Nucleosome-constrained loop extrusion model for the origin of topologically associating domains

Mary Lou P. Bailey[1,2*], Ivan Surovtsev[3,4*], Jessica F. Williams[3*], Hao Yan[2,4*], Simon G. J. Mochrie[1,2,4] & Megan C. King[2,3,5]

[1]*Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA*

[2]*Integrated Graduate Program in Physics Engineering Biology, Yale University, New Haven, Connecticut 06511, USA*

[3]*Department of Cell Biology, Yale School of Medicine, New Haven, Connecticut 06520, USA*

[4]*Department of Physics, Yale University, New Haven, Connecticut 06511, USA*

[5]*Department of Molecular, Cell and Developmental Biology, Yale University, New Haven, Connecticut 06511, USA*

* equal contributions

Chromosome conformation capture techniques (e.g Hi-C) reveal that intermediate-scale chromatin organization is comprised of "topologically associating domains" (TADs) on the tens to thousands of kb scale.[1–5] The loop extrusion factor (LEF) model[6–10] provides a framework for how TADs arise: cohesin or condensin extrude DNA loops, until they encounter boundary elements. Despite recent *in vitro* studies demonstrating that cohesin and condensin can drive loop formation on (largely) naked DNA[11–13], evidence supporting the LEF model in living cells is lacking. Here, we combine experimental measurements of

chromatin dynamics *in vivo* with simulations to further develop the LEF model. We show that the activity of the INO80 nucleosome remodeler enhances chromatin mobility, while cohesin and condensin restrain chromatin mobility. Motivated by these findings and the observations that cohesin is loaded preferentially at nucleosome-depleted transcriptional start sites[14–18] and its efficient translocation requires nucleosome remodeling[19–23] we propose a new LEF model in which LEF loading and loop extrusion direction depend on the underlying architecture of transcriptional units. Using solely genome annotation without imposing boundary elements, the model predicts TADs that reproduce experimental Hi-C data, including boundaries that are CTCF-poor. Furthermore, polymer simulations based on the model show that LEF-catalyzed loops reduce chromatin mobility, consistent with our experimental measurements. Overall, this work reveals new tenets for the origins of TADs in eukaryotes, driven by transcription-coupled nucleosome remodeling.

Chromatin organization is inextricably linked to its dynamics. The LEF model posits that multiple dynamic DNA loops appear and disappear in a stochastic fashion, leading to a statistical pattern of increased intra-domain interactions (i.e. TADs)[1–5]. To gain insights into the factors that contribute to dynamic chromatin organization *in vivo*, we developed a live-cell imaging, tracking, and analysis approach to precisely characterize the spatial and temporal dynamics of chromatin loci at short (ms-s) timescales (Fig. 1a). We combine imaging of chromosomal loci dynamics in fission yeast by *lac* operator/repressor fluorescence microscopy and subsequent motion analysis. On the seconds timescale we observe a characteristic subdiffusive behavior in fission yeast that is nearly identical at six different genomic positions (Figs. 1b, Extended Data Fig. 1a and

Supplementary Information). Although microtubules drive centromere oscillations in fission yeast, we find that microtubule dynamics play a relatively minor role on chromatin mobility outside the centromeres on this time scale (Extended Data Fig. 1b). A single mean anomalous exponent ($\alpha = 0.44 \pm 0.02$) provides a good description for all experimental MSDs for the first 29 time points (1.7 seconds) as indicated by the fit (solid lines) to the theoretical MSD (Eq. 1, see Methods). This mean value is slightly lower than that predicted for an ideal Rouse polymer of $\alpha = 0.5$[24], which is the value observed for budding yeast chromatin[25] (see also Extended Data Fig. 1c). Using $\alpha = 0.44$, the "diffusion coefficient", D, ranges from $0.0050 - 0.0062\ \mu\text{m}^2\ \text{s}^{-0.44}$. Consistent with prior studies[26–28], depletion of ATP by the addition of sodium azide decreases the diffusivity about two-fold (from 0.0056 to $0.0029\ \mu\text{m}^2\ \text{s}^{-0.44}$ for the *mmf1* locus (Fig. 1c)).

As loop extrusion by the Structural Maintenance of Chromosomes (SMC) complexes, cohesin and condensin, requires ATP, an attractive possibility is that SMC activity underlies ATP-dependent chromatin motion. Surprisingly, however, employing critical loss-of-function, temperature-sensitive alleles of the cohesin loading factor (*mis4-242*) or of the SMC2 subunit of condensin (*cut14-208*) reveals a profound increase in chromatin mobility (30 to 50% increase in D) at either the *mmf1* or *pfl5* locus (Fig. 2a-b). These observations suggest (1) that SMC-associated DNA loops constrain chromatin movement on this timescale and (2) that SMC complexes are not required for ATP-dependent chromatin mobility.

Single-molecule experiments show that nucleosomes impede cohesin translocation along DNA[29] and that nucleosome removal is required for efficient loop extrusion in *Xenopus* egg

extracts[13]. Moreover, numerous studies suggest that cohesin loading on chromosomes occurs at nucleosome-depleted regions (often associated with transcription start sites (TSSs)) and requires ATP-dependent nucleosome remodeling activities[14–18]. In addition, it has been suggested that cohesin translocation requires transcription-coupled nucleosome remodeling[19–23]. To test whether nucleosome remodeling contributes to ATP-dependent chromatin mobility, we examined cells lacking Arp8, an auxiliary component of the essential INO80 remodeling complex, and Arp9, a component of both the SWR1 and RSC remodeling complexes. Loss of Arp9 had little effect (Fig. 2c). By contrast, loss of Arp8 led to a clear suppression of chromatin mobility ($D = 0.0047\ \mu\text{m}^2\ \text{s}^{-0.44}$, representing a 40% decrease in the ATP-dependent diffusivity above the ATP-depleted background motion). These results are consistent with a prior study in which INO80 (but not SWR1 or RSC) was both required for a transcription-dependent boost in chromatin mobility and sufficient to impart increased diffusivity when artificially targeted to a chromosomal locus[30]. We next examined how loss of Arp8 would influence the elevated chromatin diffusivity observed in SMC complex mutants. Loss of Arp8 quantitatively dampens the elevated chromatin dynamics in the *cut14-208* background at the non-permissive temperature ($D = 0.0061\ \mu\text{m}^2\ \text{s}^{-0.44}$); the magnitude of the effect is qualitatively similar to that observed in WT cells (an approximately 30% decrease of the ATP-dependent diffusivity) (Fig. 2d). However, the level of absolute diffusivity remains substantially elevated compared to cells lacking only Arp8, suggesting that loss of looping interactions increases the observed chromatin mobility even the absence of Arp8.

Previously described LEF models assume random LEF loading on DNA followed by

bi-directional loop extrusion until the LEF encounters boundary elements, the permeability of which are dictated by CTCF occupancy[7–10]. As expected based on the strong correlation between CTCF binding and TAD boundaries, these models recapitulate experimentally determined TAD patterns[6,9,10]. However, if chromatin restricts where LEFs can be loaded to TSSs[14–18,31] and has to be remodeled to allow loop extrusion, then the organization of transcriptional units (position, size, and direction) necessarily affects the size and location of loops and therefore the TAD pattern. We therefore hypothesized that *in vivo* translocating LEFs require both an SMC complex and a nucleosome remodeling complex, effectively being composite objects (Fig. 3a). In the simplest scenario, a remodeler translocates along the genome, removing nucleosomes. One side of the SMC then follows in the wake of the remodeler, held close by nucleosome rebinding behind the SMC. In this model the SMC and nucleosome remodeler constitute a one-sided LEF, which translocates efficiently via a ratchet mechanism on nucleosomal DNA (see Supplementary Information). Within genes, it seems natural to identify the relevant remodeler with the transcriptional machinery.

We synthesized these ideas into Gillespie-type simulations[32] dictated by the following: (1) SMCs bind only at TSSs that are not already occupied by another SMC; (2) loop extrusion is one-sided, with one side of a composite LEF, involving both an SMC and a remodeler, translocating in the direction of transcription while the other side remains at the TSS; (3) LEFs cannot pass each other; and (4) LEF translocation continues (including across multiple genes) until it dissociates from the DNA or it encounters another LEF (see Supplementary Information for more details). (1) and (2) stand in sharp contrast to earlier LEF models, in which LEFs can bind anywhere on

the genome and loop extrusion is two-sided[6–10]. Assuming that looped chromatin behaves as a random chain polymer, we transformed the results of these simulations into contact maps (see Supplementary Information). While in existing LEF models loops positions and sizes rely on boundary elements such as CTCF, in this model the patterns of loops arise from stalling of a translocating LEF leg by another LEF leg that is retained at the TSS due to nucleosome constraint. In other words, 1D genome organization defines the TAD pattern.

We simulated several regions of the fission yeast (300 kb regions, Extended Data Fig. 2) and mouse genomes (6 Mb regions Figs. 3b, 3e, Extended Data Figs. 3 and 4 Nucleosome-constrained model) using genome annotation and the LEF activity described above as the only inputs and compared the resultant simulated Hi-C maps (rendered at 10 kb-resolution) to available experimental maps of the same regions. Although there are differences, it is striking that the simulated maps recapitulate the overall pattern of TADs including strong TAD boundaries. To make quantitative comparisons, we used the second moment of the Hi-C map at genomic location i, namely $M_i = \Sigma_{j=0}^{100} j^2 C_{i+j,i-j}$, where $C_{k,l}$ is the probability that genomic location k contacts genomic location l. We truncated the sum at $j = 100$, corresponding to 1 Mb, beyond which $C_{i+j,i-j}$ is essentially zero. We then optimized model parameters (see Supplementary Information Table S2) by maximizing the M_i cross correlation between experimental and simulated maps for three regions (*e.g.* Fig. 3b). For example, we find that LEFs must necessarily translocate across the regions between genes to make large loops. Using these optimized parameters (see Supplementary Information Table S2), we then carried out simulations for an additional seven regions (Fig. 3e and Extended Data Figs. 3 and 4). The Pearson cross correlation (R) for the regions in Fig. 3b and 3e

are $R = 0.57$ and $R = 0.63$, respectively. The mean cross correlation for all ten regions (Fig. 3 and Extended Data Figs. 3 and 4) is $R = 0.47 \pm 0.06$; if we instead compare unmatched experimental and simulated maps the cross correlation is close to zero ($R = -0.03 \pm 0.04$).

Next we simulated the LEF model described in ref.[9] for the same regions of the mouse genome following the same optimization procedure, and also using CTCF ChIP-seq data to determine the probability that a LEF passes each CTCF binding location[33]. The corresponding simulated maps (Figs. 3c and 3f, Extended Data Figs. 3, and 4 CTCF model) also qualitatively and quantitatively recapitulate many features of the experimental Hi-C map, as expected from ref.[9] ($R = 0.56$ is the mean cross correlation). By this measure as well as by visual inspection, our model explains Hi-C maps comparably well to the CTCF model, despite using entirely different input information. We also simulated a "hybrid" model which combines the nucleosome-constrained model with CTCF boundary elements (Figs. 3d and 3g and Extended Data Figs. 3, and 4). Compared to the simpler nucleosome constrained model it appears that boundaries remain at the same positions but become less permeable, suggesting that CTCF reinforces pre-existing TADs. The origin of "CTCF-less" boundaries frequently observed in experiments is an open question [1,33]. Importantly, the nucleosome-constrained model captures TAD boundaries (see arrow in Fig. 3e), which are not predicted by the CTCF-dependent LEF model (Fig. 3f). Our simulations predict a broad range of loop sizes (Fig. 4a), perhaps mirroring the inherent variability in transcriptional unit characteristics in the regions used for the simulations. In mouse, the predicted mean loop size is 222 kb, corresponding to an average of two genes per loop (Fig. 4b), which is on the same scale as TADs derived from high-resolution

mouse Hi-C data[34].

We next asked how loop extrusion influences MSD measurements by coupling loop extrusion outputs to Rouse-model polymer simulations, in which we consider every 10 kb region of the mouse (0.5 kb for yeast) genome as a bead connected to its neighbors by springs of stiffness, κ and experiencing a friction coefficient, ζ. To incorporate loops, we augment the usual nearest-neighbor Rouse model with additional springs (of the same spring constant) that connect non-adjacent beads, each representing the base of a loop. The number of loops and their locations evolve according to our loop extrusion simulations. The MSD of each gene locus is then determined from the resultant time series of bead positions. Inspection of 2D-projected polymer configurations (insets in Fig. 4c (yeast) and Fig. 4d (mouse) with (blue) and without (red) loops immediately reveals that loops drive an overall more compact chromatin configuration (see also Supplementary Video). In addition, a comparison between the MSDs with (blue) and without (red) loops in both yeast (Fig. 4c) and mouse (Fig. 4d) shows that loops constrain chromatin motion, consistent with our experimental MSD measurements of gene loci in *S. pombe* with and without disrupted SMC loading or activity (Fig. 2). We next calculated how the effective exponent, $\alpha = \frac{\log \mathrm{MSD}(t_{n+1}) - \log \mathrm{MSD}(t_n)}{\log t_{n+1} - \log t_n}$, is influenced by loop extrusion. As expected, the effective exponents for the Rouse model without loops lie close to 1/2 (Fig. 4e-f). Remarkably, the effective exponent in the presence of loops shows a stronger time dependence, declining from a value near $\alpha = 1/2$ for $t = 5\tau_P$ ($\tau_P = \frac{\zeta}{\pi^2 \kappa}$) to a value near $\alpha = 0.4$ for $t = 10^3 \tau_P$. This result is strikingly reminiscent of our finding that experimental MSDs in *S. pombe* are well-described overall by a value of $\alpha = 0.44$ (Fig. 1 and Extended Data Fig. 1a).

Whether TADs regulate transcription and/or are a consequence of transcription is currently under intense debate [35–37]. Here, we present evidence that a nucleosome-contrained LEF model can predict TADs based solely on the organization of transcriptional units, implying that transcription can be a major determinant of local chromatin organization. On top of this backdrop, further elaborations likely play important roles, including factors that recognize specific DNA sequences and/or bridge two genomic regions, such as CTCF. This picture is consistent with numerous prior observations including the prevalence of cohesin at TSSs[14–18] that is enhanced in response to transcriptional inhibition[38] and the accumulation of cohesin at sites of transcription termination, particularly at convergent genes in yeast[21]. Our model suggests that CTCF likely reinforces rather than originates TADs, consistent with the variability in the predictive nature of CTCF binding sites to TAD boundaries in some contexts[33], the observation that depletion of cohesin has a much stronger effect on TAD architecture than CTCF depletion[38,39], and the full maintenance of insulation at 20% of CTCF-positive boundaries upon CTCF depletion[40]. While cohesin and condensin have been shown to be capable of bidirectional loop extrusion on naked DNA *in vitro*[12,13] and two-sided extrusion has been suggested to be a key property for generating Hi-C maps evocative of experimental data[41], here we suggest that one-sided extrusion may frequently occur on fully chromatized DNA that dominates *in vivo*. While recent simulations highlight strong TAD corner peaks as a characteristic feature of two-sided loop extrusion,[41] such peaks are relatively rare in Hi-C data without extensive averaging (see Fig. 3 and Extended Data Figs. 3 and 4), suggesting that one-sided extrusion is compatible with most TADs, as well as with experimentally-observed "stripes"[42], which are observed in our simulations (Fig. 3 and Extended

Data Figs. 3 and 4). Importantly, our model is consistent with the observed ability of cohesin to extrude and compact DNA with a low level of incorporated nucleosomes (about 1 per 8 kb of DNA)[12], which would be permissive to cohesin loading[17] and loop extrusion between nucleosomes, achieving substantial DNA compaction. Our work also underscores that factors underlying chromatin looping are intimately linked to chromatin dynamics. A substantial fraction of ATP-dependent chromatin mobility can be accounted for by nucleosome remodeling, consistent with ATP-dependent increases in the MSD of chromatin loci in response to transcriptional activation in mammalian cells[43]. We also find that DNA looping mediated by LEFs predominantly constrains chromatin motion, leading to reduced MSDs as well as values of $\alpha < 0.5$ both *in vivo* and in simulations. Ultimately, the simple model put forward here will be useful for predicting the underlying genome organization for any eukaryote for which there is genome annotation, but particularly for those outside of bilateria that lack CTCF and/or employ other insulator proteins, such as the widely conserved TFIIIC complex[44].

Methods

Strain generation and culturing. All strains used in this study are listed in Supplementary Information Table S3. *S. pombe* were grown, maintained, and crossed using standard procedures and media.[45] Gene replacements were made by exchanging open reading frames with various MX6-based drug resistance genes.[46,47] Targeted *lacO* array insertions were generated using a modified two-step integration procedure that first creates a site-specific DNA double-strand break to increase targeting efficiency of the linearized plasmid pSR10_ura4_10.3kb[48,49]; integration sites are listed in Supplementary Information Table S3. Random *lacO* array insertions were generated by transformation of the GFP-LacI-expressing strain MKSP1120 with linearized plasmid pSR10_ura4_10.3kb followed by selection on plates lacking uracil as in ref.[48]. Strains that had successfully inserted the *lacO* array were identified by fluorescence microscopy. The *S. cerevisiae* strains used in this study were constructed, grown and imaged as described in ref.[50]

Microscopy. Cells were grown in YE5S media (yeast extract plus five supplements; (5 g/L yeast extract (BD Bacto), 30 g/L glucose (Sigma Aldrich), 0.25 g/L adenine, 0.25 g/L histidine, 0.25 g/L leucine, 0.25 g/L lysine, and 0.25 g/L uracil, (Sunrise Science Products) media at 30°C to log phase (OD_{600} 0.5-0.8). For temperature-sensitive alleles, the cells were grown at 30°C, then shifted for twenty minutes to the non-permissive temperature (36°C) prior to imaging. For pharmacological inhibition studies, cells were incubated at 30°C with sodium azide (0.02% w/v) or with 50 mg/mL MBC (methyl-2-benzimidazol-2-ylcarbamate) for 10 minutes prior to imaging. To fix cells, cells were incubated with 2% paraformaldehyde for 10 min and washed once with 2X PBS (phosphate buffer solution).

For imaging, 1.4 μL of the concentrated cell suspension was transferred to a 1 cm $\times$ 1 cm $\times$ 1 mm pad made of 1.4% agarose (Denville Scientific) in EMM media (Edinburgh minimal media) plus five supplements 0.25 g/L adenine, 0.25 g/L histidine, 0.25 g/L leucine, 0.25 g/L lysine, and 0.25 g/L uracil (Sunrise Science Products) on a microscope slide. The cell pad was covered with a #1.5 22 $\times$ 22 mm coverslip, and edges were sealed with VALAP (1:1:1 petroleum jelly, lanolin, and paraffin) to limit evaporation during imaging.

Fluorescence and bright field images were acquired on a DeltaVision widefield microscope (Applied Precision/GE) equipped with a temperature control chamber, 1.4 NA 100$\times$ objective (Olympus), solid-state-based illumination (Lumencor), and an Evolve 512 EMCCD camera (Photometrics). Fluorescence was excited at 488 nm and collected with emission filters between 500-550 nm. Typically, 1000 images were continuously acquired each with a 10 ms exposure time at a rate of 1 frame per 58 ms.

Loci tracking and trajectory analysis. Localization of each fluorescent focus (i.e. labeled locus) was done by applying a spatial bandpass filter to each raw image frame. Specifically, the centroid of each peak above a determined threshold value was calculated to obtain an approximate position. Then, the final localization was obtained by fitting a symmetric 2D Gaussian function around each centroid position. These positions were then linked into time trajectories using the algorithm for single particle tracking introduced in ref.[51]. Specifically, we did not allow gaps between frames, set the maximum displacement a focus can move between frames to 2 pixels (0.321 μm^2) and set the minimum track length to be 300 frames, or 100 frames for the experiments involving *arp8*Δ, *arp9*Δ, and *cut14*^{ts}*arp*8Δ, and sodium azide.

The mean-square displacement (MSD) versus time delay for each time-lapse image series was calculated as follows. First, the trajectories were split into sub-trajectories of 29 steps (with total duration 1.68 s). Then, individual displacements were averaged across all sub-trajectories and different time-lapse series for a given experiment. The calculated experimental MSDs were fit by the theoretical MSD for fractional Brownian motion (fBm), including static and dynamic localization errors[52]:

$$M(n) = \frac{2D(\Delta t)^{2+\alpha}\left((n-\frac{\Delta t_E}{\Delta t})^{2+\alpha} - 2n^{2+\alpha} + (n+\frac{\Delta t_E}{\Delta t})^{2+\alpha}\right)}{(1+\alpha)(2+\alpha)\Delta t_E^2} - \frac{4D\Delta t_E^{\alpha}}{(1+\alpha)(2+\alpha)} + 2\sigma^2, \qquad (1)$$

where D is an effective diffusivity, $\Delta t_E = 10$ ms is exposure time, Δt is the time step between frames, and α is an exponent ($\alpha < 1$ corresponds to subdiffusive motion and $\alpha > 1$ to superdiffusive motion), n is number of frames, and σ is the static localization error. Initial MSD fitting, in which D, α, and σ were varied, showed that the values of α grouped around a mean of 0.44 for most measured loci in unperturbed wild type cells (Supplementary Information Table S1). In all subsequent analyses (see Figs. 1 and 2 and Extended Data Fig. 1a), α was fixed at $\alpha = 0.44$, while D and σ were varied, allowing us to directly compare values of D as a measure of locus mobility across different experiments. Supplementary Information Table S1 displays all the fitting results.

Gillespie simulations. We carried out Gillespie-type simulations[32] of the chromatin loop configuration, considering the following possible reactions: SMC binding at an unoccupied transcriptional start site (TSS), LEF activity initiation, SMC unbinding, and single-step LEF

translocation into an unoccupied site. We envision that LEF activity initiation corresponds to remodeler-binding at a TSS that already is occupied by an SMC. Treating remodeler binding in this way streamlines the simulation in comparison to simulations that include explicit remodelers. Upon initiation (remodeler binding), a composite LEF forms and starts to extrude a chromatin loop until it dissociates or is blocked. We performed ten simulations each comprising 1,000,000 time steps. Loop dynamics reached a steady state over about 15,000 steps as gauged by convergence of the chromatin backbone length fluctuations at a constant value. 6 Mb (300 kb) genomic regions for mouse (yeast) were divided into 10 kb (0.5 kb) bins creating 600 sites that could be occupied by one side of a LEF. Typically, there are 100 TSSs per 6 Mb genomic region in mouse and 160 TSSs per 300 kb genomic region in yeast. TSS were allocated to specific bins using genome annotation using the genome features curated by Mouse Genome Informatics[53] (mouse) and PomBase[54] (*S. pombe*). In case of bins containing two TSS (comprising 5% of all sites), we chose that only one LEF can be present at any given time. To calculate contact probabilities from the configuration of loops, we model chromatin as a Gaussian polymer (see Supplementary Information). Thus, we calculated the effective genomic distance between two loci (see Supplementary Information), which leads to the probability density of the separation. For every 1000 steps in steady state, an instantaneous contact map was generated based on the contact probability between all pairs of loci. All these instantaneous maps were then time-averaged to create the final simulated map. The simulation parameters, namely: the stepping rate, the binding rate per molecule, the unbinding rate (inversed lifetime), the initiation rate, and the total number of LEFs per 6 Mb (300 kb) genomic region, are optimized by maximizing the

Pearson cross correlation between the second moment of the simulated map and that of the corresponding experimental Hi-C map (see Supplementary Information Table S2). Our loop-extrusion simulation code is online at the github repository: `https://github.com/nilesyan/LEF-Simulation`.

Polymer simulations. Chromatin's polymer behavior was modeled as a Rouse (bead-spring) chain, in which each 10 kb (0.5 kb) region of the mouse (yeast) genome is represented as a bead, that is connected to its nearest neighbors by springs of stiffness, κ, and experiences fluid friction with coefficient, ζ, To include loops, we augmented the model with additional springs (of the same spring constant) that connect non-adjacent beads, representing the base of a loop. For the simulations shown in Fig. 4, the characteristic polymer time, namely $\tau_P = \frac{\zeta}{\pi^2 \kappa}$, was picked on the basis of the parameters determined in ref.[55], assuming a LEF translocation velocity of 100 bp s^{-1}. However, the resultant MSDs are indistinguishable from those obtained for LEF velocities of 10 bp s^{-1} and 1000 bp s^{-1}. The number, sizes, and locations of loops evolve according to our loop extrusion simulations. Since the additional springs lead to far-from-diagonal terms in the Rouse-model dynamical matrix, the matrix was diagonalized numerically to find the eigenmodes and corresponding beads-to-eigenmode transformations. Assuming equipartition with an effective temperature, these eigenmodes (and, through transformations, the corresponding bead positions) were then evolved using the exact simulation method of ref.[56], until the loop configuration changed. This procedure was then repeated for each new loop configuration. Because this approach allows us to simulate polymer dynamics at arbitrary time points, in our simulations, we combined an equally-spaced time series (with time step shorter than the polymer time) with the

exact times of loop extrusion events, ensuring an exact simulation in the presence of loop extrusion. Finally, mean MSDs were calculated by averaging MSDs of individual beads (Fig. 4c-d) .

1. Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin domains: The unit of chromosome organization. *Mol Cell* **62**, 668–80 (2016).

2. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458–72 (2012).

3. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80 (2012).

4. Mizuguchi, T. *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *s. pombe*. *Nature* **516**, 432–5 (2014).

5. Dekker, J. Two ways to fold the genome during the cell cycle: Insights obtained with chromosome conformation capture. *Epigenetics Chromatin* **7**, 25 (2014).

6. Alipour, E. & Marko, J. F. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res* **40**, 11202–12 (2012).

7. Goloborodko, A., Marko, J. F. & Mirny, L. A. Chromosome compaction by active loop extrusion. *Biophys J* **110**, 2162–8 (2016).

8. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5** (2016).

9. Fudenberg, G. *et al.* Formation of chromosomal domains by loop extrusion. *Cell Rep* **15**, 2038–49 (2016).

10. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A* **115**, E6697–E6706 (2018).

11. Ganji, M. *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).

12. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345–1349 (2019).

13. Golfier, S., Quail, T., Kimura, H. & Brugués, J. Cohesin and condensin extrude loops in a cell-cycle dependent manner. *bioRxiv* (2019). `https://www.biorxiv.org/content/early/2019/10/29/821306.full.pdf`.

14. Lopez-Serra, L., Kelly, G., Patel, H., Stewart, A. & Uhlmann, F. The scc2-scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. *Nat Genet* **46**, 1147–51 (2014).

15. Petela, N. J. *et al.* Scc2 is a potent activator of cohesin's ATPase that promotes loading by binding scc1 without pds5. *Mol Cell* **70**, 1134–1148 e7 (2018).

16. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A* **111**, 996–1001 (2014).

17. Munoz, S., Minamino, M., Casas-Delucchi, C. S., Patel, H. & Uhlmann, F. A role for chromatin remodeling in cohesin loading onto chromosomes. *Mol Cell* **74**, 664–673 e5 (2019).

18. Garcia-Luis, J. *et al.* Fact mediates cohesin function on chromatin. *Nat Struct Mol Biol* **26**, 970–979 (2019).

19. Glynn, E. F. *et al.* Genome-wide mapping of the cohesin complex in the yeast *saccharomyces cerevisiae*. *PLoS Biol* **2**, E259 (2004).

20. Dubey, R. N. & Gartenberg, M. R. A tDNA establishes cohesion of a neighboring silent chromatin domain. *Genes Dev* **21**, 2150–60 (2007).

21. Lengronne, A. *et al.* Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573–8 (2004).

22. Schmidt, C. K., Brookes, N. & Uhlmann, F. Conserved features of cohesin binding along fission yeast chromosomes. *Genome Biol* **10**, R52 (2009).

23. Ocampo-Hafalla, M., Munoz, S., Samora, C. P. & Uhlmann, F. Evidence for cohesin sliding along budding yeast chromosomes. *Open Biol* **6** (2016).

24. Doi, M. & Edwards, S. *The Theory of Polymer Dynamics*. International series of monographs on physics (Clarendon Press, 1988).

25. Weber, S. C., Spakowitz, A. J. & Theriot, J. A. Nonthermal ATP-dependent fluctuations contribute to the *in vivo* motion of chromosomal loci. *Proc Natl Acad Sci U S A* **109**, 7338–43 (2012).

26. Marshall, W. *et al.* Interphase chromosomes undergo constrained diffusional motion in living cells. *Current Biology* **7**, 930 – 939 (1997).

27. Heun, P., Laroche, T., Raghuraman, M. K. & Gasser, S. M. The positioning and dynamics of origins of replication in the budding yeast nucleus. *J Cell Biol.* **152**, 385–400 (2001).

28. Levi, V., Ruan, Q., Plutz, M. J., Belmont, A. S. & Gratton, E. Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. *Biophysical journal* **89 6**, 4275–85 (2005).

29. Stigler, J., Camdere, G. O., Koshland, D. E. & Greene, E. C. Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin. *Cell Rep* **15**, 988–998 (2016).

30. Neumanne, F. R. *et al.* Targeted INO80 enhances subnuclear chromatin movement and ectopic homologous recombination. *Genes and Dev.* **26**, 369–83 (2012).

31. Shintomi, K., Takahashi, T. S. & Hirano, T. Reconstitution of mitotic chromatids with a minimum set of purified factors. *Nat Cell Biol* **17**, 1014–23 (2015).

32. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**, 2340–2361 (1977).

33. Bonev, B. *et al.* Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572 e24 (2017).

34. Rao, S. S. P. *et al.* Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320 e24 (2017).

35. Mir, M., Bickmore, W., Furlong, E. E. M. & Narlikar, G. Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? *Development* **146** (2019).

36. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nature Genetics* **52**, 8–16 (2020).

37. Ibrahim, D. M. & Mundlos, S. Three-dimensional chromatin in disease: What holds us together and what drives us apart? *Current Opinion in Cell Biology* **64**, 1 – 9 (2020).

38. Busslinger, G. A. *et al.* Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503–507 (2017).

39. Wutz, G. *et al.* Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* **36**, 3573–3599 (2017).

40. Nora, E. P. *et al.* Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 e22 (2017).

41. Banigan, E. J., van den Berg, A. A., Brandão, H. B., Marko, J. F. & Mirny, L. A. Chromosome organization by one-sided and two-sided loop extrusion. *bioRxiv* 815340 (2019).

42. Mehta, I. S., Kulashreshtha, M., Chakraborty, S., Kolthur-Seetharam, U. & Rao, B. J. Chromosome territories reposition during DNA damage-repair response. *Genome Biol* **14**, R135 (2013).

43. Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**, 1050–1055 (2018).

44. Van Bortle, K. & Corces, V. G. tDNA insulators and the emerging role of TFIIIC in genome organization. *Transcription* **3**, 277–84 (2012).

45. Moreno, S., Klar, A. & Nurse, P. Molecular genetic analysis of fission yeast *schizosaccharomyces pombe*. *Methods in enzymology* **194**, 795–823 (1991).

46. Bähler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in *schizosaccharomyces pombe*. *Yeast* **14**, 943–951 (1998).

47. Hentges, P., Van Driessche, B., Tafforeau, L., Vandenhaute, J. & Carr, A. M. Three novel antibiotic marker cassettes for gene disruption and marker switching in schizosaccharomyces pombe. *Yeast* **22**, 1013–9 (2005).

48. Leland, B. A. & King, M. C. Using lacO arrays to monitor DNA double-strand break dynamics in live *schizosaccharomyces pombe* cells. *Methods Mol Biol* **1176**, 127–41 (2014).

49. Rohner, S., Gasser, S. M. & Meister, P. Modules for cloning-free chromatin tagging in *saccharomyces cerevisae*. *Yeast* **25**, 235–239 (2008).

50. Colombi, P., King, D. E., Williams, J. F., Lusk, C. P. & King, M. C. LEM domain proteins control the efficiency of adaptation through copy number variation. *bioRxiv* 451583 (2018).

51. Crocker, J. C. & Grier, D. G. Methods of digital video microscopy for colloidal studies. *J. Coll. Int. Sci.* **179**, 298–310 (1996).

52. Savin, T. & Doyle, P. S. Static and dynamic errors in particle tracking microrheology. *Biophysical Journal* **88**, 623 – 638 (2005).

53. The Jackson Laboratory, M., Bar Harbor. Mouse genome database (mgd) at the mouse genome informatics website.

54. Lock, A. *et al.* Pombase 2018: user-driven reimplementation of the fission yeast database provides rapid and intuitive access to diverse, interconnected information. *Nucleic Acids Res.* D821–D827 (2019).

55. Arbona, J.-M., Herbert, S., Fabre, E. & Zimmer, C. Inferring the physical properties of yeast chromatin through Bayesian analysis of whole nucleus simulations. *Genome biology* **18**, 81–81 (2017).

56. Gillespie, D. T. The mathematics of Brownian motion and Johnson noise. *American Journal of Physics* **64**, 225–240 (1996).

Acknowledgements We thank Bryan Leland, Amanda Vines, Paolo Colombi, Dongxu Lin, Sara Siwiecki, and Ryan Brennan for generating strains used in this work and Diane King for high resolution Hi-C map processing. We also thank the Yeast Genome Resource Center at Osaka University and the many investigators who have deposited strains at this resource. This work was supported by an NSF EFRI CEE award EFMA-1830904 to S.G.J. Mochrie and M.C. King. M.P.B. was supported by NIH T32EB019941 and the NSF GRFP. H.Y. was support by the NSF Physics of Living Systems Student Research Network via PHYS 1522467.

Author Contributions J.F.W. and I.S. generated all yeast strains and raw movies for analysis of chromatin mobility. M.P.B. designed, implemented and plotted all loci tracking and trajectory analysis. H.Y. designed and performed the Gillespie and polymer simulations with input from I.S. and S.G.J.M. M.C.K. and S.G.J.M.

conceived of the project, coordinated the team strategy and provided experimental direction. The manuscript was written, edited and approved by all authors.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to SGJM (simon.mochrie@yale.edu) or MCK (megan.king@yale.edu).

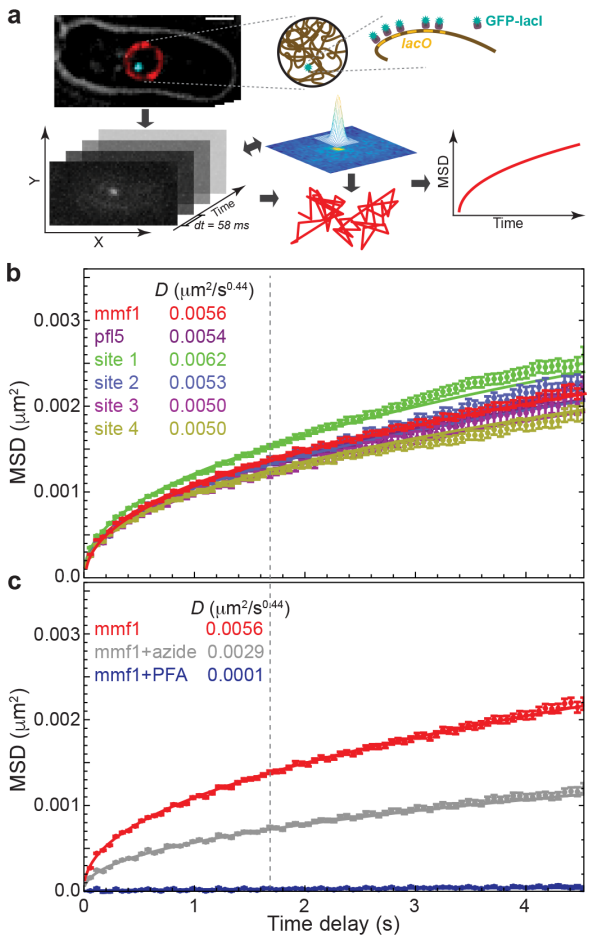

Figure 1: DNA loci display a universal sub-diffusive dynamics in live *S. pombe* cells on a seconds time scale. (a) Individual chromosomal loci are visualized in live cells using LacI-GFP that binds to a *lacO* array inserted at specific (*mmf1* and *pfl5*) or random (sites 1-4) genomic positions. The loci are tracked and population diffusive properties are quantified via the mean-square displacement (MSD) versus time. (b) Chromatin diffusivity is nearly identical regardless of specific genomic location as shown by the MSD vs time for six different genetic loci. (c) Chromatin dynamics is reduced upon cell treatment with sodium azide (*+azide*) to deplete ATP. As a control, cells fixed with formaldehyde (*+PFA*) were imaged and analyzed. In (b) and (c), lines are best-fit curves of Eq. 1 (see Methods) to the data with the exponent fixed at $\alpha = 0.44$ and the values of D corresponding to the best-fits: $D = 0.0056 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for *mmf1*, $D = 0.0054 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for *pfl5*, $D = 0.0062 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for site 1, $D = 0.0053 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for site 2, $D = 0.0050 \pm 0.0001\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for site 3, $D = 0.0050 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for site 4, $D = 0.0029 \pm 0.0002\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for *mmf1*+azide, and $D = 0.0001 \pm 0.0001\ \mu\mathrm{m}^2\,\mathrm{s}^{-0.44}$ for *mmf1*+PFA. The dashed line demarcates the time interval used for fitting. For all MSD plots in this paper, the last two terms in Eq. 1, which correspond to motion blur and static localization noise, respectively, were subtracted from both the model and the experimental MSDs, and all error bars represent standard errors of the mean.

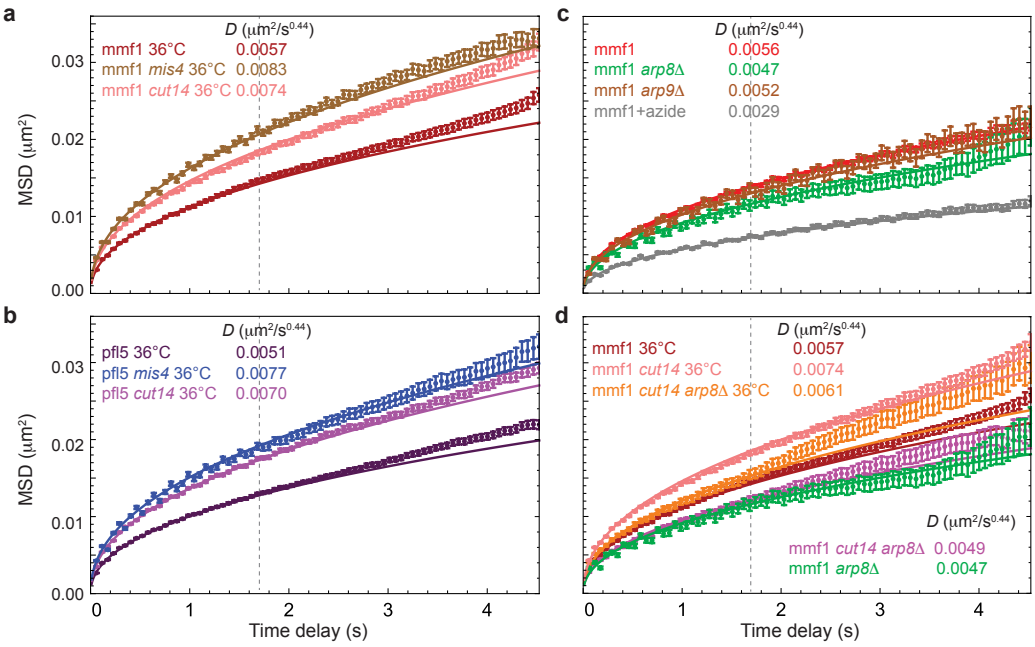

Figure 2: Loss of cohesin or condensin activity increases chromatin mobility while loss of the nucleosome remodeling complex protein, Arp8, decreases chromatin mobility. (a) MSDs of the *mmf1* locus in cells with temperature-sensitive alleles of the cohesin-loading factor Mis4 (*mis4-242*) ($D = 0.0083 \pm 0.0003\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$) or the SMC2 subunit of condensin (*cut14-208*) ($D = 0.0074 \pm 0.0002\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$) imaged at non-permissive temperature (36°C) to inactivate cohesin or condensin function, respectively, compared to wild-type mmf1 at 36°C ($D = 0.0057 \pm 0.0002\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$). (b) As in (a), but at the *pfl5* locus ($D = 0.0077 \pm 0.0004\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$ for *pfl5 mis4*, $D = 0.0070 \pm 0.0001\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$ for *pfl5 cut14*, and $D = 0.0051 \pm 0.0002\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$ for *pfl5* at 36°C). (c) MSDs of the *mmf1* locus in cells lacking Arp8, a component of the INO80 complex ($D = 0.0047 \pm 0.0003\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$), or Arp9, a component of SWI/SNF and RSC complexes ($D = 0.0052 \pm 0.0004\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$). For comparison, the MSD of the *mmf1* locus in wild type cells depleted of ATP (by treatment with sodium azide) is also shown ($D = 0.0029 \pm 0.0002\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$). (d) Loss of Arp8 dampens the elevated chromatin dynamics in cells with inactivated condensin complex. Shown are the MSDs of the *mmf1* locus in *arp8Δcut14^{ts}* cells imaged at permissive (30°C, $D = 0.0049 \pm 0.0002\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$) and non-permissive (36°C, $D = 0.0061 \pm 0.0003\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$) temperatures and of the *mmf1* locus in wild type cells at 36°C ($D = 0.0057 \pm 0.0001\ \mu\mathrm{m}^2\ \mathrm{s}^{-0.44}$). The dashed line demarcates the time interval used for fitting.

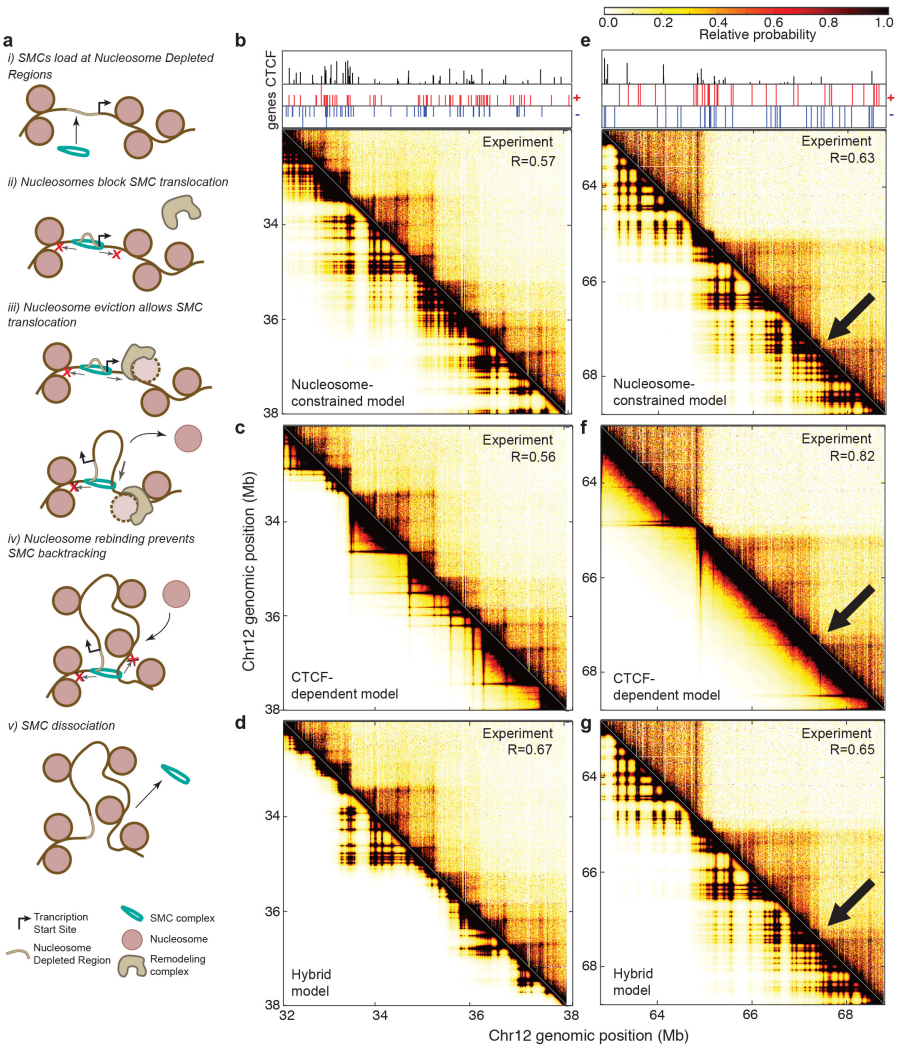

Figure 3: Nucleosome-constrained loop extrusion model yields contact maps qualitatively similar to those from murine Hi-C data. (a) The model hypothesizes that: (i) SMCs bind only at nucleosome depleted TSSs; (ii) nucleosomes block loop extrusion until (iii) nucleosome eviction allows one-sided SMC translocation and loop extrusion; (iv) nucleosome rebinding behind the SMC prevents back sliding and loop shrinkage; (v) eventually SMCs dissociate. (b-g) 10 kb-resolution experimental Hi-C maps (top right of each panel) of two 6 Mb regions of the mouse genome, compared to contact maps (bottom left) from simulations of (b and e) the nucleosome-constrained model, (c and f) the CTCF-dependent model; and (d and g) the hybrid model. Across the top of (b) and (e) are the locations and directionality of TSSs (leading and lagging strands marked as red and blue, respectively)[53] and the locations of CTCF accumulation, derived from ChIP-seq data[33]. The arrow indicates a TAD boundary predicted by the nucleosome-constrained model but not by the CTCF-dependent model. R is the Pearson cross-correlation between second moments of the simulated and experimental Hi-C maps. Maps are normalized as described in the Supplementary Information.

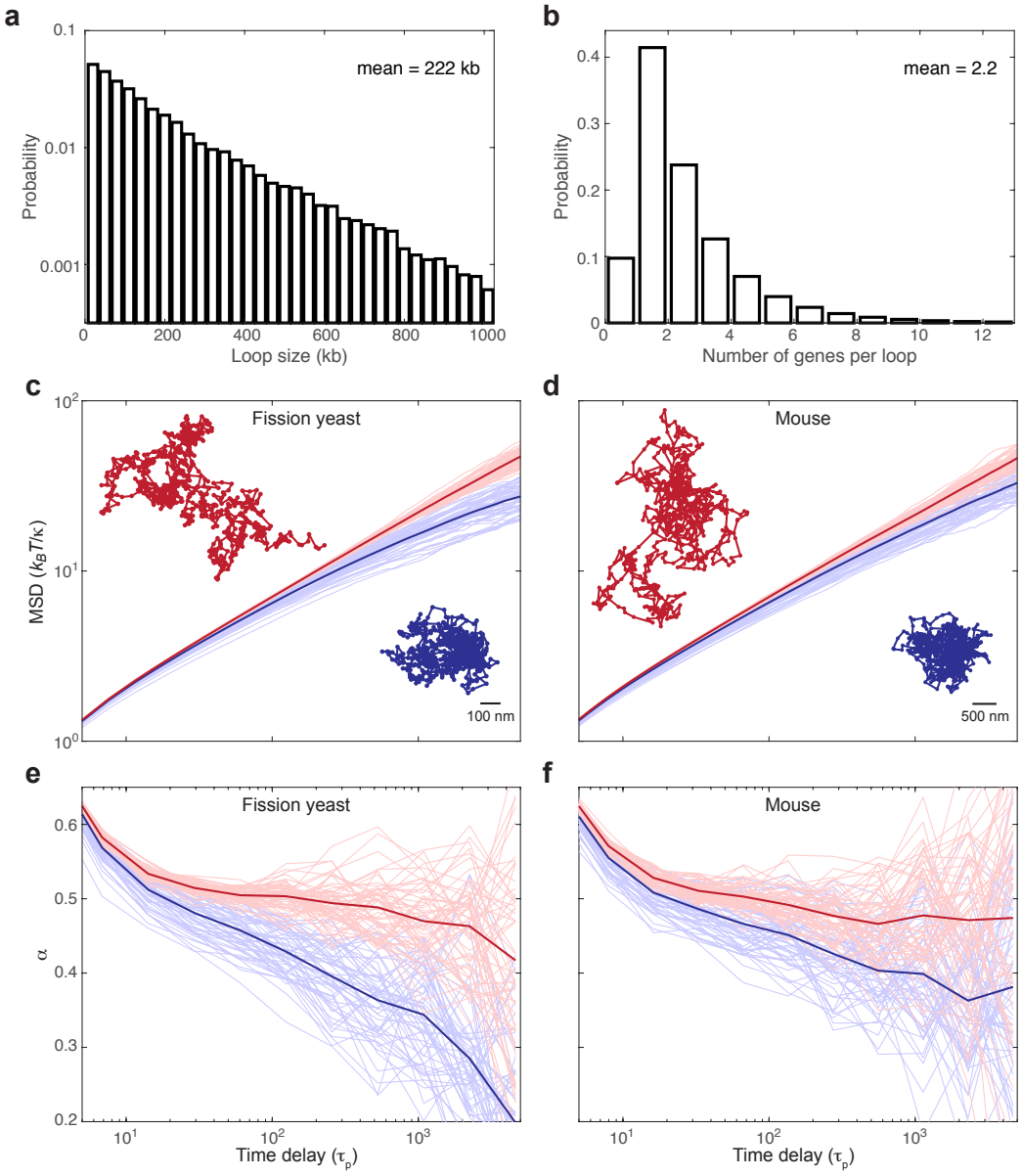

Figure 4: LEF activity constrains chromatin motion. (a and b) Nucleosome-constrained LEF model of the mouse genome (Figs. 3, S3, and S4) results in a wide distribution of chromatin loop sizes both when reported versus (a) base pairs and (b) numbers of genes. (c and d) Comparison between bead MSDs from Rouse-model (beads of friction coefficient ζ, connected by springs of spring constant κ) simulations with (blue) and without (red) time-dependent loops from the nucleosome-constrained model simulations of the yeast (c) and mouse (d) genomes. Inset are examples of the instantaneous polymer configurations from the simulations. (e) and (f) Effective exponent, α, versus time for the yeast and mouse models, respectively. In (c)-(f), thin lines correspond to individual beads, thick lines represent the average over all beads.