G(t)$ computed using simulations presented in the main text (wild type), simulations that removed chromatin-speckle interactions (speckle removal), and simulations removed chromatin-lamina interactions (lamina removal). The timescales are reported in reduced units.

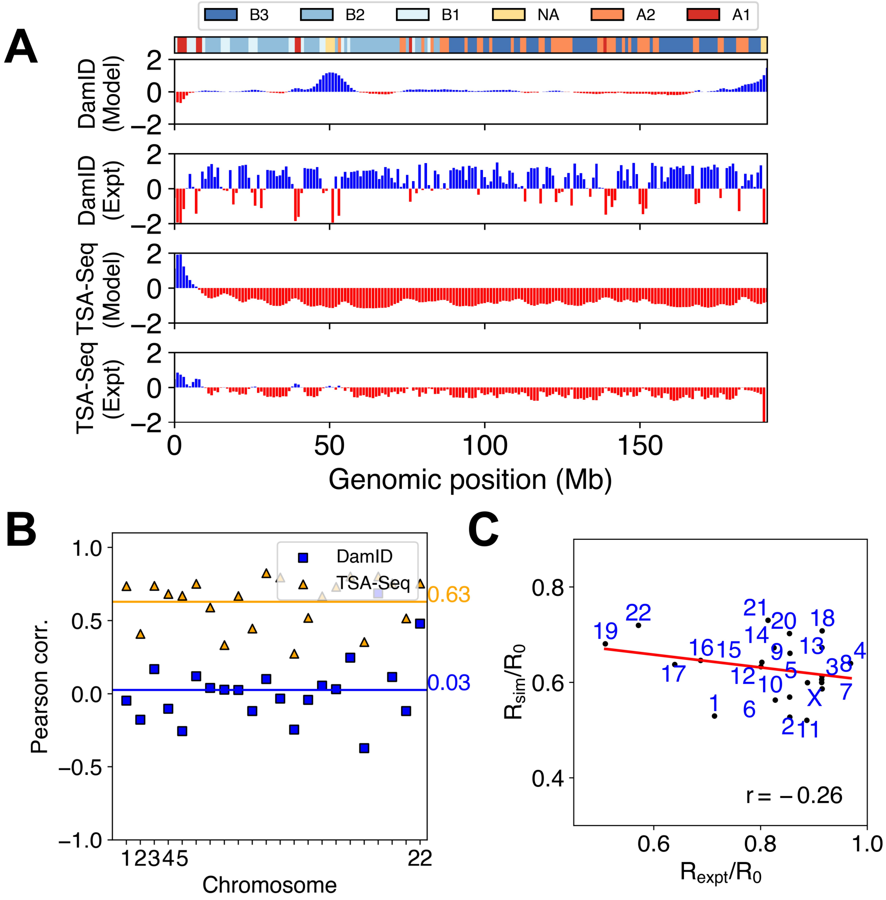


Figure S19: **Genome organization predicted by simulations that abolished chromatin-lamina interactions.** See text *Perturbing chromatin-nuclear landmark interactions* for additional discussion. (A) Comparison (for chromosome 4) between the model predicted Lamin-B DamID and SON TSA-Seq signals and the experimental data at 1MB resolution. (B) Pearson correlation coefficients between simulated and experimental Lamin-B DamID (blue) and SON TSA-Seq (yellow) data for individual chromosomes. The genome-wide averages are shown as straight lines with the corresponding values on the side. (C) Comparison of the chromosome radial positions in experiment (28) and simulations. Here R_0 is the radius of the nucleus and $R_0 = 13\sigma = 5\mu$m.

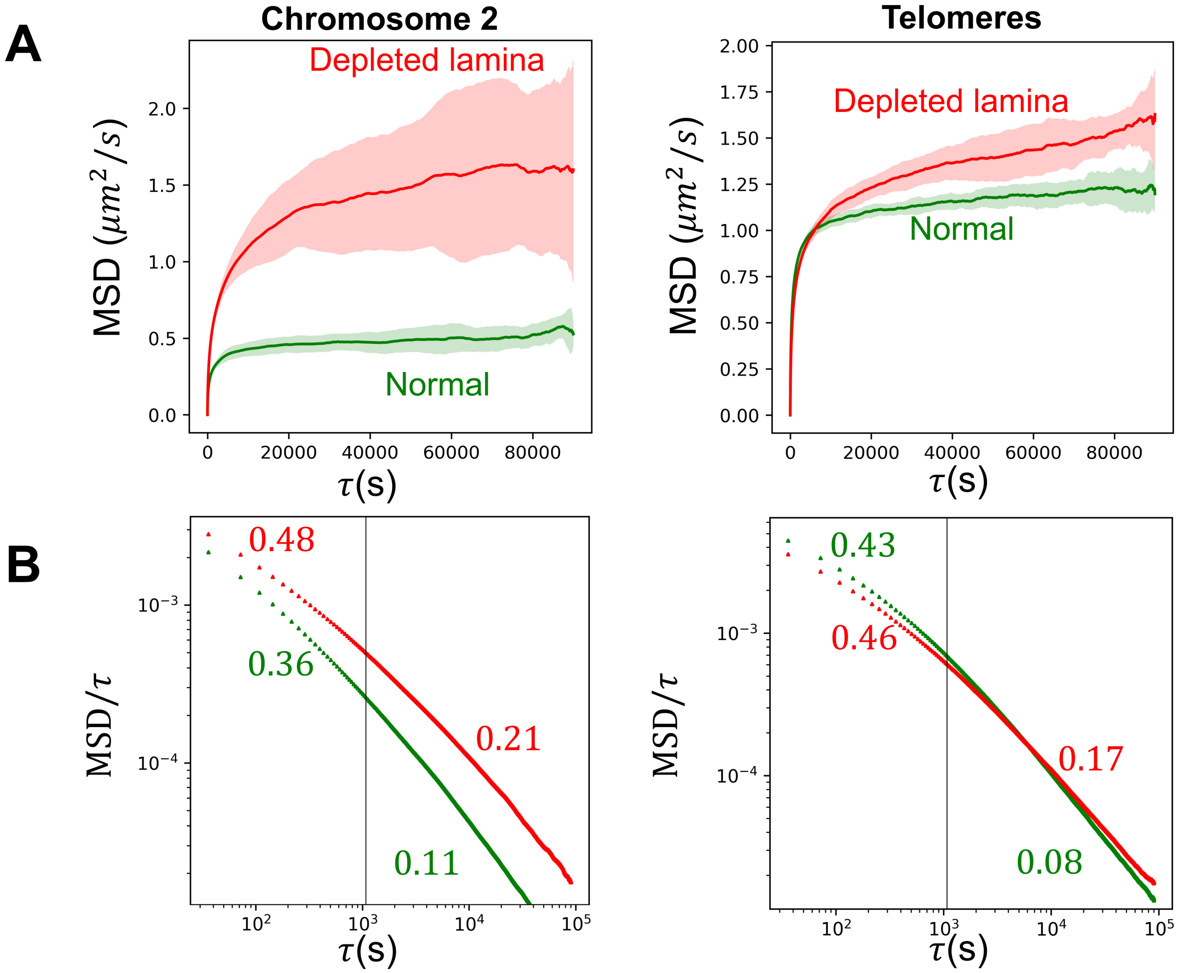


Figure S20: **Impact of chromatin-lamin interactions on the dynamics of genomic loci.** (A) The green curve is the mean square displacement (MSD) for Chromosome 2 and Telomeres averaged across 10 trajectories for the normal model. The error bars associated with the MSD is shown as a transparent region in green. Similarly, the MSD on zeroing the interaction between chromatin-lamina in the model (across 10 trajectories) is shown as red curves. Consistent with the Lamina depletion results of Chang et al., we see a pronounced effect on the dynamics of Chromosome 2 compared to the Telomeres.(16) (B) To further investigate the sub-diffusive exponent we performed careful analysis of the MSD data consistent with the procedure outlined by Bronshtein et al.(20). We converted the MSD vs τ data to MSD/τ vs τ data as the slope of $\log(\text{MSD}/\tau)$ vs $\log(\tau)$ is $\alpha - 1$ where α is exponent of the sub-diffusion process. Additionally, to estimate an average non-anomalous diffusion coefficient (in units of $\mu m^2/s$) we fit a linear line to the MSD data (skipping the first $3.6 \times 10^4 s$) We observe two regimes in the diffusion behavior (analysis boundary marked by the vertical solid black line). The exponents (α where MSD $\sim \tau^{\alpha}$) for the two regimes in the normal (green) and depleted (red) models are reported in the figure.

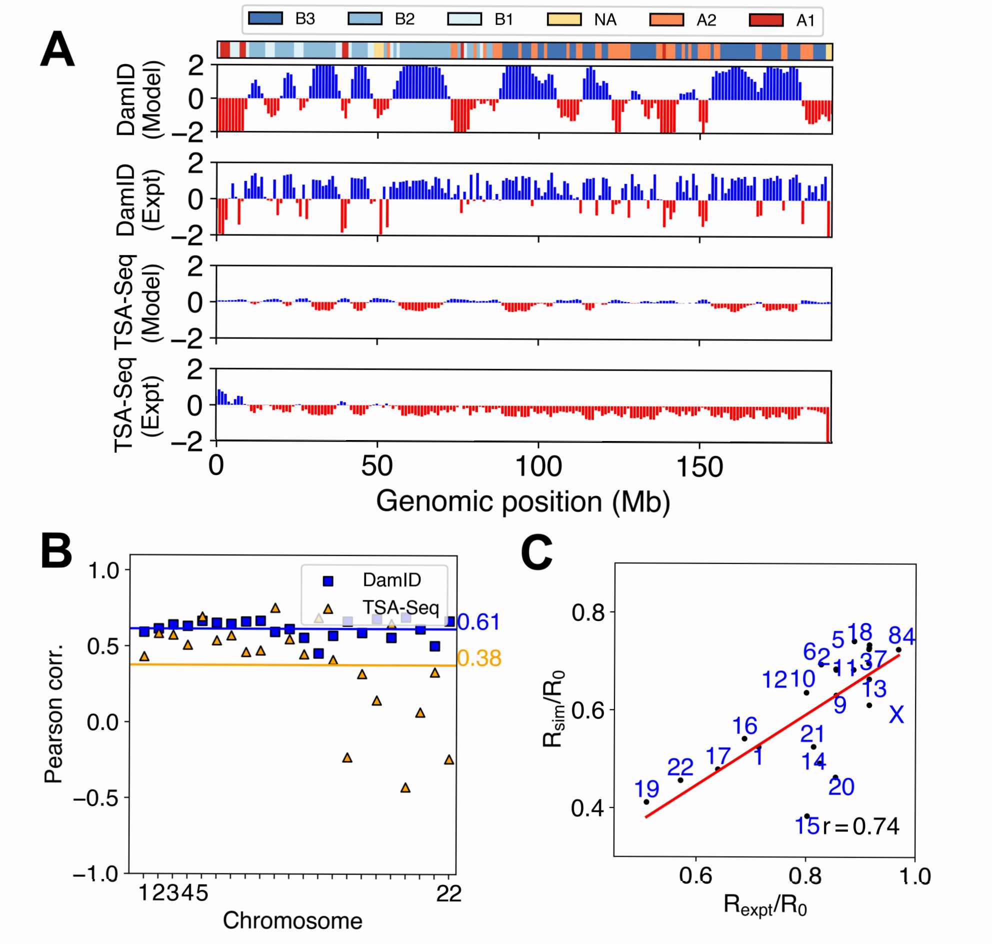


Figure S21: **Genome organization predicted by simulations that abolished chromatin-speckle interactions.** See text *Perturbing chromatin-nuclear landmark interactions* for additional discussion. (A) Comparison (for chromosome 4) between the model predicted Lamin-B DamID and SON TSA-Seq signals and the experimental data at 1MB resolution. (B) Pearson correlation coefficients between simulated and experimental Lamin-B DamID (blue) and SON TSA-Seq (yellow) data for individual chromosomes. The genome-wide averages are shown as straight lines with the corresponding values on the side. (C) Comparison of the chromosome radial positions in experiment (28) and simulations. Here R_0 is the radius of the nucleus and $R_0 = 13\sigma = 5\mu$m.

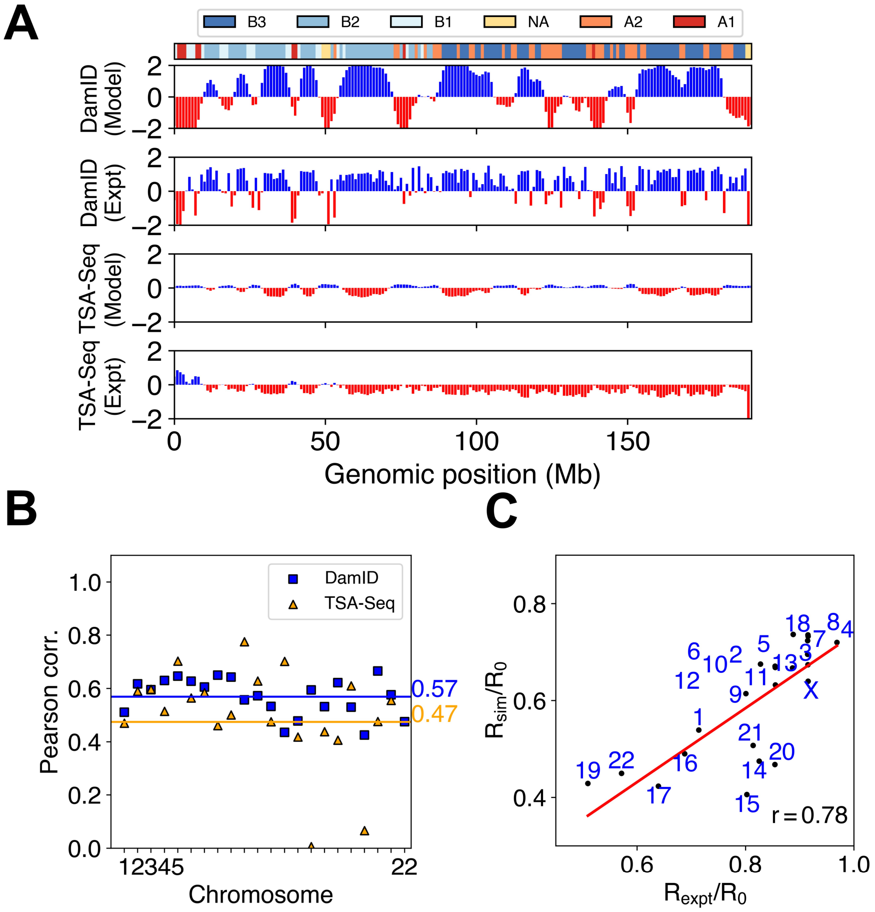


Figure S22: **Perturbing chromatin-speckle together with compartment interactions.** See text *Perturbing chromatin-nuclear landmark interactions* for additional discussion. (A) Comparison (for chromosome 4) between the model predicted Lamin-B DamID and SON TSA-Seq signals and the experimental data at 1MB resolution. (B) The model's ability to replicate TSA-Seq profiles has diminished significantly while the agreement with DamID profiles reduces slightly. (C) Although the model does not have chromatin-speckle and compartment-compartment interactions, the model preserves its capacity to replicate experimental radial positions.(28)

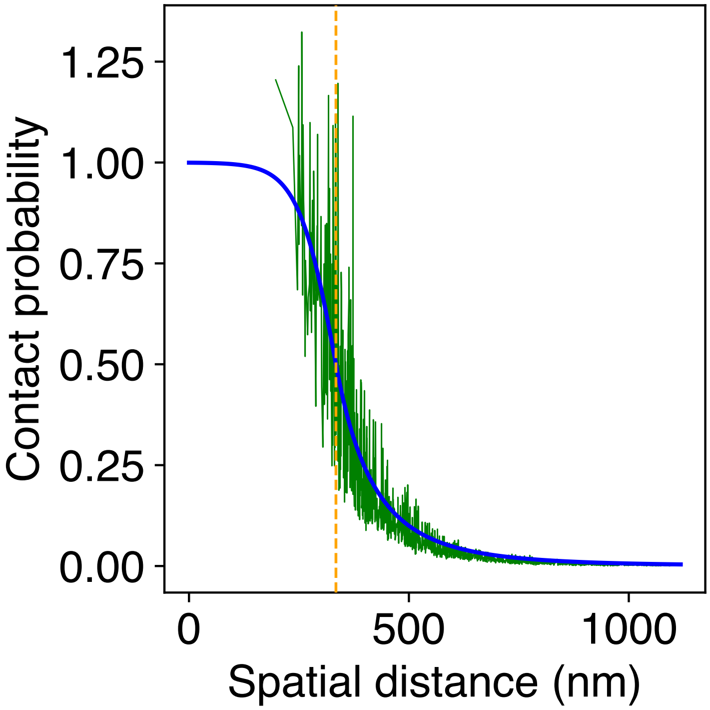


Figure S23: **Parameterizing contact function in the model using DNA-MERFISH imaging data.** The green solid lines are the raw contact probabilities from the imaging study. The blue solid line is the curve fit contact function (used in our model) that switches from a tanh function to a power law (as described by equation S10) at 334 nm (shown by the dashed orange line).

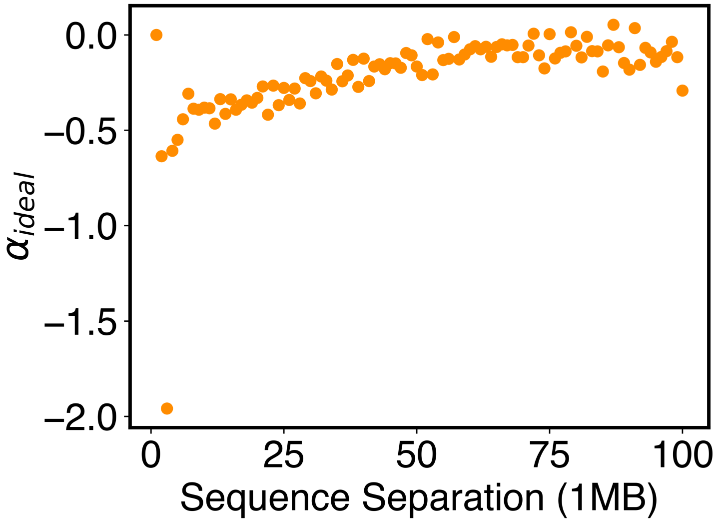


Figure S24: **Parameters in the ideal potential, α_{ideal}, as defined in Eq. S9 as a function sequence separation.** See *Section:Hi-C inspired interactions for the diploid human genome* for more text on the definition of the potential and its parameterization from Hi-C data.

REFERENCES

1. Hetzer, M. W., 2010. The nuclear envelope. *Cold Spring Harbor perspectives in biology* 2:a000539.
2. Swinbank, R., and R. J. Purser, 2006. Fibonacci grids: A novel approach to global modelling. *Quarterly Journal of the Royal Meteorological Society* 132:1769–1793.
3. Li, C., A. Ji, and Z. Cao, 2007. Stressed Fibonacci spiral patterns of definite chirality. *Applied Physics Letters* 90:164102.
4. Qi, Y., and B. Zhang, 2021. Chromatin network retards nucleoli coalescence. *Nature communications* 12:1–10.
5. Handwerger, K. E., J. A. Cordero, and J. G. Gall, 2005. Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. *Molecular biology of the cell* 16:202–211.
6. Su, J.-H., P. Zheng, S. S. Kinrot, B. Bintu, and X. Zhuang, 2020. Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* 182:1641–1659.
7. Wang, Y., Y. Zhang, R. Zhang, T. van Schaik, L. Zhang, T. Sasaki, D. Peric-Hupkes, Y. Chen, D. M. Gilbert, B. van Steensel, A. S. Belmont, and J. Ma, 2021. SPIN reveals genome-wide landscape of nuclear compartmentalization. *Genome Biol.* 22:36.
8. Brackley, C. A., J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, 2017. Nonequilibrium chromosome looping via molecular slip links. *Physical review letters* 119:138101.
9. Nomidis, S. K., E. Carlon, S. Gruber, and J. F. Marko, 2022. DNA tension-modulated translocation and loop extrusion by SMC complexes revealed by molecular dynamics simulations. *Nucleic Acids Research* 50:4974–4987.
10. Hughes, S. E., and R. S. Hawley, 2009. Heterochromatin: a rapidly evolving species barrier. *PLoS Biology* 7:e1000233.
11. Yukawa, H., 1935. On the interaction of elementary particles. I. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* 17:48–57.
12. Farley, K. I., Y. Surovtseva, J. Merkel, and S. J. Baserga, 2015. Determinants of mammalian nucleolar architecture. *Chromosoma* 124:323–331.
13. Spector, D. L., and A. I. Lamond, 2011. Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* 3:1–12.
14. Ugarte, F., R. Sousae, B. Cinquin, E. W. Martin, J. Krietsch, G. Sanchez, M. Inman, H. Tsang, M. Warr, E. Passegué, C. A. Larabell, and E. C. Forsberg, 2015. Progressive chromatin condensation and H3K9 methylation regulate the differentiation of embryonic and hematopoietic stem cells. *Stem Cell Reports* 5:728–740.
15. Johnstone, S. E., A. Reyes, Y. Qi, C. Adriaens, E. Hegazi, K. Pelka, J. H. Chen, L. S. Zou, Y. Drier, V. Hecht, N. Shoresh, M. K. Selig, C. A. Lareau, S. Iyer, S. C. Nguyen, E. F. Joyce, N. Hacohen, R. A. Irizarry, B. Zhang, M. J. Aryee, and B. E. Bernstein, 2020. Large-Scale Topological Changes Restrain Malignant Progression in Colorectal Cancer. *Cell* 182:1474–1489.
16. Chang, L., M. Li, S. Shao, C. Li, S. Ai, B. Xue, Y. Hou, Y. Zhang, R. Li, X. Fan, A. He, C. Li, and Y. Sun, 2022. Nuclear peripheral chromatin-lamin B1 interaction is required for global integrity of chromatin architecture and dynamics in human cells. *Protein Cell* 13:258–280.
17. Jiang, Z., Y. Qi, K. Kamat, and B. Zhang, 2022. Phase Separation and Correlated Motions in Motorized Genome. *J. Phys. Chem. B* `https://pubs.acs.org/doi/10.1021/acs.jpcb.2c03238`.
18. Amendola, M., and B. Steensel, 2015. Nuclear lamins are not required for lamina-associated domain organization in mouse embryonic stem cells. *EMBO Rep.* 16:610–617.
19. Zheng, X., J. Hu, S. Yue, L. Kristiani, M. Kim, M. Sauria, J. Taylor, Y. Kim, and Y. Zheng, 2018. Lamins Organize the Global Three-Dimensional Genome from the Nuclear Periphery. *Mol. Cell* 71:802–815.e7. `https://doi.org/10.1016/j.molcel.2018.05.017`.
20. Bronshtein, I., E. Kepten, I. Kanter, S. Berezin, M. Lindner, A. B. Redwood, S. Mai, S. Gonzalo, R. Foisner, Y. Shav-Tal, and Y. Garini, 2015. Loss of lamin A function increases chromatin dynamics in the nuclear interior. *Nat. Commun.* 6:8044. `http://dx.doi.org/10.1038/ncomms9044http://www.nature.com/articles/ncomms9044`.

21. Kumaran, R. I., B. Muralikrishna, and V. K. Parnaik, 2002. Lamin A/C speckles mediate spatial organization of splicing factor compartments and RNA polymerase II transcription. *J. Cell Biol.* 159:783–793.

22. Qi, Y., A. Reyes, S. E. S. Johnstone, M. M. J. Aryee, B. B. E. Bernstein, and B. Zhang, 2020. Data-Driven Polymer Model for Mechanistic Exploration of Diploid Genome Organization. *Biophys. J.* 119:1905–1916. `https://linkinghub.elsevier.com/retrieve/pii/S0006349520307190`.

23. Carrero, G., M. Hendzel, and G. De Vries, 2006. Modelling the compartmentalization of splicing factors. *Journal of theoretical biology* 239:298–312.

24. Velazquez-Dones, A., J. C. Hagopian, C.-T. Ma, X.-Y. Zhong, H. Zhou, G. Ghosh, X.-D. Fu, and J. A. Adams, 2005. Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty. *Journal of Biological Chemistry* 280:41761–41768.

25. Ester, M., H.-P. Kriegel, J. Sander, X. Xu, et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *In* kdd. volume 96, 226–231.

26. Chen, Y., Y. Zhang, Y. Wang, L. Zhang, E. K. Brinkman, S. A. Adam, R. Goldman, B. Van Steensel, J. Ma, and A. S. Belmont, 2018. Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *Journal of Cell Biology* 217:4025–4048.

27. Rao, S. S., M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–1680. `http://www.sciencedirect.com/science/article/pii/S0092867414014974`.

28. Boyle, S., S. Gilchrist, J. M. Bridger, N. L. Mahy, J. Ellis, and W. A. Bickmore, 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet.* 10:211–219. `https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/10.3.211`.