Transcription shapes 3D chromatin organization by interacting with loop-extruding cohesin complexes

Edward J. Banigan[1,*], Wen Tang[2,*], Aafke A. van den Berg[1,*], Roman R. Stocsits[2], Gordana Wutz[2], Hugo B. Brandão[3,4,5], Georg A. Busslinger[2,6], Jan-Michael Peters[2,†], and Leonid A. Mirny[1,†]

[1]Department of Physics and Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Research Institute of Molecular Pathology, Vienna BioCenter, 1030 Vienna, Austria
[3]Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA
[4]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[6]Research Center for Molecular Medicine of the Austrian Academy of Sciences and Department of Internal Medicine III, Division of Gastroenterology and Hepatology, Medical University of Vienna, Vienna 1090, Austria
*These authors contributed equally to this work
[†]Correspondence: peters@imp.ac.at (JMP) and leonid@mit.edu (LAM)

Summary

Cohesin organizes mammalian interphase chromosomes by reeling chromatin fibers into dynamic loops (Banigan and Mirny, 2020; Davidson et al., 2019; Kim et al., 2019; Yatskevich et al., 2019). “Loop extrusion” is obstructed when cohesin encounters a properly oriented CTCF protein (Busslinger et al., 2017; de Wit et al., 2015; Fudenberg et al., 2016; Nora et al., 2017; Sanborn et al., 2015; Wutz et al., 2017), and recent work indicates that other factors, such as the replicative helicase MCM (Dequeker et al., 2020), can also act as barriers to loop extrusion. It has been proposed that transcription relocalizes (Busslinger et al., 2017; Glynn et al., 2004; Lengronne et al., 2004) or interferes with cohesin (Heinz et al., 2018; Jeppsson et al., 2020; Valton et al., 2021; S. Zhang et al., 2021), and that active transcription start sites function as cohesin loading sites (Busslinger et al., 2017; Kagey et al., 2010; Zhu et al., 2021; Zuin et al., 2014), but how these effects, and transcription in general, shape chromatin is unknown. To determine whether transcription can modulate loop extrusion, we studied cells in which the primary extrusion barriers could be removed by CTCF depletion and cohesin’s residence time and abundance on chromatin could be increased by Wapl knockout. We found evidence that transcription directly interacts with loop extrusion through a novel “moving barrier” mechanism, but not by loading cohesin at active promoters. Hi-C experiments showed intricate, cohesin-dependent genomic contact patterns near actively transcribed genes, and in CTCF-Wapl double knockout (DKO) cells (Busslinger et al., 2017), genomic contacts were enriched between sites of transcription-driven cohesin localization (“cohesin islands”). Similar patterns also emerged in polymer simulations in which transcribing RNA polymerases (RNAPs)

acted as “moving barriers” by impeding, slowing, or pushing loop-extruding cohesins. The model predicts that cohesin does not load preferentially at promoters and instead accumulates at TSSs due to the barrier function of RNAPs. We tested this prediction by new ChIP-seq experiments, which revealed that the “cohesin loader” Nipbl (Ciosk et al., 2000) co-localizes with cohesin, but, unlike in previous reports (Busslinger et al., 2017; Kagey et al., 2010; Zhu et al., 2021; Zuin et al., 2014), Nipbl did not accumulate at active promoters. We propose that RNAP acts as a new type of barrier to loop extrusion that, unlike CTCF, is not stationary in its precise genomic position, but is itself dynamically translocating and relocalizes cohesin along DNA. In this way, loop extrusion could enable translocating RNAPs to maintain contacts with distal regulatory elements, allowing transcriptional activity to shape genomic functional organization.

Results

Depletion of Wapl and CTCF shows how transcription governs large-scale genome organization

Since dynamic positioning of cohesin governs global genome organization (Davidson and Peters, 2021), we investigated whether the relocalization of cohesin by transcription results in large-scale changes to chromatin contacts. We therefore performed high-throughput chromosome conformation capture (Hi-C) in DKO cells and compared it to observations in wild-type (WT), CTCF knockout (KO), Wapl KO, and Smc3 KO cells (**Extended Data Fig. 1a-b**).

The Hi-C experiments with DKO cells showed new genomic contact patterns generated by cohesin accumulated in “islands” between sites of convergent transcription. We observed new contacts between cohesin islands that appeared as Hi-C “dots” (island-island dots) that bridged distant genomic sites, consistent with the formation of cohesin-mediated chromatin loops. While cohesin frequently colocalizes with CTCF in WT cells (Busslinger et al., 2017; Parelho et al., 2008; Rao et al., 2014; Rubio et al., 2008; Wendt et al., 2008; Wutz et al., 2017), cohesin islands are not associated with CTCF sites (Busslinger et al., 2017) and contacts between CTCF sites disappear in DKO (**Fig. 1a**). Island-island dots are insulating (comparably to CTCF), and insulation is weakened in Smc3 KO cells (**Extended Data Fig. 1c-e**). Our findings indicate that cohesins that are relocalized to sites of convergent transcription continue to form large chromatin loops, consistent with ongoing active loop extrusion.

The new contact patterns are clearly distinguishable only in DKO cells (**Fig. 1a**), possibly because they have many more cohesins on chromatin that can be relocalized. CTCF KO abrogates cohesin accumulation at CTCF sites, increasing the quantity of mobile cohesins, and Wapl KO increases cohesin residence time, thus increasing the number of cohesin complexes on chromatin (Tedeschi et al., 2013) and allowing time for accumulation in islands. Nonetheless, insulation of genomic contacts, and to a lesser degree, cohesin accumulation, also emerge at sites of convergent transcription in WT cells (**Extended Data Fig. 1d**). Dots, cohesin accumulation, and insulation, at cohesin islands depend on transcription, as all are reduced in DKO by treatment with the transcription elongation inhibitor DRB (**Extended Data Fig. 1f-g**;

(Busslinger et al., 2017)). Thus, we hypothesize that active transcription may alter genome organization through its effects on loop extrusion by cohesin.

Cohesin dynamics and transcriptional activity spatially organize chromatin around genes

To directly study how the genome is organized by the interplay of transcription and extrusion, we computed average Hi-C contact maps and Scc1 (cohesin) ChIP-seq tracks centered on transcription start sites (TSSs) of genes, oriented and stratified by transcription activity and gene length (**Fig. 1b-c** and **Extended Data Fig. 2a-d**).

This revealed that individual genes are insulating, and active genes generate stronger insulation than inactive genes (**Fig. 1b** and **Extended Data Fig. 2a** and **c**). Contact enrichment and insulation correspond to cohesin accumulation at TSSs (**Fig. 1c** and **Extended Data Fig. 2a-d**). Insulation is abolished in Smc3 KO, while CTCF KO or lack of proximal CTCF only partially weakens insulation (**Fig. 1b** and **Extended Data Fig. 2e-f**). Insulation is also weakneed in Wapl KO and DKO cells (**Fig. 1b**), where increased residence time presumably allows loop-extruding cohesins to traverse the gene and bring regions upstream and downstream of the gene into contact. Thus, active genes are insulating boundaries in either the presence or absence of CTCF, and their effects on local genome organization depend on the dynamics of loop-extruding cohesins.

Near long, active genes, we observed intricate patterns of genomic contacts and cohesin accumulation, which were modulated by perturbations of cohesin dynamics. Across WT and all mutants (except Smc3 KO; **Extended Data Fig. 2e**), we observed five major features in the contact maps (**Fig. 1b** and **2a** and **Extended Data Fig. 2a**): 1) insulation (as described above), 2) lines (or “stripes”) of high contact frequency that extend upstream and downstream from the TSS, 3) lines of high contact frequency that extend upstream from the gene body that originate near the 3’ end, 4) high contact frequency within the gene and insulation of the gene, and 5) dots indicating high contact frequency between the 5’ and 3’ ends of the gene. These Hi-C features appeared with sharp ChIP-seq peaks of cohesin accumulation at the TSS, broader cohesin accumulation at the 3’ end of the gene, and a low background level of cohesin within the gene body (**Fig. 1c** and **Extended Data Fig. 2b**). The emergence of lines, dots, and insulation, along with the accumulation of cohesin at the ends of genes, suggests that TSSs and 3’ ends of active genes are barriers to loop extrusion.

Consistent with this interpretation, contact patterns are weaker for inactive genes and in cells treated with DRB, especially in DKO (**Extended Data Fig. 2c** and **g**). Furthermore, ChIP-seq shows sharp accumulation of RNAP II at the TSS and a smaller, broad accumulation near the 3’ end (**Extended Data Fig. 2h**), similar to cohesin ChIP-seq (**Fig. 1c** and **Extended Data Fig. 2b**). These observations suggest that RNAPs serve as barriers to loop extrusion, but raise the question of how the dynamic spatial organization of these barriers produces the observed genomic contact patterns.

The lines emanating from 3' ends of active genes suggest that 3' ends are effectively *asymmetric* extrusion barriers. This asymmetry is reminiscent of the lines that emanate from directionally oriented CTCF extrusion barriers (Barrington et al., 2019; Fudenberg et al., 2017, 2016; Vian et al., 2018). However, unlike with CTCF, cohesin accumulation is broad (**Fig. 1c** and **Extended Data Fig. 2b**). Together with broad, asymmetric RNAP accumulation at 3' ends (**Extended Data Fig. 2h**), these observations suggest that directional RNAP translocation in genes is central to both cohesin accumulation and asymmetric patterns of genomic contacts.

Our findings suggest that transcribing RNAPs are directionally translocating barriers to loop-extruding cohesins, which suggests a broad class of models for the dynamics and interactions of transcription and loop extrusion.

The moving barrier model for active loop extrusion can reproduce genome contact maps

The moving barrier model

We developed a model to determine how loop-extruding cohesins and their interactions with transcribing RNAPs can generate the major features of contact maps and cohesin accumulation (**Fig. 2b** and **Methods**). We modeled each cohesin as a two-sided loop-extruding complex that bridges two chromatin strands, which are independently and continuously extruded into a chromatin loop (Alipour and Marko, 2012; Fudenberg et al., 2016; Goloborodko et al., 2016b; Sanborn et al., 2015). We considered CTCF KO and DKO scenarios to focus on how transcription affects extrusion without complications from other strong extrusion barriers (*i.e.*, CTCF). In DKO simulations, cohesin residence time was increased tenfold and linear density was increased twofold due to Wapl depletion, as suggested by previous experiments (Gassler et al., 2017; Haarhuis et al., 2017; Tedeschi et al., 2013; Wutz et al., 2017) and simulations (Gassler et al., 2017; Nuebler et al., 2018).

For extrusion-transcription interactions, we extended the "moving barrier" model for interactions between bacterial condensins and RNAPs (see **Methods**; (Brandão et al., 2019)). RNAPs load at TSSs, transiently pause, and slowly translocate through the gene (~0.01-0.1 kb/s (Darzacq et al., 2007; Wang et al., 1998)) and interact with more rapidly translocating loop-extruding cohesins (0.1-1 kb/s (Banigan and Mirny, 2020; Davidson et al., 2019; Golfier et al., 2020; Kim et al., 2019)) (**Fig. 2c**). When RNAP encounters a cohesin in a head-on collision, it pushes this cohesin along the chromatin fiber in the direction of transcription, shrinking the loop from one side, while the other side of the loop continues to grow at its normal speed (**Fig. 2c**). RNAP pushing cohesin is consistent with the large difference in the stall forces of cohesin (0.1-1 pN; (Golfier et al., 2020; Kim et al., 2019)) and RNAP (~10 pN; (Wang et al., 1998)). Alternatively, when an extruding cohesin progressing toward the 3' end encounters RNAP, extrusion of this side of the loop continues more slowly behind the slower RNAP, while continuing as normal on the other side of the loop (**Fig. 2c**). In both types of collision, cohesin may stochastically bypass RNAP after a characteristic waiting time, similar to *in vitro* observations of cohesin bypassing obstacles on DNA (Pradhan et al., 2021) and predictions for bacterial condensins *in vivo*

(Brandão et al., 2019). Thus, in this model, RNAP is a weakly permeable moving barrier to loop extrusion.

We considered several models for cohesin loading and extrusion dynamics (**Fig. 2d**). Cohesin was loaded either with uniform probability on any genomic site or preferentially at promoters. The latter was suggested by ChIP-seq experiments showing cohesin and Scc2/Nipbl enrichment at TSSs (Busslinger et al., 2017; Kagey et al., 2010; Zhu et al., 2021; Zuin et al., 2014). We also considered several modes of cohesin extrusion: 1) diffusive growth or shrinking on each side of the extruded loop (Brackley et al., 2017; Yamamoto and Schiessel, 2017), similar to the earlier hypothesis that RNAPs push passive cohesins to sites of convergent transcription (Glynn et al., 2004; Lengronne et al., 2004) and *in vitro* observations (Davidson et al., 2016), or 2) active, directed loop extrusion of each of the two chromatin strands, as recently observed on DNA *in vitro* (Davidson et al., 2019; Golfier et al., 2020; Kim et al., 2019) and suggested by active extrusion models (Banigan et al., 2020; Fudenberg et al., 2016; Sanborn et al., 2015).

Loop extrusion with moving barriers generates experimentally observed genomic contact patterns

Using 3D polymer simulations coupled to stochastic 1D transcription and extrusion dynamics (see **Methods**), we simulated chromosome organization by the moving barrier mechanism with different cohesin loading scenarios, loop extrusion activities, and cohesin-RNAP bypassing times.

In models with diffusively extruding cohesins, active genes alter cohesin accumulation patterns and the spatial organization of the genome, but the simulations lacked prominent features observed via Hi-C and ChIP-seq, irrespective of the loading scenario, diffusion coefficient, and cohesin-RNAP bypassing time (**Fig. 2d** and **Extended Data Fig. 3)**. Cohesin accumulation, where it occurred, was weak and broad because diffusive cohesins do not remain localized after encountering an extrusion barrier. This resulted in weak, poorly defined features in simulation contact maps. We conclude that diffusively extruding cohesins do not reproduce the experimental observations around active genes, even when they are pushed by RNAPs.

Simulations with cohesins that actively extrude loops by directional translocation produced genome contact maps and cohesin accumulation patterns with well defined features (**Fig. 2d** and **Extended Data Fig. 4**). Active extrusion with a low, but nonzero, rate of cohesin- RNAP bypassing (~1 event per cohesin lifetime, *i.e.*, cohesin follows or is pushed by RNAP for ~100 s) gave the best agreement with experiments. Both with and without loading to the TSS, cohesin sharply accumulated at TSSs and more broadly accumulated at 3’ ends of genes, similar to experimental observations. TSS accumulation occurred because RNAPs that occupied the TSS prior to initiation acted as barriers to extrusion. Cohesin accumulated near 3’ by two mechanisms: 1) RNAPs paused, but still bound, at the 3’ end after transcription termination act as barriers and 2) translocating RNAPs that encounter extruding cohesins head-on push the cohesins back toward 3’ ends and slow down extrusion by trailing cohesins (**Fig. 2c**). Consistently with these mechanisms, cohesin accumulation at 3’ gene ends was enhanced in

DKO simulations due to their longer residence time. In both CTCF KO and DKO simulations, cohesin accumulation resulted in insulation (feature 1), lines emanating from the TSS (feature 2), and lines running upstream from 3' ends (feature 3). Consistent with the experimental observations, Hi-C lines from 3' were particularly thick in DKO simulations. We also observed enrichment of contacts within the gene and insulation of the gene (feature 4), as well as dots for contacts between gene ends (feature 5). Nonetheless, the simple moving barrier model with active, directed extrusion reproduced the major features of active gene organization surprisingly well.

Our model also allows us to differentiate between two previously proposed (Fudenberg et al., 2016; Gullerova and Proudfoot, 2008; Kagey et al., 2010) modes of cohesin loading. Extrusion with uniform cohesin loading reproduced the experimental Hi-C maps better than models with a strong preference for loading at the promoter (**Fig. 2d**). In contrast, simulations with targeted loading had an additional strong feature that is not present in the experiments: diagonal lines that emanate from the TSS, perpendicular to the main diagonal. These lines of enriched contacts formed because cohesins loaded at the TSS brought chromatin on both sides of the TSS together as they progressively extruded loops. This observation suggests that preferential loading of cohesin at all promoters is inconsistent with genome organization around active genes.

The moving barrier model also makes several testable predictions about genome contact patterns. First, it reproduces the experimental observations of cohesin islands and island-island dots at sites of convergent transcription in DKO cells (**Extended Data Fig. 5**). The simulations additionally predict that the TSSs of divergently oriented genes form contacts (dots) in both CTCF KO and DKO simulations (**Extended Data Fig. 5**). This is consistent with the idea that TSSs occupied by RNAPs are barriers to loop extrusion that can accumulate cohesin. Since cohesin also accumulates at the 3' ends (**Figs. 1c** and **2d**), genomic contacts should be enriched between two consecutive ends of active genes, regardless of orientation. We test this prediction below.

Altogether, we find that genome organization near transcriptionally active genes can be generated by active, directed loop extrusion with RNAPs acting as weakly permeable and mobile barriers to extrusion. The model suggests that both ends of active genes generically serve as weak barriers to extrusion, and thus, should form contacts with nearby ends of other active genes and CTCF sites (**Fig. 3a**). The simulations further suggest that cohesin is uniformly loaded on chromatin, without a strong preference for loading at the promoter (**Fig. 3b**). We next tested these two predictions by new ChIP-seq experiments and analysis of Hi-C data.

Transcription generates genomic contacts between gene ends

To test the prediction that nearby ends of active genes are barriers to extrusion with enriched genomic contacts, we computed average contact maps for contacts between gene ends for pairs of active genes of various orientations. Across WT and all mutants with cohesin, we observed dots of high contact frequency between proximal TSSs (**Fig. 3a**), and contacts between TSSs of adjacent active genes can be observed for all gene pairs regardless of their

orientations (**Extended Data Fig. 6a**). As predicted, 3' ends of genes can also act as extrusion barriers that enhance contacts between genes (**Extended Data Fig. 6a**). In each case, contacts are weakened when cohesin residence time is increased by Wapl depletion, presumably because cohesin may translocate through the gene. Contact enrichment depends on transcription, as dots are not observed for inactive genes (**Extended Data Fig. 6b**). These results demonstrate that cohesins generate specific genomic contacts in response to the cell's transcriptional activity.

Cohesin is not preferentially loaded at promoters

It is widely held that cohesin is loaded preferentially at promoters of active genes, but our moving barrier model predicts that agreement with the Hi-C data requires largely uniform cohesin loading along the chromatin fiber (**Figs. 2d** and **3b**). The hypothesis that cohesin loading occurs at TSSs is based on the notion that cohesin is loaded onto DNA by NIPBL (Ciosk et al., 2000) and that in ChIP-seq experiments NIPBL antibodies preferentially detect TSSs (Busslinger et al., 2017; Kagey et al., 2010; van den Berg et al., 2017; Zhu et al., 2021; Zuin et al., 2014). However, active TSSs have been identified as 'hyper-ChIP-able' regions to which some antibodies can bind in the absence of their antigen (Jain et al., 2015; Schwartz et al., 2005; Teytelman et al., 2013), and negative controls have not been reported for NIPBL ChIP-seq experiments. We therefore re-analyzed the enrichment and localization of NIPBL throughout the genome.

First, we performed ChIP-seq experiments in HeLa cells with the antibody 133M, which has been used previously (Zuin et al., 2014). It identified 6,738 peaks, of which 5,209 were located at TSSs (**Fig. 4a-b**), similar to previous observations (Zuin et al., 2014). However, after NIPBL depletion by RNAi (**Extended Data Fig. 7a**), most ChIP-seq peaks at TSSs remained (73%; **Fig. 4a-c** and **Extended Data Fig. 7b**), even though NIPBL's binding partner MAU2 could not be detected in these cells (NIPBL degradation leads to MAU2 depletion, which we analyzed because immunoblotting detects MAU2 more reliably than the 316 kDa NIPBL protein (Watrin et al., 2006)). This suggested that ChIP-seq signals at TSSs obtained with 133M antibodies did not depend on NIPBL. Importantly, the 133M antibody also identified smaller non-TSS peaks that were reduced in NIPBL-depleted cells (**Fig. 4a**), suggesting that these might be *bona fide* NIPBL sites.

To identify sites of NIPBL and MAU2 enrichment, we generated HeLa cell lines in which all NIPBL or MAU2 alleles were modified with a tag (hemagglutinin; HA) (**Fig. 4d**) and a protein domain (FK506 binding protein 12-F36V; $FKBP12^{F36V}$). The HA tag can be specifically recognized in ChIP-seq experiments and FKBP12 can be used to induce degradation of the resulting fusion proteins (Nabet et al., 2018) for critical negative control experiments (**Extended Data Fig. 7c)**. The resulting HA-$FKBP12^{F36V}$-NIPBL and MAU2-$FKBP12^{F36V}$-HA fusion proteins were functional since they supported formation of vermicelli, axial chromosomal sites at which cohesin accumulates in Wapl-depleted cells (Tedeschi et al., 2013) (**Extended Data Fig. 7d-f**).

HA-NIPBL and MAU2-HA ChIP-seq experiments identified small numbers of peaks, most of which disappeared upon dTAG-induced degradation (**Fig. 4d-e** and **Extended Data Fig. 7g**). Of

these 'high-confidence' NIPBL-MAU2 peaks, which we used as a reference set (see **Methods**), 89.5% overlapped with cohesin peaks, but only 13.8% were located at TSSs (**Fig. 4f**). Moreover, nearly all the latter NIPBL-MAU2 peaks were located at TSSs at which cohesin was also enriched (**Fig. 4e-f**). Interestingly, some of the NIPBL-MAU2 sites overlapped with the small non-TSS ChIP-seq peaks obtained with the 133M antibody, confirming that these are *bona fide* NIPBL sites (**Fig. 4a** and **c** and **Extended Data Fig. 7b**).

Since NIPBL and MAU2 co-localize with cohesin, we hypothesized that NIPBL-MAU2 complexes might be recruited to these sites by binding to cohesin. To test this possibility, we analyzed whether HA-NIPBL and MAU2-HA ChIP-seq peaks depend on cohesin by depleting SCC1 via RNAi (**Fig. 4g-h** and **Extended Data Fig. 7h**). Both HA-NIPBL and MAU2-HA ChIP-seq peaks were greatly reduced or undetectable after depletion of SCC1.

These experiments suggest that NIPBL-MAU2 complexes are not enriched at TSSs unless these are also occupied by cohesin; rather, NIPBL-MAU2 colocalizes with a subset of cohesin complexes on chromatin. Together with our simulations (**Fig. 3b**), these findings suggest that cohesin complexes are not preferentially loaded onto chromatin at TSSs. Instead, cohesin accumulation at TSSs could occur due to the barrier function of TSSs, and in turn, NIPBL could colocalize with cohesin at sites where loop extrusion stalls.

Discussion

It has been hypothesized that RNAPs push cohesin complexes that have entrapped DNA within their ring structures, displacing cohesins from their apparent loading sites at gene promoters (Kagey et al., 2010) to the 3' ends of genes (Busslinger et al., 2017; Glynn et al., 2004; Lengronne et al., 2004; Ocampo-Hafalla et al., 2016). Indeed, previous single-molecule experiments demonstrated that a transcribing RNAP could push a passively diffusing cohesin complex along DNA *in vitro* (Davidson et al., 2016). However, it is now known that cohesin can translocate by actively extruding DNA loops (Davidson et al., 2019; Golfier et al., 2020; Kim et al., 2019). Cohesin can do so without topologically encircling DNA (Davidson et al., 2019), and furthermore, it can bypass large obstacles on DNA (Pradhan et al., 2021). The mechanism introduced here can reconcile these new developments with older observations of the effects of RNAP on cohesin.

Our experiments and simulations indicate that RNAP acts as a "moving barrier" to loop-extruding cohesins. RNAP in a head-on collision with cohesin can push cohesin toward a gene's 3' end as cohesin continues to extrude at the other end of its loop. Alternatively, RNAP can slow extrusion by cohesin trailing the RNAP, while cohesin can continue to rapidly extrude the other side of the loop (**Fig. 2a-b**). RNAPs accumulated at TSSs and TTSs also act as extrusion boundaries that enrich contacts between nearby gene ends (**Fig. 3a** and **Extended Data Fig. 6**). These findings generalize the bacterial moving barrier model (Brandão et al., 2019) to eukaryotic cells and provide a detailed account of how transcription interacts with extrusion (Jeppsson et al., 2020; Rivosecchi et al., 2020; Valton et al., 2021; S. Zhang et al.,

2021) to locally modulate genome organization by stopping, hindering, and relocalizing cohesins.

Through moving barrier interactions, transcription can locally shape genome architecture. Accordingly, normal transcriptional responses or excess readthrough can disrupt cohesin-mediated looping within genes and near 3' ends in human cells (Heinz et al., 2018). Through this mechanism, transcription can also *enrich* contacts, as it does between genes in mouse cells (**Fig. 3a** and (Hsieh et al., 2020)) or between sites of convergent transcription in yeast (Dauban et al., 2020; Jeppsson et al., 2020). Furthermore, in mammalian cells, TSSs and 3' ends of active genes insulate genomic contacts (Hsieh et al., 2020; Valton et al., 2021; You et al., 2021; Zhang et al., 2020; H. Zhang et al., 2021) by acting as extrusion boundaries (**Figs. 1b-c**, **2,** and **3a** and **Extended Data Figs. 6** and **8**; (Valton et al., 2021)).

Functionally, RNAP barriers could regulate genes through their effects on loop extrusion. Pausing of extrusion at TSSs would allow cohesin to linearly scan chromatin for proximal enhancers near other boundaries, such as CTCFs, to bring them into contact with the TSS (Blackwood and Kadonaga, 1998). Subsequently, cohesin may track the transcribing RNAP through the gene, maintaining continuous contact between the enhancer and the transcription complex, as previously observed (Lee et al., 2015). Further, cohesin-mediated contacts between nearby active promoters (**Fig. 3a** and (Hsieh et al., 2020)) may mediate mutual regulation, possibly facilitating the spreading of histone marks and transcription factors. Linear scanning by cohesin in gene regulation would be consistent with cohesin's proposed role in other contexts, including V(D)J recombination (Hill et al., 2020; Lin et al., 2018; Peters, 2021; Wood and Tonegawa, 1983; Zhang et al., 2019), alternative protocadherin choice (Canzio et al., 2019), and double strand break repair (Arnould et al., 2021; Yang et al., 2021). Furthermore, since moving RNAP extrusion barriers are transcription-dependent, chromosomal interactions could be rapidly modulated in a locus-specific manner. For instance, histone marks or transcription factors could regulate genomic contacts by locally activating transcription, thus modulating functional interactions in *cis* via extrusion. Similarly, even non-protein-coding genes could have this effect through, for example, transcription of non-coding RNAs and eRNAs (Joung et al., 2017). Therefore, in contrast to static barriers like CTCFs, moving RNAP barriers can dynamically regulate loop extrusion and functional interactions.

Our modeling predicts that cohesin is not preferentially loaded at promoters (**Figs. 2d** and **3b**), in contrast to previous reports (Busslinger et al., 2017; Kagey et al., 2010; Zhu et al., 2021; Zuin et al., 2014). In agreement with this prediction, our new ChIP-seq experiments (**Fig. 4**) show that the enrichment of the "cohesin loader" NIPBL at TSSs may have been, at least in part, an artifact, possibly because active TSSs are 'hyper-chippable,' especially when the antibody has a poor specificity (Jain et al., 2015; Schwartz et al., 2005; Teytelman et al., 2013). Furthermore, we found that NIPBL occupancy depends on the presence of cohesin (**Fig. 4g-h**), consistent with the requirement of NIPBL for *in vitro* loop extrusion (Davidson et al., 2019; Kim et al., 2019) and *in vivo* loop lengthening (Haarhuis et al., 2017; Rhodes et al., 2017) (**Extended Data Fig. 7e-f**) by cohesin. This further suggests that Nipbl may serve as an extrusion processivity factor

for cohesin rather than only as a loading factor (Davidson et al., 2019; Davidson and Peters, 2021).

Even though RNAP relocalizes and slows cohesin in our model, cohesin can bypass RNAP approximately once per cohesin lifetime (~100 s). This is consistent with experiments indicating that cohesin does not topologically enclose DNA while it extrudes loops (Davidson et al., 2019; Pradhan et al., 2021). However, our model's bypassing time is 10-fold longer than predicted for bacterial condensins bypassing RNAPs (Brandão et al., 2019) and measured for SMC complexes bypassing obstacles on DNA *in vitro* (Pradhan et al., 2021). This discrepancy suggests differences between loop extrusion on nucleosomal fibers versus DNA or that cohesin may have some affinity for RNAP. The former could be due to steric interactions imposed by nucleosomes and/or large nascent RNA molecules trailing the RNAP, while the latter could facilitate the linear scanning processes described above. However, much like the mechanism of extrusion itself, the molecular mechanisms of interactions with RNAP and bypassing remains unclear.

Our results indicate that RNAP belongs to a growing list of elements that dynamically structure the genome by acting as barriers to loop extrusion. However, while boundaries such as CTCF sites are stationary (Busslinger et al., 2017; Fudenberg et al., 2016; Nora et al., 2017; Sanborn et al., 2015; Wutz et al., 2017), RNAPs are mobile and can be dynamically controlled by transcriptional regulators**.** Together with emerging evidence that extrusion might also be obstructed by other mobile complexes such as replication machinery (Dequeker et al., 2020; Jeppsson et al., 2020)**,** this suggests that in addition to *structural* functions, cohesin has important *dynamic* functions in the spatiotemporal organization of the genome.

Methods

Generation of HeLa cell lines

All cell lines used in this study were generated by homology-directed repair using CRISPR Cas9 (D10A) paired nickase (Ran et al., 2013). Based on the cell line SCC1-GFP (Wutz et al., 2017), we introduced a Halo-AID tag to the N-terminus of WAPL, generating Halo-AID-WAPL/SCC1-GFP. Subsequently, Tir1 expression was introduced by transducing a homozygous cell clone with lentiviruses using pRRL containing the constitutive promoter from spleen focus forming virus (SFFV) followed by Oryza sativa Tir1- 3xMyc-T2A-Puro (Wutz et al., 2017). Based on this cell line, HA-FKBP12F36V-NIPBL/Halo-AID-WAPL/SCC1-GFP and MAU2-FKBP12 F36V-HA/Halo-AID-WAPL/SCC1-GFP were created by introducing a HA-FKBP12F36V tag to the N terminus of NIPBL or a FKBP12F36V-HA tag to the C-terminus of MAU2 individually. As a control cell line, HA-FKBP12 F36V-WAPL was generated using the same gRNAs as described above but with a different repair template. The following gRNAs were used for WAPL: CACCGCTAAGGGTAGTCCGTTTGT and CACCGTGGGGAGAGACCACATTTA; NIPBL: CACCGTCCCCGCAAGAGTAGTAAT and CACCGGTCTCACAGACCGTAAGTT; MAU2: CACCGCTTGGAACTGCACGGGGGG and CACCGCTCCTGTGAGGCCTTGATG. Clones were selected after verification of homozygous

integration by PCR of genomic DNA (primers used for WAPL: TGATTTTTCATTCCTTAGGCCCTTG and TACAAGTTGATACTGGCCCCAA; NIPBL: GCAGTGCTTGTCGAGGTTGAT and GCTCAGCCTCAATAGGTACCAACA; MAU2: ATGTCGGTACAGCTGTGGTC and GTGCCACGCACTCTAAGCTA).

Antibodies and Reagents

Antibodies against MAU2 (Peters laboratory A974) were reported previously (Watrin et al., 2006) and rabbit anti-NIPBL antibodies (Peters laboratory A870 or 133M) were used in a previous publication (Zuin et al., 2014). The following commercial antibodies were used: HA (Abcam, ab9110 for ChIP; BioLegend, 901501 for Immunofluorescence and Western Blot), GFP (Abcam, ab13970 for Immunofluorescence; Abcam, ab290 for ChIP), RNA polymerase II S2 (Abcam, ab5095), SCC1 (Millipore, 05-908), α-tubulin (Sigma-Aldrich, T5168), and histone H3 (Cell Signaling, 9715L). The following secondary antibodies were used: goat anti-chicken IgG Alexa-Fluor-488, anti-mouse IgG Alexa-Fluor-568 (Molecular Probes) for Immunofluorescence and anti-rabbit or mouse Ig, and HRP-linked whole antibody (GE Healthcare) for Western Blot. Auxin were purchased from Sigma (I5148) and dTAG-7 was a gift from Georg Winter (CeMM, Austria).

Whole cell extract

Cells were re-suspended in modified RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Na-deoxycholate and 0.1% SDS), which additionally contained pepstatin, leupeptin and chymostatin (10 µg/ml each), and PMSF (1 mM). The protein concentration was determined with Bradford assay (Bio-Rad, #5000006). SDS-PAGE and Western Blot were applied to detect individual protein with specific antibodies.

Chromatin fractionation

Cells were re-suspended with extraction buffer (25 mM Tris pH 7.5, 100 mM NaCl, 5mM MgCl2, 0.2% NP-40, 10% glycerol) supplemented with an EDTA-free protease inhibitor tablet (Roche, 05056489001) and PMSF (1 mM). The chromatin pellet was obtained by centrifugation and re-suspended with extraction buffer containing Benzonase (Merck, 70664) on ice for 30 min, and the protein concentration was determined with Bradford assay.

Immunofluorescence microscopy

Cells were fixed with 4% formaldehyde for 20 min at room temperature and permeabilized with 0.2% Triton X-100 in PBS for 10 min. After blocking with 3% BSA in PBS-T 0.01% for 30 min, the cells were incubated with the primary antibodies for 1 h and subsequently incubated with the secondary antibodies for 1 h at room temperature. After counterstaining the DNA by 10 min incubation with DAPI, the coverslips were mounted with ProLong Gold (Thermo Fisher) before imaging.

RNA interference

For RNAi experiments, the cells were transfected as described previously (Wutz et al., 2017). Briefly, the cells were transfected by incubating 100 nM duplex siRNA with RNAi-MAX transfection reagent in antibiotic-free growth medium. After 48 h of RNAi treatment, cells were harvested for experiments. The following target sequences of siRNAs (Ambion) were used: SCC1 (5′-GGUGAAAAUGGCAUUACGGtt -3′), NIPBL (5′-GCAUCGGUAUCAAGUCCCAtt-3′).

Calibrated ChIP followed by next-generation sequencing

ChIP was performed as described previously (Wendt et al., 2008). Before crosslinking, 10 million HeLa cells were spiked in with 5% MEFs cells (except for RNA polymerase II S2 ChIP, which used only 10 million MEFs cells, without spike-in). Cells were crosslinked with 1% formaldehyde at room temperature for 10 min and subsequently quenched with 125 mM glycine for 5 min. Cells were washed with PBS and then lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS, protease inhibitors) on ice for 10 min. DNA was sonicated by 6 cycles (30 sec on/off) using Biorupter. 10 volumes of dilution buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 150 mM NaCl, 1 mM PMSF) was added to the lysate, and followed by pre-clear with 100 µl Affi-Prep Protein A beads at 4°C. Immunoprecipitation was performed with rabbit IgG or antibody overnight, followed by 3-hour incubation with Affi-Prep Protein A beads. Beads were washed twice with Wash buffer 1 (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 150 mM NaCl, 0.1% SDS, 1 mM PMSF), twice with Wash buffer 2 (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 500 mM NaCl, 0.1% SDS, 1 mM PMSF), twice with Wash buffer 3 (10 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate), twice with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0), and eluted twice with 200 µl elution buffer (25 mM Tris-HCl pH 7.5, 5 mM EDTA pH 8.0, 0.5% SDS) by shaking at 65°C for 20 min. The eluates were treated with RNase-A at 37°C for 1 hour and proteinase K at 65°C overnight. Addition of 1 µl glycogen (20 mg/ml) and 1/10th volume sodium acetate (3 M, pH 5.2) was followed by extraction with phenol/chloroform/isoamyl alcohol (25:24:1), precipitation with ethanol. DNA was re-suspended in 100 l of H2O, and ChIP efficiency was quantified by quantitative PCR (qPCR). The DNA samples were submitted for library preparation and Illumina deep sequencing to the Campus Science Support Facility.

ChIP-seq peak calling and calculation of peak overlaps

Illumina short read sequencing data were aligned against a fusion genome template consisting of merged hg19 and mm9 assemblies. Alignments were filtered for reads which mapped uniquely to the human genome allowing up to two mismatches. The uniquely mappable mouse genome fraction of the spike-in was used for estimation of material loss and calibration by random subsampling of read counts after deduplication.

Peak callings with MACS software versions 1.4 and 2 were applied on full, merged, filtered, and deduplicated replicate data sets of HA-tagged samples with corresponding depleted conditions (+dTAG) as input controls (Zhang et al., 2008). No model-building was performed and output for

visualization as bedGraph was normalized by signal strength per million reads for further comparability.

Reference peak set was chosen by searching for these positions in both NIPBL and MAU2 data which could be detected by peak calling via MACS 1.4 using a p-value threshold of 1e-10 in NIPBL data as well as with a p-value threshold of 1e-5 in MAU2 data, and which were confirmed also by MACS2 with default stringencies in both sets. Further, the reference was extended to also include highly scoring NIPBL sites that did not have corresponding signals passing the threshold in the MAU2 fraction (which generally shows weaker signals).

All genomic overlaps as well as area-proportional threefold Venn diagrams have been calculated using multovl version 1.3 (Aszódi, 2012) and were drawn with eulerAPE (Micallef and Rodgers, 2014). Since occasionally more than one site from one dataset overlaps with a single site in a second dataset, resulting coordinates of such an overlap contribute to one single entry - a so-called union. Consequently, the overall site counts drop slightly if displayed in union overlaps.

Heatmap plots have been made using deeptools (Ramírez et al., 2016). bedGraph output from peak callings was converted to bigwig input for processing as heatmaps.

RNA polymerase II ChIP-Seq in mouse cells was performed without calibration through spike-in, and short-read sequencing data were mapped against the mm9 template using bowtie2. Only uniquely mappable de-duplicated reads were used further, allowing up to two mismatches.

Hi-C protocol for MEFs

Hi-C was performed as described previously (Silva et al., 2020). Briefly, 30x106 cells were cross-linked in 2% formaldehyde for 10 minutes and quenched with ice-cold glycine (0.125 M final concentration). Cells were snap-frozen and stored at -80°C before cell lysis. Cells were lysed for 30 min in ice cold lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA, and 0.2% NP-40) in the presence of protease inhibitors. Chromatin was solubilized in 0.6% SDS at 37°C for 2h minutes, quenched by 3.3% Triton X-100. Chromatin was digested with 400 units of HindIII overnight at 37°C. Fill-in of digested overhangs by DNA polymerase I, large Klenow fragment in the presence of 250 nM biotin-14-dATP for 90min was performed prior to 1% SDS based enzyme inactivation and dilute ligation with T4DNA ligase for 4 hours at 16°C. Cross-links of ligated chromatin were reversed overnight by 1% proteinase K incubation at 65°C. DNA was isolated with 1:1 phenol:chloroform, followed by 30 minutes of RNase A incubation. Biotin was removed from unligated ends by incubation with 15 units of T4 DNA polymerase. DNA was sheared using an E220 evolution sonicator (Covaris, E220) and size selected to 150-350 bps by using AMPure XP beads. After end repair in a mixture of T4 polynucleotide kinase, T4 DNA polymerase and DNA polymerase I, large (Klenow) fragment at room temperature for 30 minutes, dATP was added to blunted ends polymerase I, large fragment (Klenow 3’ → 5’ exo-) at 37°C for 30 minutes. Biotinylated DNA was collected by incubation in the presence of 10 µl of streptavidin coated myOne C1 beads and Illumina paired-end adapters were added by ligation with T4 DNA ligase for 2 hours at room

temperature. A PCR titration (primers PE1.0 and PE2.0) was performed prior to a production PCR to determine the minimal number of PCR cycles needed to generate a Hi-C library. Primers were separated from the library using AMPure XP size selection prior to 50 bp paired-end sequencing (HiSeqv4, Illumina).

Scc1 and CTCF ChIP-seq

ChIP-seq data for Scc1 and CTCF ((Busslinger et al., 2017); GSE76303) were mapped using the mm9 assembly and normalized by input (logfe) using a workflow based on the ENCODE pipeline (https://github.com/ENCODE-DCC/chip-seq-pipeline). CTCF sites were selected by overlapping ChIP-seq peaks (Busslinger et al., 2017) with CTCF motifs, which were obtained by scanning the position weight matrix (PWM) of the M1 CTCF motif (Schmidt et al., 2012) using FIMO (Grant et al., 2011). Cohesin islands were identified as previously (Busslinger et al., 2017).

Hi-C mapping and analysis

Hi-C data were mapped using the mm9 genome assembly and distiller pipeline (https://github.com/mirnylab/distiller-nf; version 0.0.3 for all datasets except for DRB and DRB release, which used version 0.3.1). The mapped data were then converted to cooler files (Abdennur and Mirny, 2019) and balanced by iterative correction as described previously (Imakaev et al., 2012). Contact probability scalings, $P_c(s)$, and insulation were computed using cooltools (https://github.com/mirnylab/cooltools; (Abdennur and Mirny, 2019)). Pile ups were computed from Python scripts by collecting snippets of maps (“observed”) around sites of interest (such as ends of genes, CTCF sites, or island-island contacts), normalizing each diagonal by the value of the scaling (“expected”) at that diagonal, and averaging “observed-over-expected” values across the collected snippets (https://github.com/mirnylab/moving-barriers-paper). To select Hi-C regions around genes based on transcription levels and gene length, we combined gene annotations for genes with a known transcription status from GENCODE (https://www.gencodegenes.org/) with previously reported GRO-seq for MEFs ((Busslinger et al., 2017); GEO accession number GSE76303). Unless noted, we considered only genes isolated from other genes by at least 10 kb. Dot strengths are computed by summing observed-over-expected within a 50 kb of the dot and dividing by the background, taken to be the mean number of contacts in two windows of the same size centered 150 kb upstream and downstream of the dot. For analyzing genomic loci, such as genes, that are “away” from CTCF sites, unless noted, we excluded sites within 5 kb of the top 50%, by motif score, of identified CTCF sites.

Polymer simulations with loop extrusion

Polymer simulations with loop extrusion were performed using OpenMM (Eastman et al., 2017; Eastman and Pande, 2010) and the openmm-polymer library (https://github.com/mirnylab/openmm-polymer-legacy), as described previously (Banigan et al., 2020; Fudenberg et al., 2016; Goloborodko et al., 2016a). The simulation code implementing the moving barrier model with these packages is freely and publicly available

(https://github.com/mirnylab/moving-barriers-paper). Loop extrusion dynamics with RNAP moving barriers are computed through the 1D model described below. Genomic positions of loop extruders as a function of time determine which monomeric subunits of the polymer are bridged at any particular instant in 3D simulations.

Computation of loop extrusion dynamics

Cohesin dynamics

The chromosome is modeled as a 1D array of $L=10^4$ genomic (lattice) sites, each of which represents 1 kb of chromatin. Cohesin complexes are modeled as loop-extruding factors (LEFs) with two linked components. Each component of a cohesin complex occupies a distinct, single lattice space. A LEF is loaded onto a pair of adjacent lattice sites that is not occupied by another LEF or RNAP. At each subsequent timestep, each component of the LEF may translocate to an unoccupied adjacent site with a probability determined by the type of extrusion dynamics simulated. Each of the two components in an individual LEF translocates (or not) independently of the other. For active, directed extrusion, each LEF component translocates away from its initial loading site (growing the loop) onto the next adjacent site with probability $v=1$, provided that the new lattice site is unoccupied; LEF components stop when they encounter another LEF component. For passive, diffusive extrusion, a LEF component may translocate in either direction along the chromosome lattice with equal probability, again provided that the new lattice site is unoccupied. Each component translocates in a single direction during each timestep, with probability $v=0.5$ per possible move. In both cases, extrusion proceeds in this manner (modulated by interactions with RNAPs; see below) until unbinding from the chromosome, which occurs with probability $1/\tau$, where τ is the mean residence time. Upon unbinding, the LEF is instantaneously reloaded to another pair of lattice sites. Loading biases are implemented by increasing the probability of binding to the TSS to 100-fold that of other lattice sites.

For active extrusion, $\tau=100$ for CTCF KO and $\tau=1000$ for DKO simulations, giving mean LEF processivities of $\lambda=200$ kb and 2000 kb, respectively, similar to previous studies (Banigan et al., 2020; Nuebler et al., 2018). In passive extrusion simulations, standard parameters were $\tau=5000$ or 50000 for CTCF KO or DKO, respectively. Setting the extrusion speed to $v=1$ kb/s, sets 1 timestep = 1 s in directed extrusion simulations. In simulations with passive extrusion, rescaling the lifetimes to match the directed case sets 1 timestep = 0.02 s and the LEF diffusion coefficient to $D=12.5$ kb^2/s, consistent with previous simulations (Banigan et al., 2020) and experimental measurements (Davidson et al., 2016; Kanke et al., 2016; Stigler et al., 2016). For simulations with slow diffusive extrusion, $\tau=100$ for CTCF KO and $\tau=1000$ for DKO, and $v=0.07$ kb/s ($D=0.005$ kb^2/s).

Simulations are performed with $N=L/d=50$ LEFs in CTCF KO and $N=100$ LEFs in DKO (Nuebler et al., 2018), giving a mean separation $d=200$ kb and 100 kb, respectively.

RNA polymerase dynamics

RNAP barriers were incorporated by adapting the moving barrier model for bacteria (Brandão et al., 2019). To simulate contact patterns near genes, we simulated chromosomes with 9 genes,

with TSSs located at s=950, 1950, 2950, …, 8950 kb and TTSs located 110 lattice sites (110 kb) downstream. For simulations of genes in convergent orientations, convergently oriented gene pairs had TSSs at (s_1, s_2)= (840 kb, 1160 kb), (1640 kb, 1960 kb), (2440 kb, 2760 kb), …, (8840 kb, 9160 kb); thus, the sites of convergent transcription used for pile ups were located at s=1000, 1800, 2600, …, 9000 kb. RNAP may be loaded onto an unoccupied TSS with at rate k_{load}. It remains paused at the TSS until it unpauses at rate $k_{unpause}$. After unpausing, the RNAP stochastically translocates by at most one lattice site per timestep toward the TTS at rate v_p, provided the site is unoccupied by another RNAP. After passing the TTS, the RNAP stalls at rate k_{stall}, and subsequently unbinds at rate k_{unbind}. Only one RNAP may occupy a single lattice site.

In simulations with directed extrusion, k_{load}=0.001, $k_{unpause}$=0.002, v_p=0.1, k_{stall}=0.001, and k_{unbind}=0.002. These correspond to k_{load}=0.001 s^{-1}, $k_{unpause}$=0.002 s^{-1}, v_p=0.1 kb/s (Darzacq et al., 2007; Wang et al., 1998), k_{stall}=0.001 s^{-1}, k_{unbind}= 0.002 s^{-1}. Times (and thus, rates) are rescaled in passive extrusion simulations according to the cohesin residence time, so there, k_{load}=$2\cdot10^{-5}$, $k_{unpause}$=$4\cdot10^{-5}$, v_p=0.002, k_{stall}=$2\cdot10^{-5}$, and k_{unbind}=$4\cdot10^{-5}$.

Cohesin-polymerase interactions

When a RNAP and a LEF arrive at adjacent lattice sites and the RNAP is translocating toward the LEF, they are in a head-on collision (**Fig. 2c**). The LEF may translocate past the RNAP at rate k_{bypass}. However, at any timestep for which LEF remains in the site adjacent to the RNAP and the RNAP translocates toward the LEF, the LEF is pushed by one site along the lattice in the direction of RNAP translocation. If one or more LEF components are on the lattice immediately behind the pushed LEF component, those LEFs are also pushed in the direction of RNAP translocation. In the case where a LEF component is at a lattice site adjacent to the RNAP and the RNAP is translocating *away* from the LEF (*e.g.*, they are both moving toward the TTS; **Fig. 2c**), the LEF component may only translocate if it bypasses the RNAP (at rate k_{bypass}) or if the RNAP vacates (by translocation or unbinding) the lattice site. We focus on results for simulations with characteristic bypassing time t_{bypass}=1/k_{bypass}=100 s, *i.e.*, we present k_{bypass}=0.01 (active) and k_{bypass}=0.0002 (passive), but results for other bypassing rates are shown in **Extended Data Figs. 3-4**.

Polymer molecular dynamics simulations

3D molecular dynamics simulations with Langevin dynamics were performed with openmm-polymer, as described previously (Banigan et al., 2020; Fudenberg et al., 2016; Goloborodko et al., 2016a), with timestep dt=80. L=10000 monomeric subunits were arranged into a linear polymer through pairwise harmonic bond interactions. Monomeric subunits repel each other through weak excluded-volume-like interactions.

Each monomeric subunit corresponds to a 1 kb genomic locus simulated in the loop extrusion model. The positions of LEFs indicate additional monomers that are bridged by harmonic bonds. For each configuration of LEFs, the polymer was evolved for 200 timesteps, after which the bonds due to LEFs are updated according to the next step in the extrusion dynamics computation.

Simulations were run for at least 40 LEF residence times, τ, before data collection, after which simulations were run for at least 90τ. At least 4 simulations were performed per parameter set, with 3000 configurations from each simulation included in the analysis. To generate contact maps, we used a cut-off radius r_c=2 monomers, which may be taken to correspond to r_c~50-100 nm (since 1 monomer is 1 kb or a few nucleosomes).

Data availability statement

Hi-C, Pol II ChIP-seq, and NIPBL ChIP-seq data are available upon request to the authors, and the data will be made available in the Gene Expression Omnibus (GEO). GRO-seq and Scc1 and CTCF ChIP-seq data were previously reported in (Busslinger et al., 2017), and they are available on GEO, accession number GSE76303.

Software availability statement

Codes used for analysis of ChIP-seq and Hi-C data and moving barrier model simulations are freely and publicly available at https://github.com/mirnylab/moving-barriers-paper. Other relevant analysis and simulation codes were previously published and are available as described in the **Methods**.

Figure Legends

Figure 1. Transcription and cohesin generate characteristic patterns of contacts and cohesin accumulation near active genes. a, *Top,* ChIP-seq tracks for CTCF in WT (black) and Scc1 (cohesin) in WT, CTCF KO, Wapl KO, and DKO cells (purple, blue, red, and orange, respectively) for a 5 Mb region of chromosome 1, with the corresponding gene track below. *Bottom,* Hi-C contact maps for the corresponding region. Boxes in DKO identify examples of island-island dots, with arrows pointing to the corresponding “cohesin islands” in the ChIP-seq tracks. *Insets,* Averages of observed-over-expected contacts (see **Methods**), centered on island-island dots separated by genomic distances 50 kb < s < 350 kb (n=1314), plotted with $\log_{10}$ color scale. Numbers indicate dot strengths (see **Methods**). **b,** Average observed-over-expected Hi-C contact maps centered on the TSS for oriented short active genes (length 10 kb < L < 30 kb; GRO-seq TPM > 3; n=644 except for DKO with n=666) and long active genes (80 kb < L < 120 kb; n=200 except DKO n=177). **c,** Cohesin (Scc1) ChIP-seq heatmaps and average tracks near short and long active genes oriented and aligned at their TSSs. Heatmaps depict the longest 50% of genes, sorted by decreasing length from top to bottom. Dotted lines in average plots indicate the length of the longest gene in the respective set.

Figure 2. Transcription as a moving barrier for loop extrusion recapitulates major features of genome organization and cohesin accumulation around active genes. a, Observed-over-expected contact maps around long active genes for CTCF KO and DKO with five major features identified and illustrated. **b,** Schematics of the moving barrier model. Cohesins (yellow and pink) bind to chromatin and extrude loops until unbinding. RNAPs (open

ellipses) are loaded at the promoter, translocate through the gene (purple), and are unloaded at the 3' end. **c,** Arch diagrams illustrating time series of two types of collisions between extruding cohesins and translocating RNAP that may occur in genes in the model. Yellow circles depict the two genomic positions at the base of the extruded loop, bridged by a cohesin. During head-on collisions, RNAP pushes cohesin until the cohesin bypasses the RNAP, the RNAP stops translocating (beyond the 3' end), or either the RNAP or cohesin unbinds. During co-directional collisions, extrusion by cohesin translocation is slowed by the RNAP barrier moving toward 3'. In both cases, interactions between RNAP and cohesin only alter extrusion on one side of cohesin; collisions do not affect growth of the other side of the extruded loop or RNAP translocation. **d,** Average observed-over-expected maps and cohesin accumulation tracks near active genes in CTCF KO and DKO simulations. Results shown for simulations with either active, directed loop extrusion or passive, diffusive loop extrusion, each with either uniform cohesin loading or preferential loading at TSSs. Gene positions are indicated by purple bars on the *x*-axes. *Inset*, illustrations of cohesin loading and translocation.

Figure 3. Predictions of the moving barrier model. a, *Left,* Drawing showing contacts between ends of nearby genes in four pairs of orientations. *Right,* Average observed-over-expected maps centered on contacts between nearest ends of active (TPM > 2) genes separated by 50 kb < *s* < 350 kb. At least one gene in each pair is not near a CTCF site (see **Methods**). **b,** Zoomed-in views of average observed-over-expected maps for CTCF KO and DKO in experiments and simulations with active, directed loop extrusion, with and without preferential cohesin loading at TSSs. Arrows indicate the presence or lack of an extra line of enriched genomic contacts characteristic of cohesin loading at the TSS. Gene positions are shown by purple bars on the axes.

Figure 4. NIPBL and MAU2 colocalize with cohesin but not TSS throughout the genome. a, Enrichment profiles of NIPBL (133M, -/+ NIPBL RNAi) along an exemplary 330 kb region of chromosome 2, illustrating typical distribution and co-localization of sequencing read pileups. Genes within this region are depicted above. Arrow indicates a *bona fide* NIPBL site. **b,** Area-proportional threefold eulerAPE Venn diagram illustrating overlap between NIPBL (133M, -/+ NIPBL RNAi) and TSS or SCC1. **c,** Average signal profiles of NIPBL (133M, -/+ NIPBL RNAi) binding around TSS or NIPBL/MAU2 co-sites. **d,** ChIP-Seq profiles of HA-NIPBL (-/+dTAG), MAU2-HA (-/+dTAG), HA-WAPL (-/+dTAG), GFP-WAPL, SCC1-GFP and CTCF-GFP along an exemplary 581 kb region of chromosome 2, illustrating the typical distribution and co-localization of sequencing read pileups. Genes within this region are depicted above. The red rectangle on the left indicates one region of interest and a zoom-in view is shown on the right. **e,** Heatmaps of HA-NIPBL (-/+dTAG) and SCC1 ChIP-Seq signal prominence at TSSs overlapping cohesin, TSSs without cohesin, and cohesin sites not at TSSs. **f,** Area-proportional threefold Venn diagram illustrating overlap between NIPBL/MAU2 co-sites, SCC1 and TSS. **g,** Enrichment profiles of HA-NIPBL (-/+ SCC1 RNAi) and MAU2-HA (-/+ SCC1 RNAi) along an exemplary 185 kb region of chromosome 2, illustrating typical distribution and co-localization of sequencing read pileups. Genes within this region are depicted above. **h,** Average signal profiles of HA-NIPBL (-/+ SCC1 RNAi) or MAU2 (-/+ SCC1 RNAi) around NIPBL/MAU2 co-sites.

Extended Data Figure Legends

Extended Data Figure 1. Hi-C in WT, CTCF KO, Wapl KO, DKO, and Smc3 KO cells shows genomic contacts depend on extrusion dynamics. a, Western blots for Wapl, Smc3, Scc1, CTCF, Smc3Ac, and loading control α-tubulin in wildtype (control) and different mutants. **b,** Contact probability, $P_c(s)$, and log-derivative, dlog $P_c(s)$ / dlog s, for WT and mutants, as a function of genomic distance, s (left and right, respectively). Dotted lines in log-derivative plot indicate inferred mean loop sizes (Gassler et al., 2017). **c,** Observed-over-expected maps piled up on the top 10% scoring CTCF sites (n=4524), showing insulation around CTCF sites. **d,** Observed-over-expected maps piled up on previously identified cohesin islands (Busslinger et al., 2017). **e,** Scc1 ChIP-seq and Hi-C contact map of Smc3 KO from the same representative 5 Mb region as in **Fig. 1a**. WT CTCF ChIP-seq track is shown for reference. *Inset*, Observed-over-expected pile ups on island-island dots for Smc3 KO, with dot strength indicated. **f,** Observed-over-expected pile ups on island-island dots for WT and DKO cells treated with DRB for 3 h and after DRB release by washing 3x with PBS and incubating cells without DRB for 24 h. **g,** Observed-over-expected pile ups centered on cohesin islands for WT and DKO DRB and DRB release.

Extended Data Figure 2. Cohesin dynamics and transcription govern genomic contact patterns near genes. a, Hi-C observed-over-expected pile ups centered on TSSs for all active genes (top row) and two medium-length sets genes (30-55 kb and 55-80 kb, middle and bottom, respectively) for WT, CTCF KO, Wapl KO, DKO, and Smc3 KO. **b,** Scc1 ChIP-seq heatmaps and average tracks for the same sets of genes. Heat maps show individual gene tracks from the corresponding, sorted by decreasing length from top to bottom, for the longest 50% of genes, except for the set of all active genes, which shows heat maps for the longest 300 genes. Dotted lines show median gene length (49 kb) for all genes or longest gene length for the other two sets. **c,** Hi-C observed-over-expected pile ups centered on TSSs for all inactive genes (TPM<1, top row), short inactive genes (middle), and long inactive genes (bottom). **d,** Scc1 ChIP-seq heatmaps and average tracks for the same sets of genes, with heat maps showing the longest 300 genes for the set of all genes and the longest 50% of genes for short and long genes. Dotted lines show median gene length (49 kb) for all genes or longest gene length for the other two sets. **e,** Pile ups on TSSs for short and long active genes for Smc3 KO. **f,** Pile ups for short and long active genes for genes that are at least 5 kb away from the top 20% of CTCF sites. **g,** Pile ups on TSSs for all, short, and long active genes (using TPM>3 from WT GRO-seq or, for DKO, DKO GRO-seq) in WT with DRB treatment or DRB treatment and release and in DKO with DRB treatment or DRB treatment and release. DRB treatment and release were as described in **Extended Data Fig. 1**. **h,** RNAP II ChIP-seq heat maps and average tracks for short and long active genes. Heat maps show 300 short genes (here, 20-30 kb) and 168 long genes. All non-overlapping genes were included in RNAP II stack ups. Dotted lines indicate shortest and longest gene lengths in the corresponding sets.

Extended Data Figure 3. Moving barrier simulations with diffusive cohesins with different diffusion coefficients, RNAP bypassing rates, and loading biases. a, Pile ups centered on TSSs showing observed-over-expected maps for simulations with passive cohesins that grow

loops by diffusion and cannot bypass RNAP with loading preferentially at TSSs or **b,** uniform probability of loading along the genome. **c,** Simulations with slowly diffusing cohesins (D=0.005 kb^2/s) that cannot bypass RNAP and preferentially load at TSSs or **d,** load with uniform probability along the genome. **e,** Slowly diffusing cohesins that *can* bypass RNAPs with cohesin loading preferentially at TSSs or **f,** cohesin loading uniformly throughout the genome.

Extended Data Figure 4. Moving barrier simulations with actively extruding cohesins with different bypassing rates and loading biases. a, Observed over expected pile ups and average cohesin accumulation tracks centered on TSSs for active extrusion simulations with fast cohesin-RNAP bypassing (k_{bypass}=0.04 s^{-1}, as opposed to k_{bypass}=0.01 s^{-1} in the main text) with preferential cohesin loading at TSSs or **b,** with uniform loading probability across the genome. **c,** Simulations with active extrusion, slow cohesin-RNAP bypassing (k_{bypass}=0.001 s^{-1}), and preferential cohesin loading at TSSs or **d,** uniform loading probability throughout the genome. **e,** Simulations with active extrusion in which cohesin-RNAP bypassing is not permitted (*i.e.*, k_{bypass}=0 s^{-1}) with preferential cohesin loading at TSSs or **f,** uniform loading.

Extended Data Figure 5. Simulations of convergently oriented genes produce cohesin islands and island-island dots in DKO conditions. *Top,* Pile ups centered on sites of convergent transcription showing observed over expected contacts for CTCF KO and DKO simulations. Gene track (top) indicates positions, lengths, and orientations of genes (purple arrows) on the chromosome polymer fiber (gray). Boxes identify features of interest, such as contacts formed between cohesin islands or contacts between gene ends. *Bottom*, Average cohesin occupancy along the polymer fiber in simulations in a small region around a convergent gene pair. Dotted lines indicate the 5' and 3' ends of genes.

Extended Data Figure 6. 3' and 5' ends of active genes act as extrusion barriers and generate genomic contacts. a, Pile ups centered on contacts between nearest ends of pairs of adjacent active genes (TPM>2) within a genomic distance 50 kb< s < 350 kb of each other, separated by relative orientation: convergent, divergent, and parallel (top, middle, and bottom, respectively). Schematic drawings illustrate the orientations. Purple bars indicate positions of genes in the pile ups. **b,** Pile ups centered on contacts of nearest ends of pairs of adjacent inactive genes (TPM<1) within a genomic distance 50 kb< s < 350 kb of each other, where at least one gene in each pair is not near a CTCF site (see **Methods**).

Extended Data Figure 7. NIPBL and MAU2 colocalize with cohesin but not TSSs throughout the genome. a, Immunoblotting analysis of HeLa cells treated with or without NIPBL siRNAs for 48 hours. **b,** Heatmaps of NIPBL (133M, -/+ NIPBL RNAi) and SCC1 ChIP-Seq signal prominence at TSS or NIPBL/MAU2 co-sites. **c,** Chromatin fractionation of HA-FKBP-NIPBL/Halo-AID-WAPL/SCC1-EGFP, MAU2-AID-FKBP/Halo-AID-WAPL/SCC1-EGFP, and HA-FKBP-WAPL cells were analyzed by immunoblotting with or without of dTAG for 24 hours. **d,** Genotype analysis of HeLa wildtype, homozygous HA-FKBP-NIPBL/ Halo-AID-WAPL/SCC1-EGFP, MAU2-AID-FKBP/Halo-AID-WAPL/SCC1-EGFP and HA-FKBP-WAPL cells. **e,** Representative immunofluorescence images of HA-FKBP-NIPBL/ Halo-AID-WAPL/SCC1-EGFP cells treated with auxin or/and dTAG for 24 hours. DNA was

counterstained with DAPI. *Inset*, magnified images of boxed regions. Scale bar shows 10 μm. **f,** Representative immunofluorescence images of MAU2-FKBP-HA/Halo-AID-WAPL/SCC1-EGFP cells treated with auxin or/and dTAG for 24 hours. DNA was counterstained with DAPI. *Inset*, magnified images of boxed regions. Scale bar shows 10 μm. **g,** Heatmaps of HA-NIPBL, MAU2-HA (-/+dTAG), and SCC1 ChIP-Seq signal prominence at TSSs with cohesin, TSSs without cohesin, and cohesin without TSSs. **h,** Immunoblotting analysis of HA-NIPBL or MAU2-HA cells treated with or without SCC1 siRNA for 48 hours.

Extended Data Figure 8. Transcription termination sites of active genes generate insulation. Observed over expected Hi-C contact maps for short and long active genes in WT, CTCF KO, Wapl KO, and DKO, piled up and centered on their transcription termination sites (TTSs).

Acknowledgements

We thank Nezar Abdennur, Sameer Abraham, Anton Goloborodko, and Maxim Imakaev for instructive discussions and sharing code. We also thank Job Dekker, Anne-Laure Valton, and Sergey Venev for helpful conversations. We are grateful to the Sequencing Team at IMP Vienna for their assistance in the sequencing experiments. HBB was partially supported by R00GM130896. LAM acknowledges funding from R01GM114190 of the National Institutes of Health, the NIH Common Fund 4D Nucleome Program (UM1HG011536), the Human Frontier Science Program (grant RGP0057/2018), and Chaire Internationales d'excellence Blaise Pascal. JMP acknowledges funding from Boehringer Ingelheim, the Austrian Research Promotion Agency (Headquarter grant FFG-852936), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements No 693949 and No 101020558), the Human Frontier Science Program (grant RGP0057/2018), and the Vienna Science and Technology Fund (grant LS19-029). For the purpose of Open Access the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission. JMP is also an adjunct professor at the Medical University of Vienna.

Author Contributions

EJB, AAB, HBB, and LAM conceived and designed the simulation model. EJB and AAB performed and analyzed the simulations. LAM supervised the simulation research. WT, GW, GAB, and JMP conceived and designed the experiments. WT and GW performed the experiments. EJB, WT, AAB, RRS, and GW analyzed the experimental data. JMP supervised the experimental research. LAM and JMP supervised analysis of the experimental data. EJB drafted the manuscript with input from LAM and JMP. All authors discussed the results and commented on the manuscript.

Figures

Figure 1

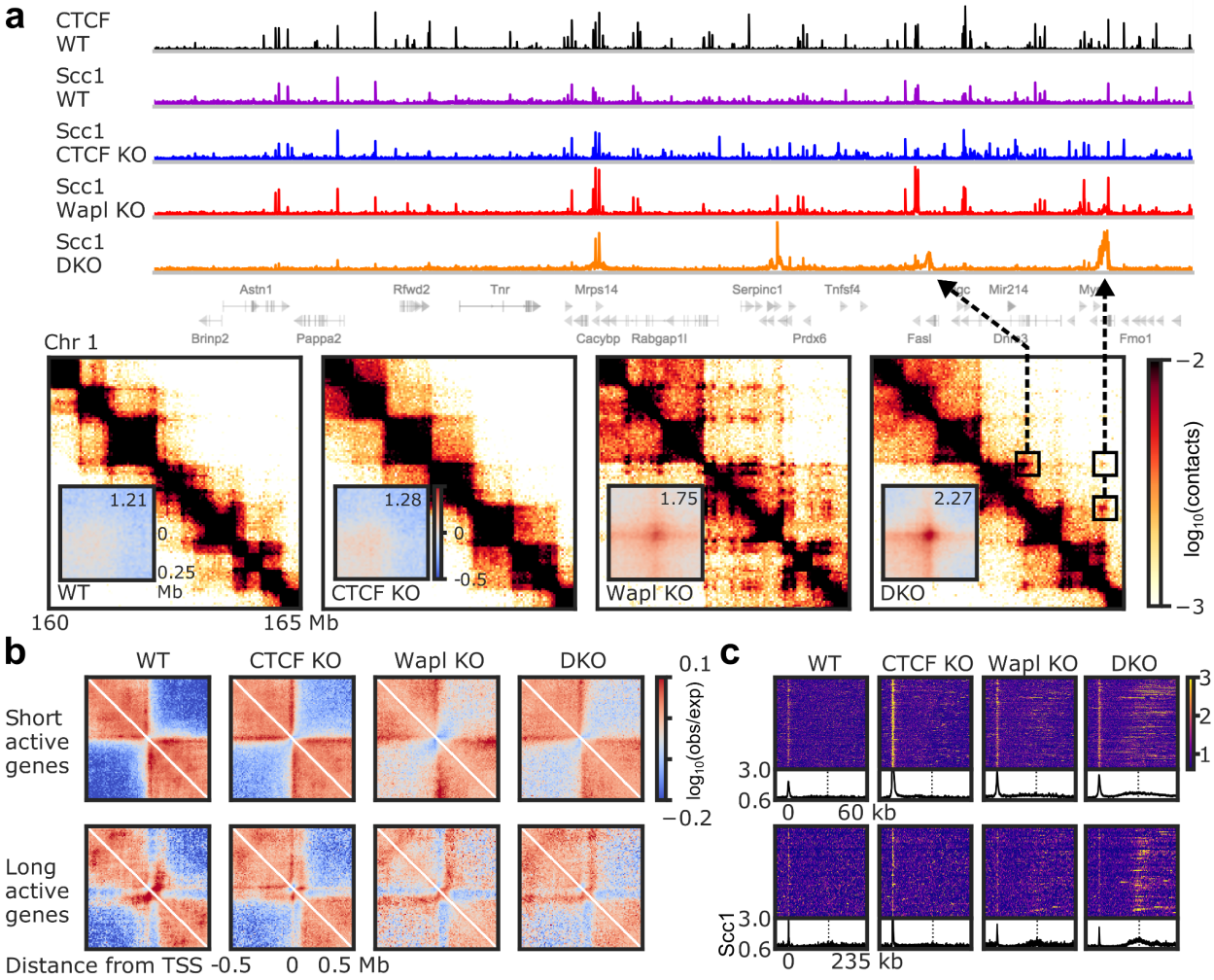


Figure 2

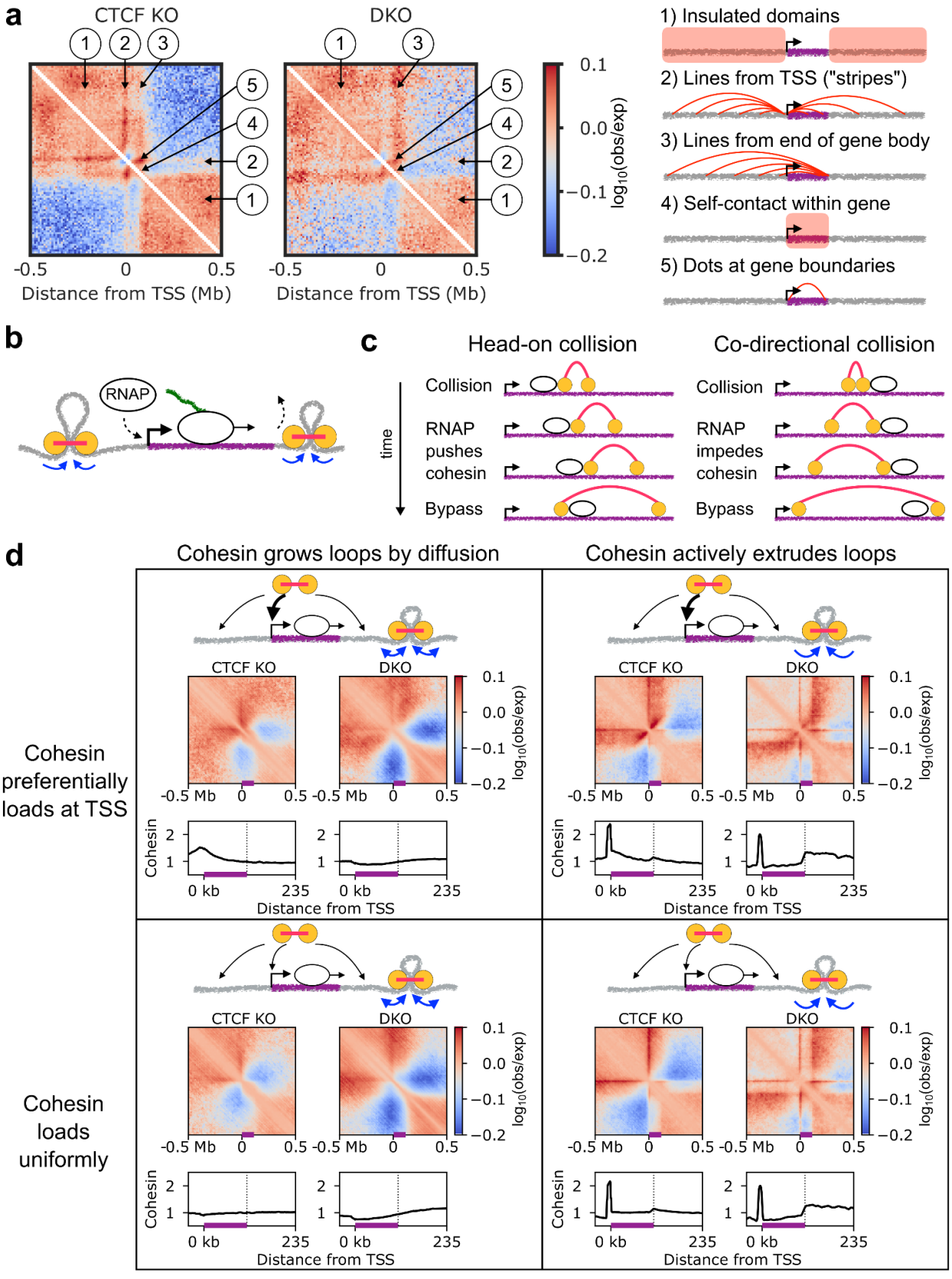


Figure 3

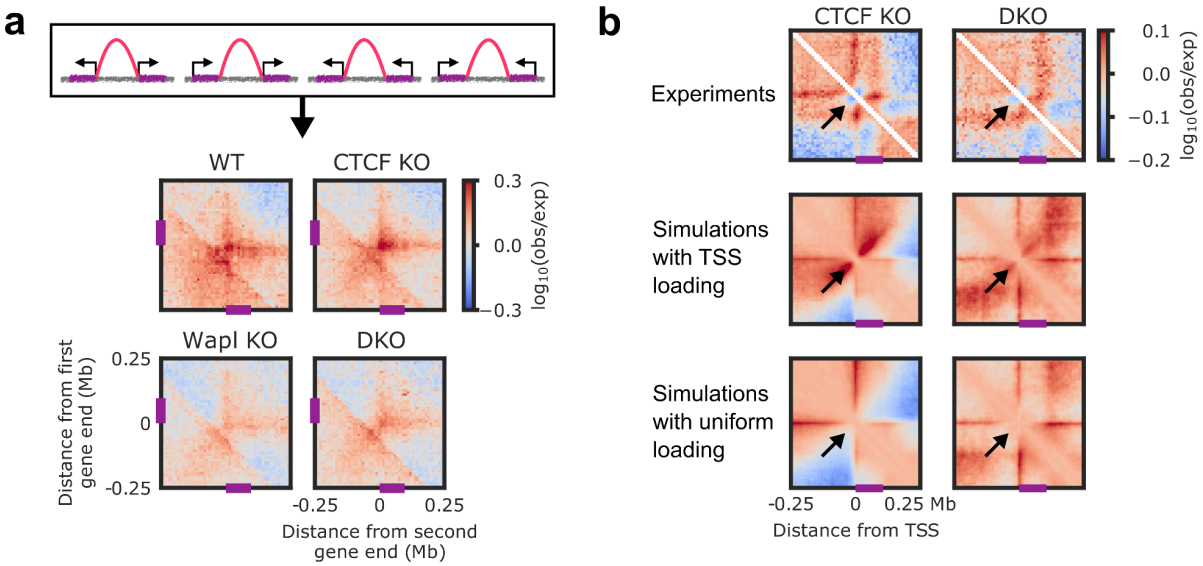
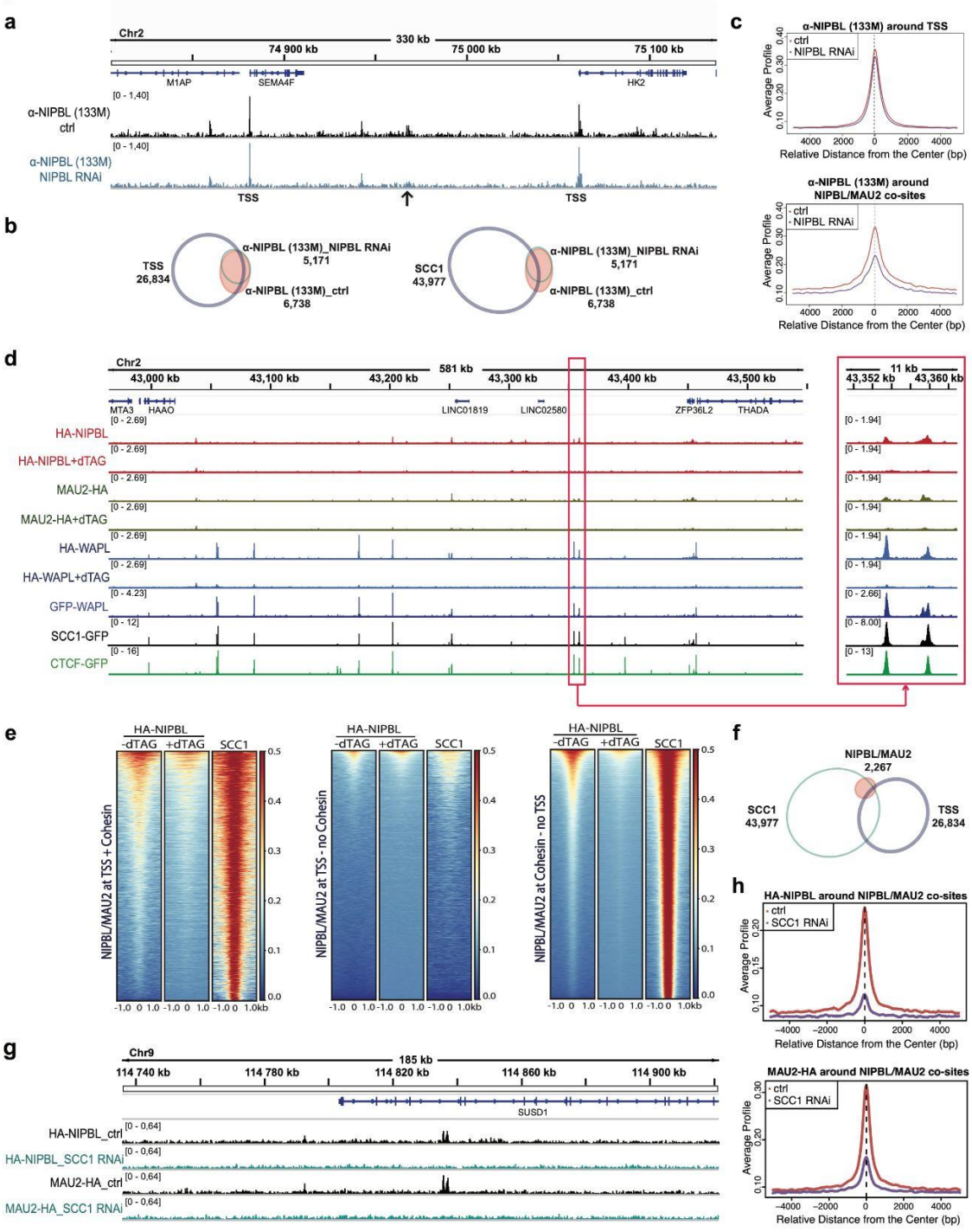
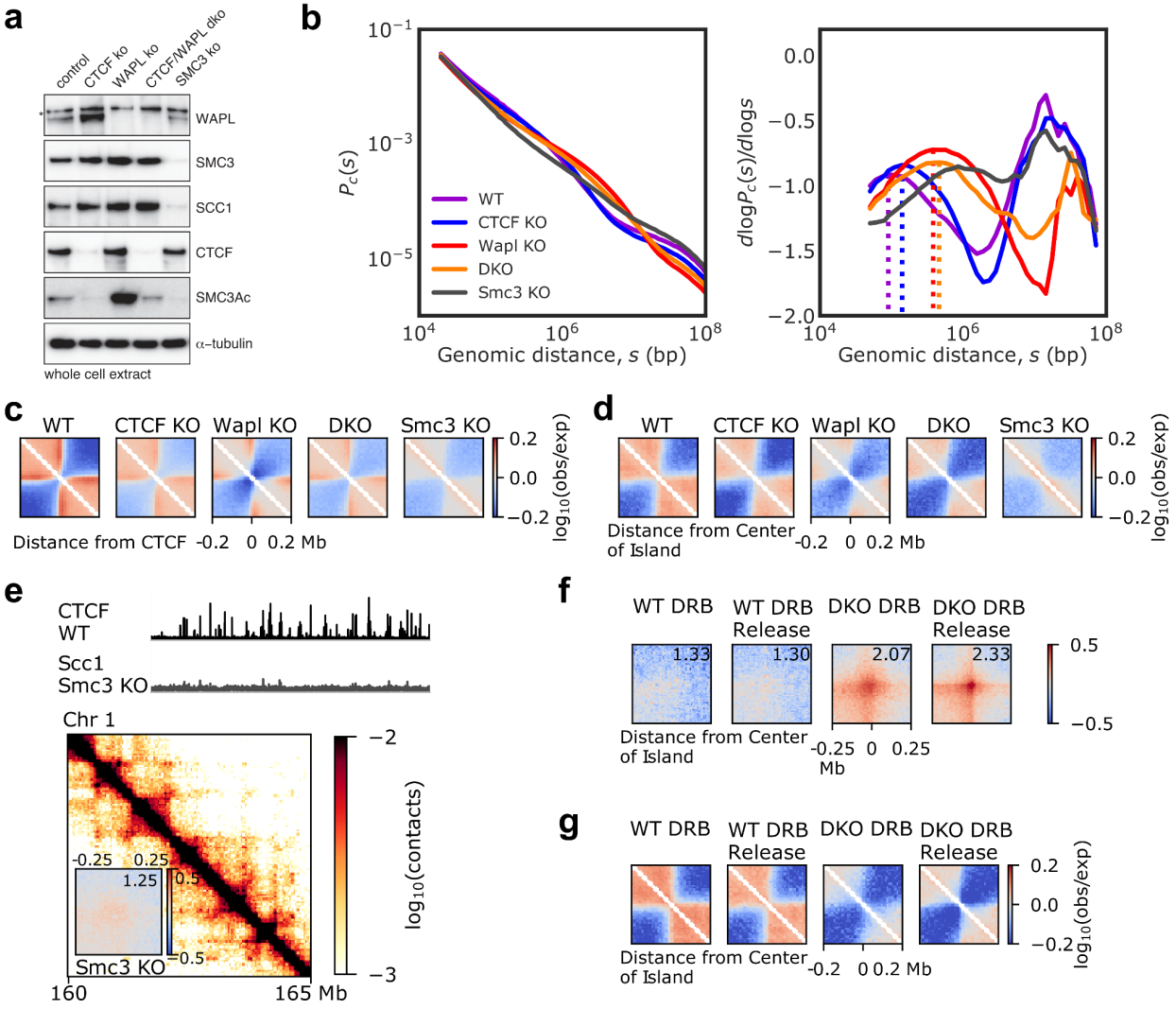


Fig 4

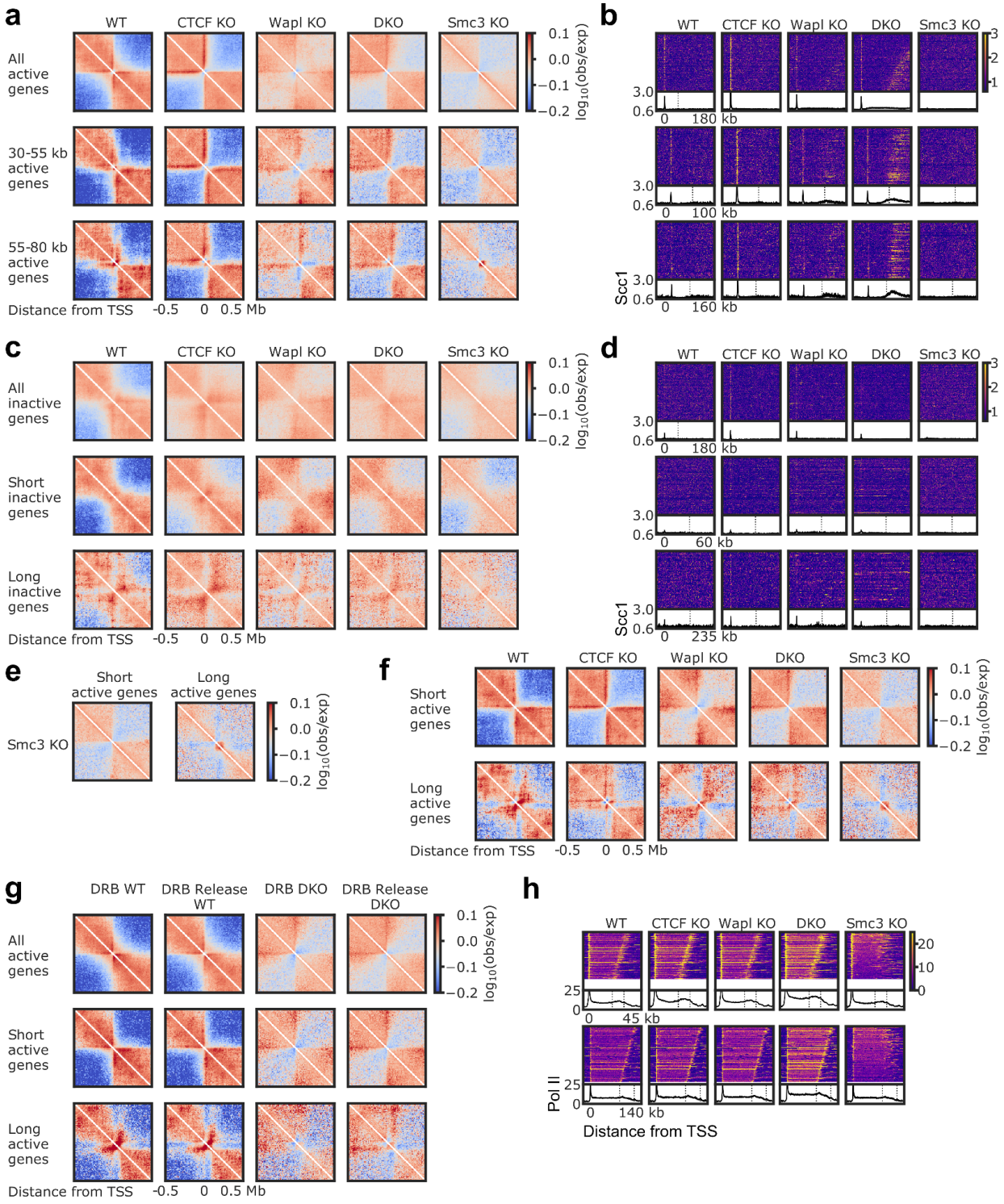


Extended Data Figures

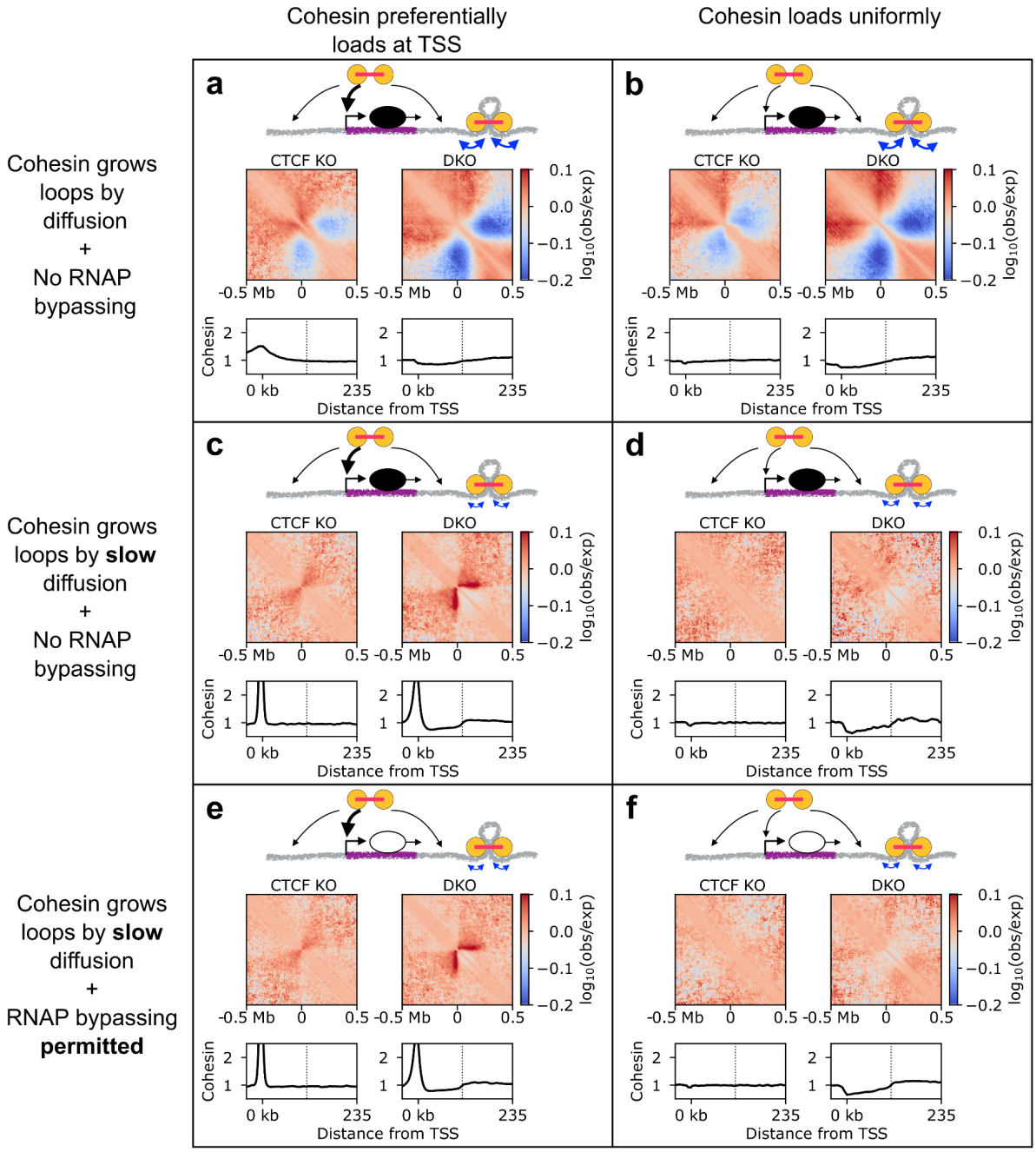
Extended Data Figure 1



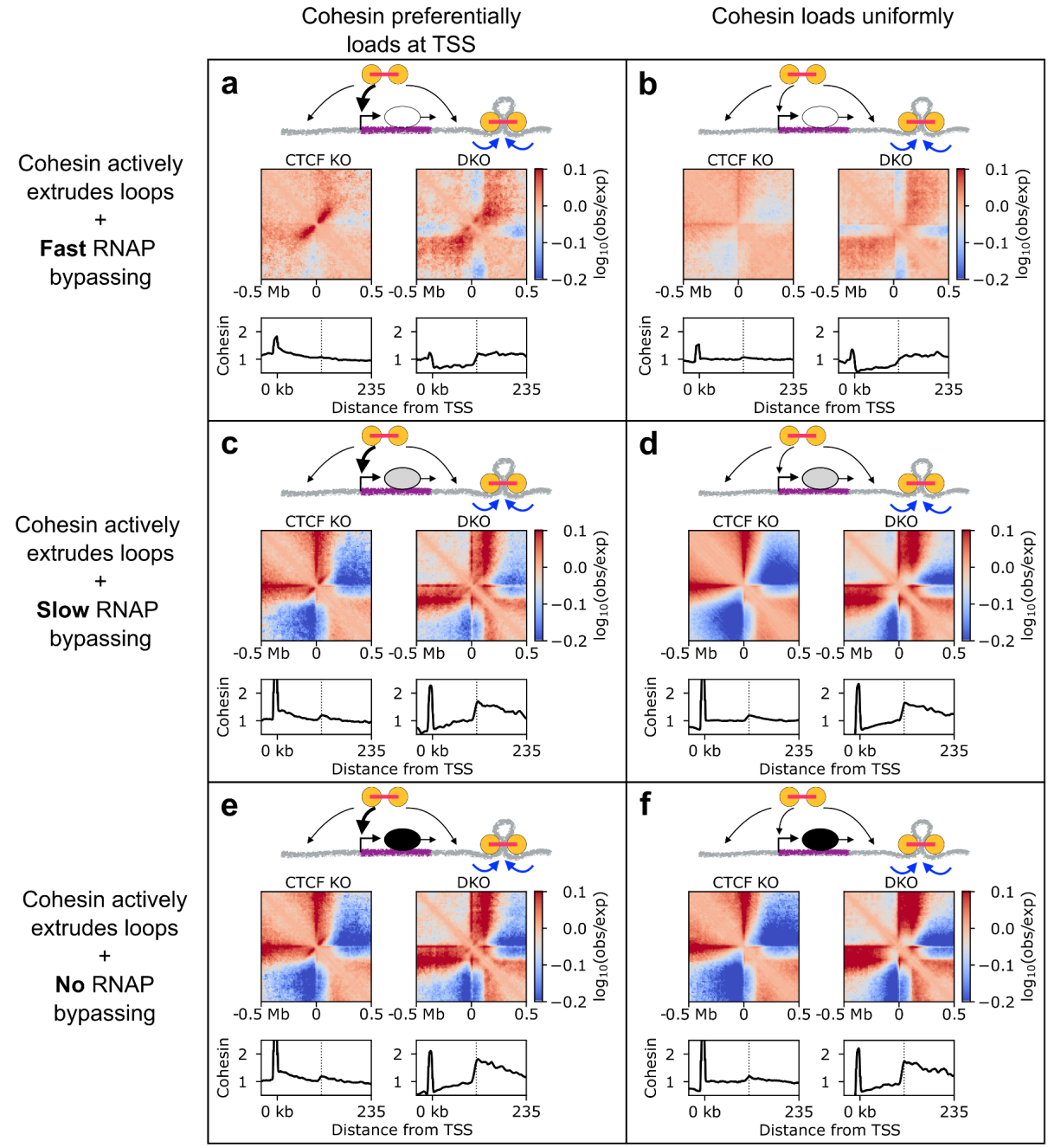
Extended Data Figure 2



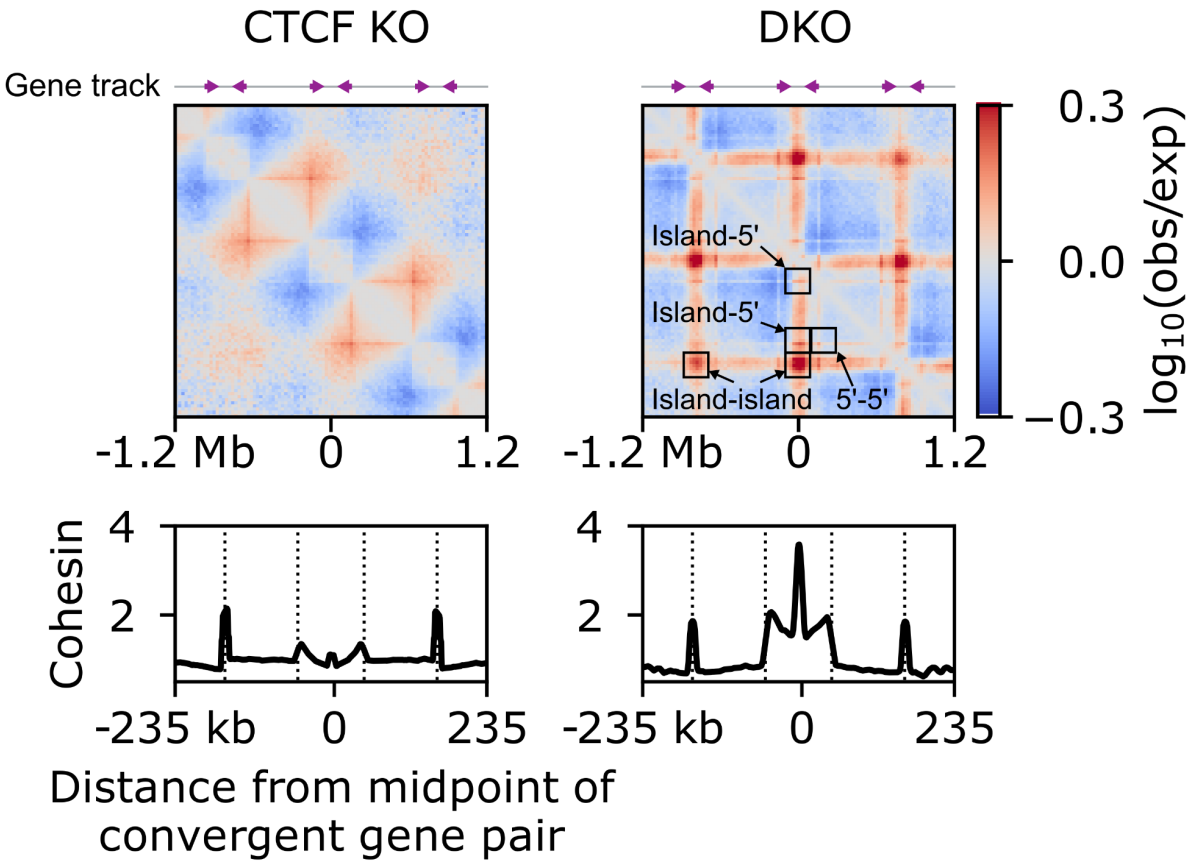
Extended Data Figure 3



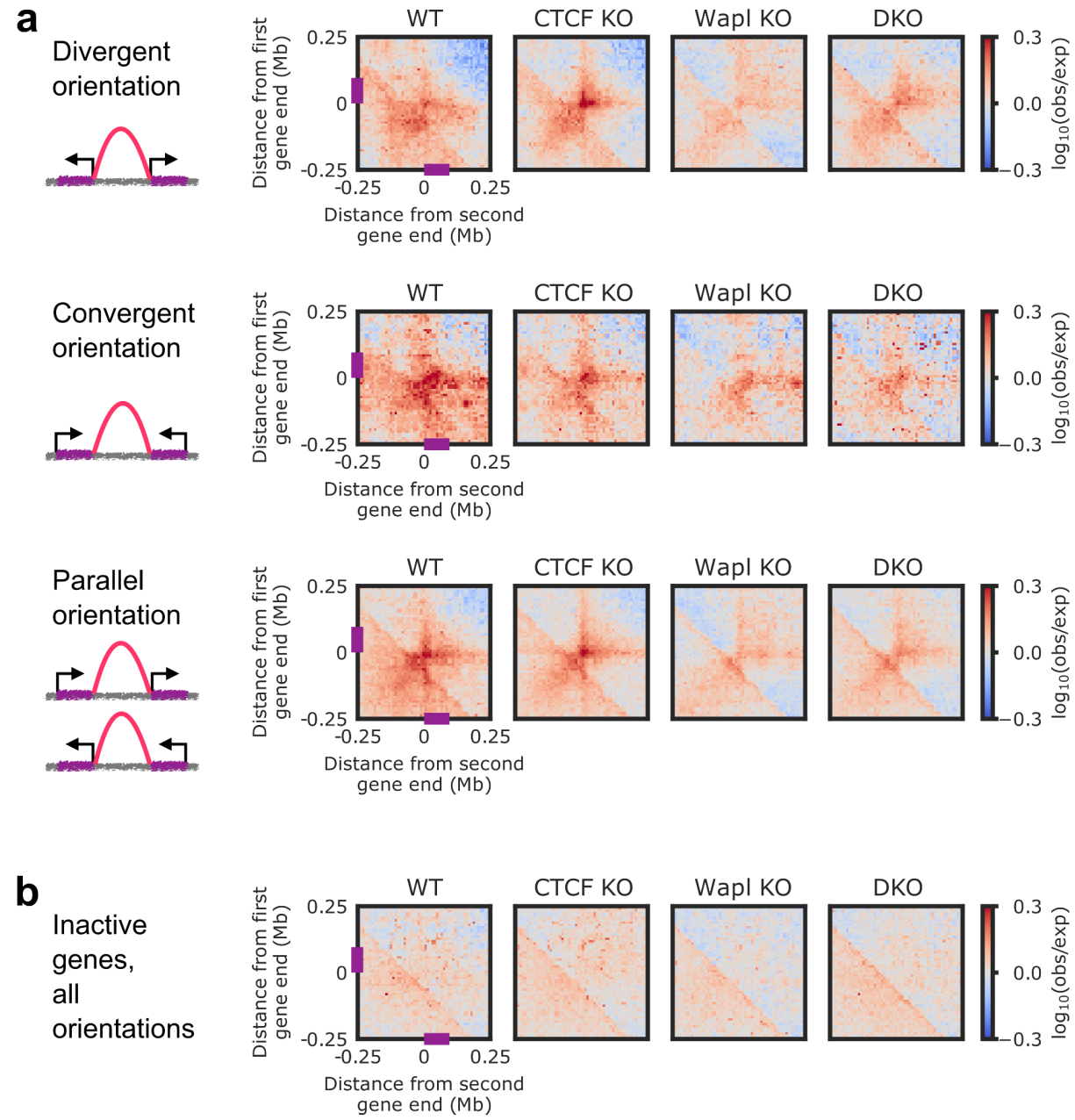
Extended Data Figure 4



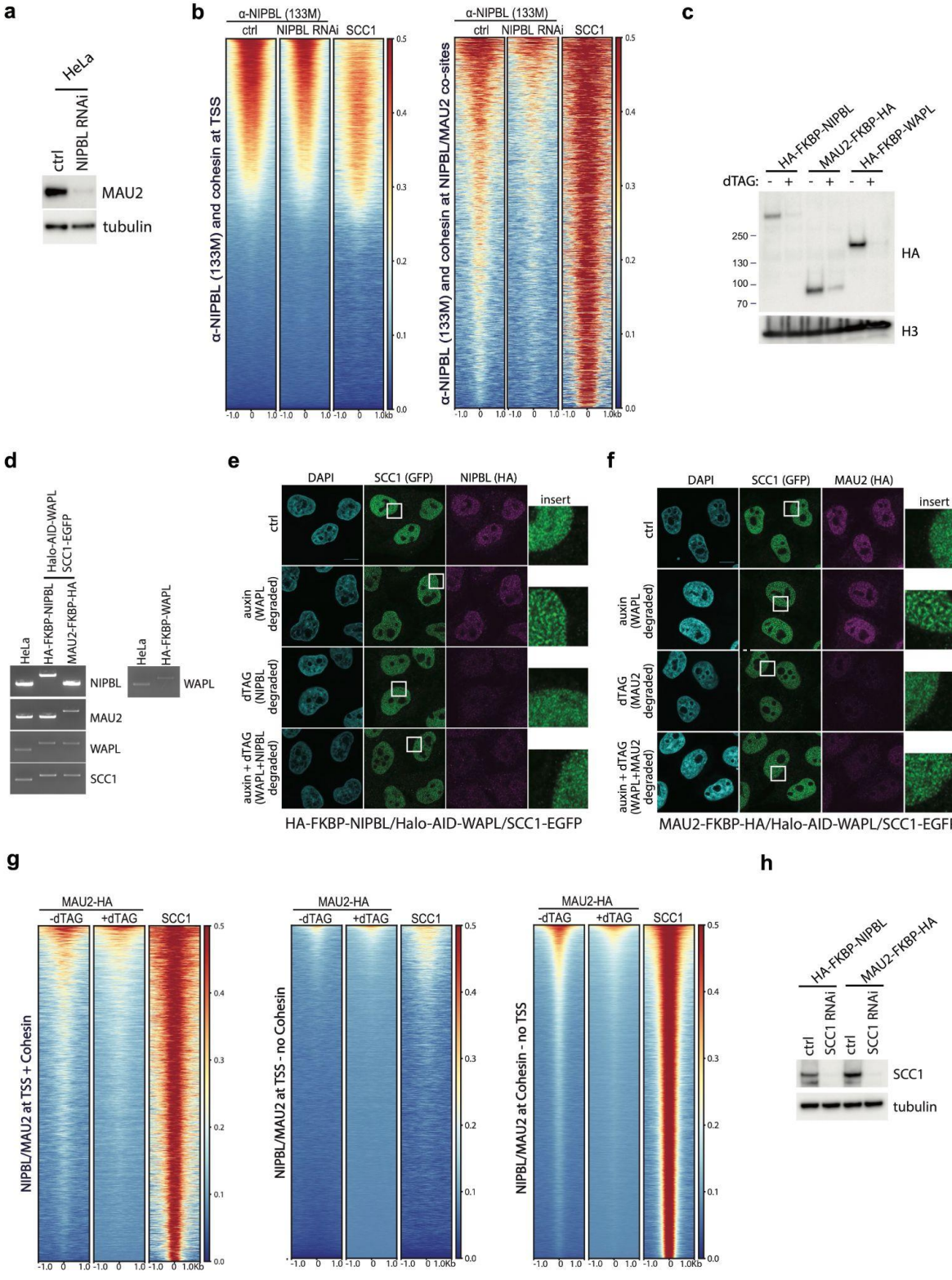
Extended Data Figure 5



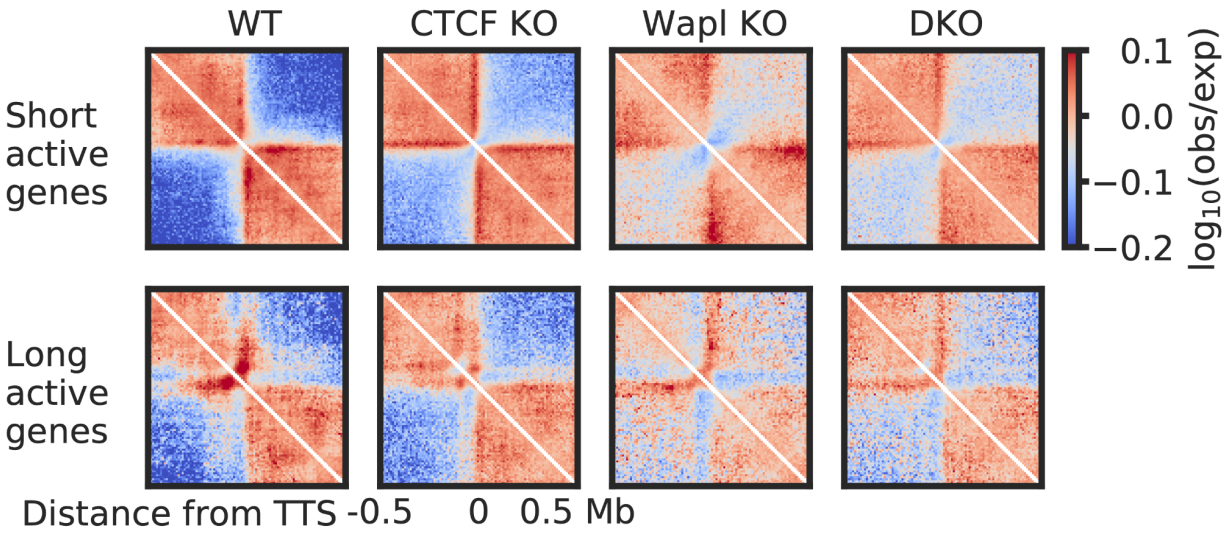
Extended Data Figure 6



Extended Data Figure 7



Extended Data Figure 8



References

Abdennur N, Mirny L. 2019. Cooler: scalable storage for Hi-C data and other genomically-labeled arrays. *Bioinformatics*. doi:10.1093/bioinformatics/btz540

Alipour E, Marko JF. 2012. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res* **40**:11202–11212.

Arnould C, Rocher V, Finoux A-L, Clouaire T, Li K, Zhou F, Caron P, Mangeot PE, Ricci EP, Mourad R, Haber JE, Noordermeer D, Legube G. 2021. Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature* **590**:660–665.

Aszódi A. 2012. MULTOVL: fast multiple overlaps of genomic regions. *Bioinformatics* **28**:3318–3319.

Banigan EJ, Mirny LA. 2020. Loop extrusion: theory meets single-molecule experiments. *Curr Opin Cell Biol* **64**:124–138.

Banigan EJ, van den Berg AA, Brandão HB, Marko JF, Mirny LA. 2020. Chromosome organization by one-sided and two-sided loop extrusion. *Elife* **9**. doi:10.7554/eLife.53558

Barrington C, Georgopoulou D, Pezic D, Varsally W, Herrero J, Hadjur S. 2019. Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nat Commun* **10**:2908.

Blackwood EM, Kadonaga JT. 1998. Going the distance: a current view of enhancer action. *Science* **281**:60–63.

Brackley CA, Johnson J, Michieletto D, Morozov AN, Nicodemi M, Cook PR, Marenduzzo D. 2017. Nonequilibrium Chromosome Looping via Molecular Slip Links. *Phys Rev Lett* **119**:138101.

Brandão HB, Paul P, van den Berg AA, Rudner DZ, Wang X, Mirny LA. 2019. RNA polymerases as moving barriers to condensin loop extrusion. *Proc Natl Acad Sci U S A* **116**:20489–20499.

Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N, Peters J-M. 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**:503–507.

Canzio D, Nwakeze CL, Horta A, Rajkumar SM, Coffey EL, Duffy EE, Duffié R, Monahan K, O'Keeffe S, Simon MD, Lomvardas S, Maniatis T. 2019. Antisense lncRNA Transcription Mediates DNA Demethylation to Drive Stochastic Protocadherin α Promoter Choice. *Cell* **177**:639–653.e15.

Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K. 2000. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell* **5**:243–254.

Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH. 2007. In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* **14**:796–806.

Dauban L, Montagne R, Thierry A, Lazar-Stefanita L, Bastié N, Gadal O, Cournac A, Koszul R, Beckouët F. 2020. Regulation of Cohesin-Mediated Chromosome Folding by Eco1 and Other Partners. *Mol Cell*. doi:10.1016/j.molcel.2020.01.019

Davidson IF, Bauer B, Goetz D, Tang W, Wutz G, Peters J-M. 2019. DNA loop extrusion by human cohesin. *Science*. doi:10.1126/science.aaz3418

Davidson IF, Goetz D, Zaczek MP, Molodtsov MI, Huis in 't P, Weissmann F, Litos G, Cisneros DA, Ocampo-Hafalla M, Ladurner R, Uhlmann F, Vaziri A, Peters J. 2016. Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J* **35**:2671–2685.

Davidson IF, Peters J-M. 2021. Genome folding through loop extrusion by SMC complexes. *Nat Rev Mol Cell Biol* **22**:445–464.

Dequeker BJH, Brandão HB, Scherr MJ, Gassler J, Powell S, Gaspar I, Flyamer IM, Tang W, Stocsits R, Davidson IF, Peters J-M, Duderstadt KE, Mirny LA, Tachibana K. 2020. MCM

complexes are barriers that restrict cohesin-mediated loop extrusion. *bioRxiv*. doi:10.1101/2020.10.15.340356

de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, Splinter E, Wijchers PJ, Krijger PHL, de Laat W. 2015. CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell* **60**:676–684.

Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science Engineering* **12**:34–39.

Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, Brooks BR, Pande VS. 2017. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* **13**:e1005659.

Fudenberg G, Abdennur N, Imakaev M, Goloborodko A, Mirny LA. 2017. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb Symp Quant Biol* **82**:45–55.

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep* **15**:2038–2049.

Gassler J, Brandão HB, Imakaev M, Flyamer IM, Ladstätter S, Bickmore WA, Peters J, Mirny LA, Tachibana K. 2017. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J* **36**:3600–3618.

Glynn EF, Megee PC, Yu H-G, Mistrot C, Unal E, Koshland DE, DeRisi JL, Gerton JL. 2004. Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. *PLoS Biol* **2**:E259.

Golfier S, Quail T, Kimura H, Brugués J. 2020. Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. *Elife* **9**. doi:10.7554/eLife.53885

Goloborodko A, Imakaev MV, Marko JF, Mirny L. 2016a. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5**. doi:10.7554/eLife.14864

Goloborodko A, Marko JF, Mirny LA. 2016b. Chromosome Compaction by Active Loop Extrusion. *Biophys J* **110**:2162–2168.

Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**:1017–1018.

Gullerova M, Proudfoot NJ. 2008. Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* **132**:983–995.

Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, Brummelkamp TR, de Wit E, Rowland BD. 2017. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**:693–707.e14.

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, Marazzi I, García-Sastre A, Shaw ML, Benner C. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell* **174**:1522–1536.e22.

Hill L, Ebert A, Jaritz M, Wutz G, Nagasaka K, Tagoh H, Kostanova-Poliakova D, Schindler K, Sun Q, Bönelt P, Fischer M, Peters J-M, Busslinger M. 2020. Wapl repression by Pax5 promotes V gene recombination by Igh loop extrusion. *Nature* **584**:142–147.

Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R, Darzacq X. 2020. Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol Cell* **78**:539–553.e8.

Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9**:999–1003.

Jain D, Baldi S, Zabel A, Straub T, Becker PB. 2015. Active promoters give rise to false positive “Phantom Peaks” in ChIP-seq experiments. *Nucleic Acids Res* **43**:6959–6968.

Jeppsson K, Sakata T, Nakato R, Milanova S, Shirahige K, Björkegren C. 2020. Transcription and replication organize cohesin-dependent chromosome loops. *bioRxiv*.

Joung J, Engreitz JM, Konermann S, Abudayyeh OO, Verdine VK, Aguet F, Gootenberg JS, Sanjana NE, Wright JB, Fulco CP, Tseng Y-Y, Yoon CH, Boehm JS, Lander ES, Zhang F. 2017. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**:343–346.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**:430–435.

Kanke M, Tahara E, Nishiyama T. 2016. Cohesin acetylation and Wapl-Pds5 oppositely regulate translocation of cohesin along DNA. *EMBO J*.

Kim Y, Shi Z, Zhang H, Finkelstein IJ, Yu H. 2019. Human cohesin compacts DNA by loop extrusion. *Science* **366**:1345–1349.

Lee K, Hsiung CC-S, Huang P, Raj A, Blobel GA. 2015. Dynamic enhancer--gene body contacts during transcription elongation. *Genes Dev* **29**:1992–1997.

Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F. 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**:573–578.

Lin SG, Ba Z, Alt FW, Zhang Y. 2018. RAG Chromatin Scanning During V(D)J Recombination and Chromatin Loop Extrusion are Related Processes. *Adv Immunol* **139**:93–135.

Micallef L, Rodgers P. 2014. eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses. *PLoS One* **9**:e101717.

Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, Scott TG, Vittori S, Perry JA, Qi J, Winter GE, Wong K-K, Gray NS, Bradner JE. 2018. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14**:431–441.

Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**:930–944.e22.

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

Ocampo-Hafalla M, Muñoz S, Samora CP, Uhlmann F. 2016. Evidence for cohesin sliding along budding yeast chromosomes. *Open Biol* **6**. doi:10.1098/rsob.150178

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**:422–433.

Peters J-M. 2021. How DNA loop extrusion mediated by cohesin enables V(D)J recombination. *Curr Opin Cell Biol* **70**:75–83.

Pradhan B, Barth R, Kim E, Davidson IF, Bauer B, van Laar T, Yang W, Ryu J-K, van der Torre J, Peters J-M, Dekker C. 2021. SMC complexes can traverse physical roadblocks bigger than their ring size. *bioRxiv*. doi:10.1101/2021.07.15.452501

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**:W160–5.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**:2281–2308.

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**:1665–1680.

Rhodes J, Mazza D, Nasmyth K, Uphoff S. 2017. Scc2/Nipbl hops between chromosomal cohesin rings after loading. doi:10.7554/eLife.30000

Rivosecchi J, Vachez L, Gautier F, Bernard P, Vanoosthuyse V. 2020. RNA polymerase backtracking drives the accumulation of fission yeast condensin at active genes. *Cold Spring Harbor Laboratory*. doi:10.1101/2020.11.13.381434

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. 2008. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* **105**:8309–8314.

Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, Geeting KP, Gnirke A, Melnikov A, McKenna D, Stamenova EK, Lander ES, Aiden EL. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A* **112**:E6456–65.

Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C, Brown GD, Marshall A, Flicek P, Odom DT. 2012. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148**:335–348.

Schwartz YB, Kahn TG, Pirrotta V. 2005. Characteristic Low Density and Shear Sensitivity of Cross-Linked Chromatin Containing Polycomb Complexes. https://journals.asm.org/doi/abs/10.1128/mcb.25.1.432-439.2005

Silva MCC, Powell S, Ladstätter S, Gassler J, Stocsits R, Tedeschi A, Peters J-M, Tachibana K. 2020. Wapl releases Scc1-cohesin and regulates chromosome structure and segregation in mouse oocytes. *J Cell Biol* **219**. doi:10.1083/jcb.201906100

Stigler J, Çamdere GÖ, Koshland DE, Greene EC. 2016. Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. *Cell Rep* **15**:988–998.

Tedeschi A, Wutz G, Huet S, Jaritz M, Wuensche A, Schirghuber E, Davidson IF, Tang W, Cisneros DA, Bhaskara V, Nishiyama T, Vaziri A, Wutz A, Ellenberg J, Peters J-M. 2013. Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature* **501**:564–568.

Teytelman L, Thurtle DM, Rine J, van Oudenaarden A. 2013. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proceedings of the National Academy of Sciences* **110**:18602–18607.

Valton AL, Venev SV, Mair B, Khokhar E, Tong AHY. 2021. A cohesin traffic pattern genetically linked to gene regulation. *bioRxiv*.

van den Berg DLC, Azzarelli R, Oishi K, Martynoga B, Urbán N, Dekkers DHW, Demmers JA, Guillemot F. 2017. Nipbl Interacts with Zfp609 and the Integrator Complex to Regulate Cortical Neuron Migration. *Neuron* **93**:348–361.

Vian L, Pękowska A, Rao SSP, Kieffer-Kwon K-R, Jung S, Baranello L, Huang S-C, El Khattabi L, Dose M, Pruett N, Sanborn AL, Canela A, Maman Y, Oksanen A, Resch W, Li X, Lee B, Kovalchuk AL, Tang Z, Nelson S, Di Pierro M, Cheng RR, Machol I, St Hilaire BG, Durand NC, Shamim MS, Stamenova EK, Onuchic JN, Ruan Y, Nussenzweig A, Levens D, Aiden EL, Casellas R. 2018. The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* **173**:1165–1178.e20.

Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM. 1998. Force and velocity measured for single molecules of RNA polymerase. *Science* **282**:902–907.

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters J-M. 2006. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* **16**:863–874.

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters J-M. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**:796–801.

Wood C, Tonegawa S. 1983. Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation: implications for the joining

mechanism. *Proc Natl Acad Sci U S A* **80**:3030–3034.
Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, Schoenfelder S, Jessberger G, Muhar M, Hossain MJ, Walther N, Koch B, Kueblbeck M, Ellenberg J, Zuber J, Fraser P, Peters J-M. 2017. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* **36**:3573–3599.
Yamamoto T, Schiessel H. 2017. Osmotic mechanism of the loop extrusion process. *Phys Rev E* **96**:030402.
Yang JH, Brandão HB, Hansen AS. 2021. DNA double-strand break end synapsis by DNA loop extrusion. *bioRxiv*.
Yatskevich S, Rhodes J, Nasmyth K. 2019. Organization of Chromosomal DNA by SMC Complexes. *Annu Rev Genet* **53**:445–482.
You Q, Cheng AY, Gu X, Harada BT, Yu M, Wu T, Ren B, Ouyang Z, He C. 2021. Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution. *Nat Biotechnol* **39**:225–235.
Zhang D, Huang P, Sharma M, Keller CA, Giardine B, Zhang H, Gilgenast TG, Phillips-Cremins JE, Hardison RC, Blobel GA. 2020. Alteration of genome folding via contact domain boundary insertion. *Nat Genet* **52**:1076–1087.
Zhang H, Lam J, Zhang D, Lan Y, Vermunt MW, Keller CA, Giardine B, Hardison RC, Blobel GA. 2021. CTCF and transcription influence chromatin structure re-configuration after mitosis. *Nat Commun* **12**:5157.
Zhang S, Uebelmesser N, Josipovic N, Forte G, Slotman JA, Chiang M, Gothe H, Gusmao EG, Becker C, Altmueller J, Houtsmuller AB, Roukos V, Wendt KS, Marenduzzo D, Papantonis A. 2021. RNA Polymerase II is Required for Spatial Chromatin Reorganization Following Exit from Mitosis. *Science Advances* **7**:eabg8205.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137.
Zhang Y, Zhang X, Ba Z, Liang Z, Dring EW, Hu H, Lou J, Kyritsis N, Zurita J, Shamim MS, Presser Aiden A, Lieberman Aiden E, Alt FW. 2019. The fundamental role of chromatin loop extrusion in physiological V(D)J recombination. *Nature* **573**:600–604.
Zhu Y, Denholtz M, Lu H, Murre C. 2021. Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. *Genes Dev* **35**:65–81.
Zuin J, Franke V, van Ijcken WFJ, van der Sloot A, Krantz ID, van der Reijden MIJA, Nakato R, Lenhard B, Wendt KS. 2014. A cohesin-independent role for NIPBL at promoters provides insights in CdLS. *PLoS Genet* **10**:e1004153.