Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes

Jennifer M. Luppino, Daniel S. Park, Son C. Nguyen, Yemin Lan, Zhuxuan Xu, Rebecca Yunker and Eric F. Joyce

The human genome can be segmented into topologically associating domains (TADs), which have been proposed to spatially sequester genes and regulatory elements through chromatin looping. Interactions between TADs have also been suggested, presumably because of variable boundary positions across individual cells. However, the nature, extent and consequence of these dynamic boundaries remain unclear. Here, we combine high-resolution imaging with Oligopaint technology to quantify the interaction frequencies across both weak and strong boundaries. We find that chromatin intermingling across population-defined boundaries is widespread but that the extent of permissibility is locus-specific. Cohesin depletion, which abolishes domain formation at the population level, does not induce ectopic interactions but instead reduces interactions across all boundaries tested. In contrast, WAPL or CTCF depletion increases inter-domain contacts in a cohesin-dependent manner. Reduced chromatin intermingling due to cohesin loss affects the topology and transcriptional bursting frequencies of genes near boundaries. We propose that cohesin occasionally bypasses boundaries to promote incorporation of boundary-proximal genes into neighboring domains.

Chromosomes are hierarchically folded within the nuclei of eukaryotic cells. At the largest scale, chromosomes are packaged into spatially distinct chromosome territories. Chromosome conformation capture-based methods, including Hi-C, have further subdivided the genome into compartments, domains and chromatin loops. Domains are typically defined from population-averaged chromatin interactions and have been proposed to function as regulatory units that delimit the genomic extent from population-averaged chromatin interactions and have been proposed to function as regulatory units that delimit the genomic extent. Consistent with this model, population-averaged domains show remarkable coordination with gene regulation. In addition, deletion of CTCF binding sites at the boundary of domains can result in ectopic transcriptional activation of one or more flanking genes via formation of a functional interaction across the deleted boundary. Together, these observations have led to the hypothesis that CTCF dimerization halts cohesin-mediated chromatin exclusion from the boundaries of topological domains. Indeed, CTCF and cohesin colocalize on chromatin at the anchors of loops and the boundaries of domains, and depletion of either greatly perturbs loop and domain formation at the population-average level.

Importantly, the insulation of gene expression programs is believed to function via the spatial separation of population-defined domains. However, recent single-cell Hi-C datasets and imaging-based approaches have suggested extensive heterogeneity in domain organization at the single-cell level. In particular, super-resolution microscopy studies in which DNA is traced via sequential FISH have shown that domain-like structures exist in individual fixed cells, with large cell-to-cell heterogeneity in their boundary positions. Whereas this would suggest that population-defined domains represent an ensemble of several chromatin configurations, these approaches have only been tested at a small number of loci across the human genome. Therefore, it remains unclear whether boundary variability is a widespread feature across the genome and whether the extent of variability differs across different chromatin types and boundary strengths. In addition, the way in which these heterogeneous interactions impact gene regulation remains unknown.

To address these issues, we generated Oligopaint-based FISH probes to precisely target population-defined domains in single cells. We used a combination of high- and super-resolution microscopy to quantify the frequency and extent of domain intermingling across regions of different length scales, chromatin types and boundary strengths. These approaches revealed extensive heterogeneity across many population-defined boundaries with locus-specific differences in the extent of permissibility. Furthermore, we found that interactions across these boundaries were facilitated by the cohesin complex and antagonized by WAPL and CTCF. Reduced chromatin intermingling due to cohesin loss affects domain intermingling and transcriptional bursting frequencies of genes close to architectural boundaries. Therefore, we propose that, rather than by strictly forming spatially insulated domains, cohesin frequently bypasses population-defined boundaries to ensure proper regulation of boundary-proximal genes.

Results
Boundary permissibility is a widespread feature of the human genome. To measure the frequency and extent of interactions
across architectural boundaries at the single-allele level, we designed an Oligopaint FISH-based assay that tiles probes along population-defined domains (Fig. 1a)30–34. We applied our FISH assay in HCT-116, a human colorectal carcinoma cell line from which chromatin loops and domains have been previously defined on the basis of high-resolution Hi-C data. To rank-order boundary strengths between domains and subdomains, we determined insulation scores across the genome by using the mean contact frequency with a 25-kb sliding window35. Boundaries were further refined by local insulation score (IS) minima and colocalization of CTCF and RAD21, according to ENCODE chromatin immunoprecipitation sequencing (ChIP-seq) data sets36.

We designed Oligopaint libraries targeting a total of 17 domain pairs, which represented a range of gene densities, expression status, chromatin modifications and boundary strengths across six different chromosomes (Extended Data Fig. 1 and Supplementary Table 1). Cells were synchronized in G1 to avoid heterogeneity due to the cell cycle or the presence of sister chromatids (Extended Data Fig. 2a). We used custom 3D segmentation37 to trace the edges of each domain’s signal and generate a distribution of domain volumes across a minimum of 1,500 alleles per domain pair (Fig. 1b). The overlap volume per allele was normalized to the volume of each domain to control for the varying genomic lengths at the loci tested. If population-defined domains exist as spatially separate structures, we would expect little-to-no overlap between adjacent domains. This was indeed the case for 2–35% of alleles across all loci tested (Fig. 1c–e and Extended Data Fig. 3a–n). Thus, the majority of domains overlapped by up to 50–61%, which indicates a high level of intermingling differed in a locus-specific manner. Similar results were also observed in asynchronous cell populations, which indicates that this is not a feature that is specific to cells in G1 (Extended Data Fig. 2b).

To compare our FISH data to those of Hi-C, we plotted the frequency of domain contact to the IS of their intervening boundary. We found a good correlation between these two metrics (R² = 0.56; Fig. 1f), which suggests that Hi-C and our FISH assay are in agreement when comparing relative contact frequencies across different boundaries. Moreover, as the IS of the boundary can predict contact between domains by FISH, we reasoned that the majority of interactions most probably occur near the population-defined boundary. Indeed, when we subdivided upstream domains into three subdomains anchored by CTCF/RAD21 sites, the boundary-proximal regions exhibited the most contact and overlap with the downstream domain (Fig. 1g–j and Extended Data Fig. 4c–f).

Across all loci tested, the strongest and weakest boundaries exhibited an approximate twofold difference in their inter-domain contact (Fig. 1f). To measure interactions across a strong and weak boundary simultaneously, we labeled 3 regions of ~500-kb on chromosome 22 (Fig. 1k,l). As expected, overlap across the weak subdomain boundary occurred more frequently and to a greater extent than across the stronger domain boundary (Fig. 1m,n). Specifically, we observed almost twofold more contact across the weak boundary as compared to the strong boundary. This is remarkably similar to the approximately twofold genome-wide average increase in intra-domain contacts that was recently estimated from Hi-C data38. Surprisingly, we observed only a modest correlation (R² = 0.24) between domain overlap across the strong and weak boundaries on the same allele (Extended Data Fig. 2g), which suggests that interactions across each boundary occur independently from one another. Together, these data suggest that boundary permissibility is a widespread feature of the human genome but that the extent of permissibility differs in a locus-specific manner.

Variable interactions across boundaries occur independently from intra-domain compaction. Our data suggest that interactions across population-defined boundaries are frequent and extensive events. To validate our results at a higher resolution and determine the impact of regional compaction on these interactions, we applied super-resolution microscopy using 3D stochastic optical reconstruction (3D-STORM) to visualize adjacent domains with <50 nm error in their localization and <5% error in their physical sizes39. We chose two pairs of consecutive domains that flanked relatively weak (IS = 142) and strong (IS = 103) boundaries located on chromosomes 12 and 22, respectively (Fig. 2a). These domain pairs were also chosen on the basis of their differing chromatin modification landscapes (Extended Data Fig. 1). In particular, the shared boundary between the domain pair on chromosome 12 is contained within an active A compartment, whereas the domain pair on chromosome 22 is mostly contained within a silent B compartment.

We observed a wide diversity in the shape, volume, and particle density of each domain across the cell population (Fig. 2b,c and Extended Data Fig. 4a). Whereas alleles within the same cell showed a moderate correlation in domain volume, no such correlation was observed between neighboring domains across the same chromosome (Extended Data Fig. 2b–e), which indicates that stochastic compaction rates are intrinsic to each domain.

The spatial overlap between neighboring domains recapitulated the results observed by diffraction-limited microscopy. Notably, spatially separated alleles were observed twice as frequently across the strong boundary on chromosome 22 (34%) versus the relatively weaker boundary on chromosome 12 (15%) (Fig. 2d). However, the majority of chromosomes (66–85%) at both loci showed some level of intermingling between the probed domain pairs. Neighboring domains overlapped by up to 50–61%, which indicates a high level of inter-domain interactions across the cell population. Importantly, the amount of spatial overlap between neighboring domains did not correlate with the particle density of either domain (Fig. 2e and Extended Data Fig. 4f). This indicates that interactions between domains are not a direct result of regional decompaction. These results align with recent reports of heterogeneous domain intermingling19–21, and together suggest that an independent process is driving chromatin folding in a region-specific manner.

Cohesin promotes interactions within and across domain boundaries. To test the contribution of cohesin to inter-domain interactions, we carried out an acute depletion of RAD21, a core component of the cohesin ring, via indole-3-acetic acid (auxin)-inducible degradation (AID) in HCT-116 cells. Previous work using Hi-C analysis has shown that RAD21 degradation leads to a complete loss of loop domains (Fig. 3a), which can be attributed to a randomization of boundary positions at the single-cell level by using sequential FISH analysis37. To determine whether this randomization was associated with ectopic interactions across boundaries, we repeated our FISH assay in synchronized G1 cells that were treated with auxin for 6 h. This treatment resulted in a greater than 95% reduction in chromatin-bound RAD21 levels (Extended Data Fig. 5a–c).

Despite the loss of recurrent boundaries identified by Hi-C and FISH, all 17 domain pairs we assayed exhibited reduced contact frequencies following RAD21 degradation (Fig. 3b,c). Although cohesin loss reduced contact between some domain pairs more than others, these locus-specific differences did not seem to reflect the size or chromatin type of the domain pair being tested. Instead, there was a moderate but significant correlation between the fold change in contact frequency and the boundary strength before treatment (Fig. 3d). Although this cannot fully explain the locus-specific differences in effect size, it does suggest that weaker boundaries were more sensitive to cohesin loss. When overlap was observed in the absence of cohesin, domains intermixed significantly less frequently at all loci tested (P < 0.001; Fig. 3e,f and Extended Data Fig. 6a–u). We validated these findings using 3D-STORM, which revealed an increase in spatial separation between domains following cohesin loss across both a strong and weak boundary.
**Fig. 1 | Boundary permissibility is a widespread feature of the human genome.**

**a.** Oligopaint design to label two population-defined TADs. **b.** Representative image of neighboring domains at chr12:11.6–13.6 Mb with corresponding 3D image reconstruction. Dashed line represents nuclear edge. Scale bar, 5 μm (left) or 1 μm (zoomed images, right). **c.** Distribution of spatial overlap between the neighboring domains (D1 and D2) at chr3:44.2–47.55 Mb. Overlap normalized to the volume of the upstream domain. n = 1,642 chromosomes. IS, insulation score of intervening boundary. **d.** Distribution of spatial overlap between the neighboring domains (D1 and D2) at chr12:11.6–13.6 Mb, normalized to the volume of the upstream domain. n = 3,986 chromosomes. **e.** Distribution of spatial overlap between the neighboring domains (D1 and D2) at chr22:33.2–36.8 Mb, normalized to the volume of the upstream domain. n = 2,835 chromosomes. **f.** Contact frequency of neighboring domains by FISH as a function of their boundary IS by Hi-C (n = 17 boundaries). Each point represents the average of two biological replicates. **g.** Hi-C contact matrix of chr22:33.2–36.8 Mb and Oligopaint design corresponding to h-j. **h.** Representative FISH image illustrating interactions between each of the three subdomains (S1–S3) and the downstream domain D2. Scale bar, 1 μm. Corresponding 3D segmentation of FISH signals below each image. **i.** Cumulative distribution plot of spatial overlap between subdomains (S1 (n = 1,552); S2 (n = 1,660) and S3 (n = 1,822)) and D2, normalized to the volume of D2. ***P < 0.001, two-tailed Mann–Whitney U-test. **j.** Frequency of contact between each subdomain and D2 from data in i. **k.** Representative three-color FISH image of chr22:33.2–36.8 Mb and Oligopaint design corresponding to l-n. **l.** Distribution of spatial overlap across the strong domain boundary (green, n = 1,610) and weak subdomain boundary (purple, n = 1,644). Overlap normalized to the volume of the boundary-proximal subdomain (blue probe). ***P < 0.001, two-sided Fisher's exact test. **m.** Frequency of contact across the strong and weak boundary from data in n. ***P < 0.0001, two-sided Fisher's exact test.
WAPL and CTCF restrict cohesin-dependent interactions across domain boundaries. Loop extrusion models would predict that altering the processivity of cohesin could affect the extent of inter-domain interactions. Therefore, we next depleted regulators of cohesin to determine their effect on chromatin intermingling across domain boundaries. Cohesin is loaded onto chromatin by NIPBL, whereas the negative regulator of cohesin, WAPL, opens the cohesin ring to release it from DNA\cite{1}. We depleted NIPBL and WAPL using RNAi in HCT-116 cells to decrease and increase cohesin occupancy, respectively, and then measured the overlap between neighboring domains at two loci by FISH (Fig. 4a and Extended Data Fig. 8a–b). NIPBL-depleted chromatin showed significantly less contact and more spatial separation between domains, in a manner similar to that observed following RAD21 degradation. NIPBL knockdown also showed a more severe domain separation across the weak boundary on chromosome 12 as compared to the strong boundary on chromosome 22. By contrast, WAPL-depleted chromatin showed significantly more contact and greater overlap (P < 0.001) between the domains and to a similar extent at both loci (Fig. 4b–d). Double knockdown of NIPBL and WAPL phenocopied NIPBL depletion alone, which indicates that the increase in overlap was cohesin-dependent (Extended Data Fig. 8c,d).

We next depleted the insulator protein CTCF, which, like cohesin, is necessary for loops and domains at the population-average level\cite{2}. In contrast to cohesin loss, however, we observed significantly increased contact and spatial overlap at 5 out of 6 domain pairs tested (P < 0.001), in a manner similar to that seen following WAPL depletion (Fig. 4a,e,f and Extended Data Fig. 8e–k). Therefore, depletion of cohesin and CTCF show opposite effects on inter-domain interactions by FISH, despite their comparable loss of population-defined domains by Hi-C\cite{3,4}. In addition, in contrast to cohesin, the fold change in contact between neighboring domains was negatively correlated with the IS at the boundary before RNAi (R² = 0.28), such that stronger boundaries were more dependent on CTCF (Fig. 4g). Notably, all domain pairs tested had a CTCF site at their shared boundary (Extended Data Fig. 1). Auxin-mediated depletion of RAD21 following CTCF knockdown mimicked the phenotype of RAD21 depletion alone, which indicates that the increase in overlap was cohesin-dependent (Extended Data Fig. 8f–h–k). Taken together, these results indicate that CTCF and WAPL restrict but do not eliminate cohesin-dependent interactions across population-defined boundaries.
Cohesin alters the topological context of boundary-proximal genes. Chromatin loops and domains are implicated in the regulation of gene expression and previous work using nascent RNA-sequencing revealed that 4,196 genes were differentially expressed in HCT-116 cells following RAD21 degradation, albeit with relatively modest changes\(^1\) (Fig. 5a). Several differentially expressed genes (DEGs) are associated with RAD21/CTCF binding sites and are thus in close proximity to domain boundaries. Given that domains are less likely to interact in the absence of cohesin, we re-examined the spatial position of these genes relative to their neighboring population-defined domains in control and RAD21-depleted cells. This analysis was conducted for two DEGs, CREBL2 and MCM5, whose transcription start sites (TSSs) are within 125 kb and 50 kb of a domain boundary, respectively (Fig. 5b and Extended Data Fig. 9a).

We defined four topological configurations according to the position of the gene relative to either domain: (1) the gene interacts with its expected contact domain (domain maintenance); (2) the gene interacts with the neighboring domain (domain switching); (3) the gene no longer interacts with either domain (domain exclusion) or (4) the gene interacts simultaneously with both domains (domain sharing) (Fig. 5c). Although all four configurations occurred throughout the cell population, <1% of cells exhibited a domain-sharing configuration in which the gene interacted with both domains simultaneously. This indicates that stochastic domain intermingling is not due to complete domain merging but instead arises from the asymmetric incorporation of boundary-proximal chromatin with its neighboring regions, consistent with a shifting boundary position across individual cells. Indeed, the MCM5 gene interacted with the upstream portion of its own domain in 37% of alleles and interacted exclusively with the neighboring domain at a similar frequency (22%; Fig. 5d). Most commonly, the gene was spatially excluded from either domain (40%). Similar results were obtained for the boundary-proximal DEG CREBL2 on chromosome 12 (Extended Data Fig. 9b,c). These results further highlight the variable nature of domain boundaries and suggest that genes near boundaries are frequently located outside their expected population-defined topological context.

Following auxin treatment to deplete RAD21, the MCM5 locus was more frequently excluded from either domain (61% of alleles; Fig. 5d). A similar increase in spatial exclusion was found for CREBL2 (Extended Data Fig. 9c). At both loci, this was accompanied by a significant reduction in both domain maintenance and switching. Therefore, in the absence of cohesin, domain separation preferentially induces increased exclusion of these boundary-proximal genes from neighboring domains in a fraction of cells.

Boundary proximity correlates with gene expression changes following cohesin dysfunction. Nearly all regulatory elements and their target promoters are mapped within the same population-defined domain\(^11-14\). We therefore reasoned that the expression of genes near boundaries may be especially sensitive to cohesin loss due to increased exclusion from their neighboring regulatory domains. We plotted the distance between the TSS of each
expressed gene and the nearest domain boundary as a function of their log2(fold change) in expression using data from Rao et al. Overall, TSSs of expressed genes were found at varying distances from domain boundaries and spanned the full length of the domain size (Fig. 5e). However, DEGs with a higher fold change in expression were enriched near domain boundaries as compared to DEGs with a lower fold change or unchanged genes (P < 0.0025; Fig. 5e). For example, most DEGs with a greater than twofold change in expression following cohesin loss were located within 50 kb of a domain boundary, despite an average domain size of ~350 kb. Importantly, there was no difference in the size of domains harboring either DEGs or non-DEGs (Extended Data Fig. 10a). To control for the high gene density near domain boundaries, we calculated the percentage of all expressed genes in HCT-116 cells with greater than twofold change in expression following cohesin loss were located within 50 kb of a domain boundary (Fig. 5f). We note a similar enrichment for genes with >30% fold change following cohesin degradation (Fig. 5g). In total, 42% of expressed genes within 5 kb of a domain boundary are differentially regulated following depletion of RAD21 with nearly equal representation of upregulated and downregulated genes (Fig. 5h and Extended Data Fig. 10b).

Interestingly, although DEGs are also close to super-enhancers in a nonrandom manner as reported previously, we found no correlation between the fold change in expression and distance to the nearest super-enhancer (Extended Data Fig. 10c). Boundary-proximal DEGs also show no enrichment for specific biological processes or pathways (Extended Data Fig. 10d,e). Only 7% of DEGs within 5 kb of a domain boundary are classified as housekeeping genes, which suggests that the majority of boundary-proximal DEGs are probably cell-type-specific.

To determine whether this is a general signature of cohesin loss, we also analyzed data from lymphoblastoid cells derived from patients with NIPBL mutations, which cause a rare developmental disorder called Cornelia de Lange syndrome (CdLS). Previous studies have found decreased chromatin-bound RAD21 levels and a total of 1,500 genes that are recurrently differentially expressed across patient-derived lymphoblastoid cells. In a manner similar to that observed in DEGs following RAD21 degradation in HCT-116 cells, the extent of misexpression of CdLS-associated DEGs is correlated to their distance from a boundary (Extended Data Fig. 10f,g). Together, these results indicate that genes at the boundaries of domains are more likely to be differentially expressed and to a larger extent following cohesin dysfunction.

Cohesin alters the transcriptional bursting frequency of boundary-proximal genes. Cohesin loss may alter transcription of boundary-proximal genes either by influencing how frequently
...the gene is transcribed (transcription burst frequency) or how much RNA polymerase II is loaded during each burst (burst size). To determine at which level of transcriptional regulation cohesin functions, we performed intronic RNA FISH on six upregulated and eight downregulated boundary-proximal genes (Fig. 6a). RNA FISH analysis showed that all 14 genes exhibited bursty transcription, as calculated by the fraction of active alleles, and are positioned between 363 bp and 125 kb of a domain boundary (Fig. 6b,c).

With the exception of KLF4, we observed altered bursting frequency in all genes following RAD21 degradation with a large range of effect sizes (Fig. 6d). A few upregulated DEGs, including KLF4, exhibited moderate increases in burst size (Extended Data Fig. 10). By contrast, all eight downregulated genes exhibited significant decreases in burst frequency, but the burst size was largely unaffected. Overall, the shifts in bursting frequency across all 14 DEGs were consistent with their directionality and relative changes in expression, as shown by precision nuclear run-on sequencing (PRO-seq) with a correlation of $R^2 = 0.90$ (Fig. 6e). On the basis of these data, we propose that cohesin promotes proper gene expression at the level of transcriptional bursting frequency at genes near domain boundaries.

### Discussion

Subsequent to their discovery, early models suggested that TADs are functional regulons, which spatially separate into 3D structures to insulate gene regulation. Using Oligopaint FISH technology to precisely target population-defined domains, our data add to an emerging theme from recent single-cell-based assays that genome packaging is extremely dynamic and heterogeneous across a cell population. We find that, on average, ~45% of alleles show some degree of intermingling between adjacent population-defined domains. We also show that interactions are enriched near their shared boundaries with approximately twofold less contact across the stronger versus the weakest elements (Fig. 1f). Using three-color FISH, we show that boundary-proximal chromatin is asymmetrically incorporated with neighboring domains. Thus, our data are consistent with variable boundary positions between population-defined domains as reported by Bintu et al. (Fig. 6f). The simplicity of our domain-FISH assay allowed us to extend this conclusion to many loci of different chromatin types, and our data suggest that boundaries flanked by domains of the same or different compartments are similarly permissible. Although we cannot rule out the possibility that some boundaries in the genome remain invariant, our data...
suggest that interactions across population-defined boundaries are a widespread feature of the human genome.

The loop extrusion model, in which cohesin complexes extrude DNA until halted by a convergent pair of CTCF motifs, has been proposed to explain the formation of loops and domains at the population-average level\textsuperscript{12,14}. Indeed, we find that cohesin promotes intermingling within and between population-defined domains in single cells, as has been predicted from polymer simulations\textsuperscript{16}. We also find increased interactions between domains following WAPL or CTCF knockdown, which is consistent with their role in restricting cohesin-based chromatin extrusion\textsuperscript{12,14}. Combined with the dynamic nature of cohesin and CTCF binding\textsuperscript{12,14}, we propose that cohesin-mediated stochastic boundary bypass toggles boundary incorporation between neighboring domains across a single cell cycle (Fig. 6f).

Finally, we consider our results in the context of gene expression and the proposed role of TADs in transcriptional regulation. In particular, if population-defined domains contain the appropriate regulatory elements for the genes that lie within, why would cells permit such variability in boundary positions? This might be a consequence of cohesin processivity and CTCF binding dynamics, which, according to recent estimates\textsuperscript{1}, are too rapid to facilitate such precise and prolonged chromatin interactions. Alternatively, shifting boundary positions may offer advantages over an invariable system. We identified a signature of RAD21 loss in which genes near population-defined boundaries are more likely to be misexpressed, and to a greater extent, than genes that are not near such boundaries. We find a similar signature in cells derived from NIPBL-deficient CdLS patients, which suggests that this is a general feature of cohesin dysfunction.

It is not entirely clear why boundary-proximal genes would be more sensitive to cohesin loss. However, cohesin-mediated incorporation of boundary-proximal chromatins with either of its neighboring domains would ensure a high probability of contact between all portions of a regulatory domain over time. This would be especially important for boundary-proximal genes, as these genes may need to travel up to twice the maximum distance, compared to a gene near the TAD center, to contact a distal regulatory element. Indeed, increased exclusion of these genes from neighboring domains following cohesin loss could explain both downregulation and upregulation of DEGs if a boundary-proximal gene were looped out away from a distal enhancer or silencer, respectively.

**Fig. 6 | Cohesin alters the transcriptional bursting frequency of boundary-proximal genes.** a. Representative images of intronic RNA FISH to the HS3ST1 transcript before and after auxin treatment. Dashed lines represent nuclear edges. Scale bars, 5 μm. b. Scatterplot indicating the gene expression changes previously reported by PRO-seq\textsuperscript{8} and distance to nearest boundary for genes assayed by RNA FISH. The mean domain size denoted by a dashed line is 343.9 kb. c, Hi-C contact matrices of the loci surrounding GALNT5 and HS3ST1. Hi-C maps shown for HCT-116 cells before and after auxin treatment to degrade RAD21. d, Change in bursting frequency of each gene following auxin treatment by intronic RNA FISH. n = 227 chromosomes. An average of three biological replicates per gene is shown. e, Change in gene expression previously reported by PRO-seq\textsuperscript{8} versus change in bursting frequency detected by intronic RNA FISH ($R^2=0.9047$, two-sided Pearson correlation). $P<0.0001$, n = 14 boundaries. f, Model of single-cell variability in domain formation. Two architectural domains are depicted in green and magenta, with arrows indicating a boundary-proximal promoter in each domain. Colored rectangles represent the appropriate enhancer for each gene. Cohesin promotes variable boundary bypass such that the boundary-proximal chromatin is asymmetrically incorporated with the neighboring domains in a large fraction of cells. The boundary-proximal promoters thus alternate their contact with regulatory elements in their respective domains, which can result in a transcriptional burst. In the absence of cohesin, the boundary-proximal region is more often excluded from either domain, such that promoters in this region are less frequently exposed to their regulatory elements. This could explain both downregulation and upregulation of DEGs if a boundary-proximal gene were looped out away from a distal enhancer or silencer, respectively.
Articles

21. Splinter, E. et al. CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes Dev.* 20, 2349–2354 (2006).

22. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 284–290 (2009).

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

24. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51–56 (2017).

25. 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).

26. Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64 (2013).

27. Binta, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* 362, eaau1783 (2018).

28. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* 176, 1502–1515.e10 (2019).

29. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oct-oct-telo transition. *Nature* 544, 110–114 (2017).

30. Beliveau, B. J. et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* 6, 7147 (2015).

31. Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl Acad. Sci. USA* 109, 21301–21306 (2012).

32. Beliveau, B. J. et al. OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc. Natl Acad. Sci. USA* 115, E2183–E2192 (2018).

33. Beliveau, B. J., Apostolopoulos, N. & Wu, C. T. Visualizing genomes with Oligopaint FISH probes. *Curr. Protoc. Mol. Biol.* 105, 14.23.1–14.23.20 (2014).

34. Beliveau, B. J. et al. In situ super-resolution imaging of genomic DNA with OligoSTORM and OligoDNA-PAINT. *Methods Mol. Biol.* 1663, 231–252 (2017).

35. Shin, H. et al. TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res.* 44, e70 (2016).

36. ENCODE Project Consortium *An integrated encyclopedia of DNA elements in the human genome.* *Nature* 489, 57–74 (2012).

37. Ollion, J., Cochemecc, J., Loll, F., Escude, C. & Boudier, T. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics* 29, 1840–1841 (2013).

38. Chang, L.-H., Ghosh, S. & Noordermeer, D. TADs and their borders: free movement or building a wall? *J. Mol. Biol.* 432, 643–652 (2020).

39. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.* 15, 210–218 (2016).

40. Nasmuth, K. & Haering, C. H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 43, 525–558 (2009).

41. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet.* 29, 569–574 (2013).

42. Dorsett, D. & Krantz, I. D. On the molecular etiology of Cornelia de Lange syndrome. *Annu. N. Y. Acad. Sci.* 1151, 22–37 (2009).

43. Liu, J. et al. Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS Biol.* 7, e1000119 (2009).

44. Nir, G. et al. Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS Genet.* 14, e1007872 (2018).

45. Cardozo Gizzi, A. M. et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* 74, 212–222.e5 (2019).

46. Nuelber, J., Fudenberg, G., Imaeva, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extension and compartmental segregation. *Proc. Natl Acad. Sci. USA* 115, E6697–E6706 (2018).

47. Haahr, H. J. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* 169, 693–707.e14 (2017).

48. Thiecke, M. J. et al. Cohesin-dependent and independent mechanisms support chromosomal contacts between promoters and enhancers. Preprint at bioRxiv https://doi.org/10.1101/2020.02.10.941989 (2020).

49. Bartman, C. R., Hsu, S. C., Hsiung, C. C., Raj, A. & Blobel, G. A. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol Cell* 62, 237–247 (2016).

Published online

Received: 27 August 2019; Accepted: 18 May 2020; Online content

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Methods

TopDom insolation score calculation. Published Hi-C data were downloaded from the Gene Expression Omnibus database (GSE104334, GSE63525) and the 25-kb window contact matrix was extracted with the Juicer toolbox (v1.9.8, ‘dump observed’ option). TopDom (v0.2.3) was then used to identity topological domains and to define insolation scores from the contact matrix, with the optimization parameter ‘window.size’ set at ten (includes data within ten windows when computing local topological domains). Domains were further processed to include only those between 250 kb and 2 Mb in size.

Compartment analysis. To designate compartments in the HCT-116-RAD21-AID cell line, we determined eigenvectors from Hi-C data reported in Rao et al.9. We applied the eigenvector feature annotation package included in the Juicer software (https://github.com/aidenlab/juicer/wiki/Eigenvector) with KR normalization at 500-kb resolution.50 Although the sign of the eigenvector typically reflects the A or B compartment definition, we further confirmed compartment designation by comparing the eigenvector coordinates with ChIP–seq data for various histone modifications marking active and inactive chromatin from previously published studies.9

Oligopaint design and synthesis. To label domains in the HCT-116-RAD21-AID cell line, we first applied the TopDom TAD algorithm to Hi-C data reported in Rao et al.9 to define insolation scores. We then identified RAD21 and CTCF colocalized sites at the boundaries between domains from ChIP–seq data available on ENCODE (ENCSCR00885 and ENCSR000D100, respectively). We defined the domains and subdomain boundaries coordinates by the center of the corresponding RAD21 ChIP–seq peak. We then applied the OligoMiner design pipeline to design DNA FISH probes to the coordinates found in Supplementary Table 1 (ref.9).

Oligopaints were designed to have either 42 or 80 bases of homology with an average of 5 probes per kb and were purchased from CustomArray (42-mers) or Twist Bioscience (80-mers). Additional bridge probes were designed to the subdomains and boundary proximal genes to amplify their signal.19 Oligopaints were synthesized as described previously.24,25

Cell culture. HCT-116-RAD21-AID cells were obtained from the authors of Natsume et al.39. The cells were cultured in McCoy’s 5A medium supplemented with 10% fetal bovine serum, 2 mM l-glutamine, 100 U ml\(^{-1}\) penicillin, and 100 µg ml\(^{-1}\) streptomycin at 37 °C with 5% CO\(_2\). Cells were selected with 100 µg ml\(^{-1}\) G418 and 100 µg ml\(^{-1}\) HygroGold before experiments. To avoid heterogeneity due to the cell cycle, HCT-116-RAD21-AID cells were synchronized at the G1/S transition. First, to arrest cells in the S-phase, cells were grown in medium supplemented with 2 mM thymidine (Sigma-Aldrich T8195) for 12 h. Cells were then resuspended in fresh medium and allowed to grow for 12 h to exit S-phase. To arrest at the G1/S transition, cells were grown in medium supplemented with 400 µM mimosine (Sigma-Aldrich M025) for 12 h. Lastly, the medium was replaced with either 400 µM mimosine + 500 µM auxin (Sigma-Aldrich I5148)–supplemented medium to degrade RAD21 or 400 µM mimosine-supplemented medium alone as an untreated control; cells were incubated with or without auxin for 6 h and then collected for experiments. Synchronization was confirmed by immunofluorescence (Extended Data Fig. 2a).

RNAi. HCT-116 cells were cultured in McCoy’s 5A medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U ml\(^{-1}\) penicillin, and 100 µg ml\(^{-1}\) streptomycin at 37 °C with 5% CO\(_2\). The following siRNAs (Dharmacon) were used: non-targeting control, NIPBL, WAPL and CTFCF. siRNA sequences can be found in Supplementary Table 2. Duplex siRNAs were incubated for 20 min at room temperature (RT) with RNAiMAX transfection reagent (Thermo Fisher Scientific). The cells were then harvested for experiments.

DNA FISH. Cells were settled on poly(λ-lysine)-treated glass slides for 2 h, or uncapped high-pressure 22 × 22 mm coverslips for 6 h. Cells were then fixed to the slide or coverslip for 10 min with 4% formaldehyde in phosphate-buffered saline (PBS) with 0.1% Tween 20, followed by three washes in PBS for 5 min each wash. Slides and coverslips were stored in PBS at 4 °C until use.

For experiments imaged with widefield microscopy, FISH was performed on slides. Slides were warmed to RT in PBS for 10 min. Cells were permeabilized in 0.5% Triton-PBS for 15 min. Cells were then dehydrated in an ethanol row, consisting of 2-min consecutive incubations in 70%, 90% and 100% ethanol. The slides were then allowed to dry for about 2 min at RT. Slides were incubated for 5 min each in 2x SSC (0.3 M NaCl, 0.03 M sodium citrate and 0.1% Tween 20) and 2x SSC/50% formamide at RT, followed by a 1-h incubation in 2x SSC/50% formamide at 37 °C. Hybridization buffer containing primary Oligopaint probes, hybridization buffer (10% dextran sulfate, 2x SSC, 50% formamide and 4% polyvinylsulfonic acid (PVSA)), 500 µM mimosine was applied to slides, covered with a coverslip, and sealed with rubber cement. Fifty pmol of probe was used per 25 µl hybridization buffer. Slides were then denatured on a heat block in a water bath set at 80 °C for 30 min, then transferred to a humidified chamber and incubated overnight at 37 °C. The following day, the coverslips were removed and slides were washed in 2x SSC at 60 °C for 15 min, 2x SSC at RT for 10 min, and 0.2x SSC at RT for 10 min. To stain DNA slides were washed with Hoechst (1:10,000 in 2x SSC) for 5 min. Slides were then mounted in SlowFade Gold Antifade (Invitrogen). For experiments imaged by 3D-STORM, FISH was performed on coverslips as described above for slides, without DNA staining.

Immunofluorescence. Slides were prepared as for DNA FISH. Cells were permeabilized in 0.1% Triton-PBS for 15 min, then washed three times in PBS-T (PBS with 0.1% Tween 20) for 10 min each wash. Proteins were blocked in 1% bovine serum albumin (BSA) in PBS-T for 1 h at RT. Primary antibodies diluted in 1% BSA-PBS-T were added to the slide, covered with a coverslip, and sealed with rubber cement. Slides were transferred to a humidified chamber and incubated overnight at 4 °C. The following day, slides were washed three times in PBS-T for 10 min each wash. Secondary antibody was diluted in 1% BSA-PBS-T and added to the slide, which was covered with a coverslip and sealed with rubber cement. Slides were transferred to a humidified chamber and incubated at RT for 2 h. Slides were then washed twice in PBS-T for 10 min, and once in PBS for 10 min. To stain the DNA, slides were washed with Hoechst (1:10,000 in 2x SSC) for 5 min. Slides were then mounted in SlowFade Gold Antifade Mountant (Invitrogen).

Widefield microscopy, image processing and data analysis. Images were acquired on a Leica widefield microscopy microscope, using a 1.4 NA 63 oil-immersion objective (Leica) and Andor iXon Ultra emCCD camera. All images were deconvolved with Huygens Essential v18.10 (Scientific Volume Imaging), using the CMLSE algorithm, with a signal to noise ratio of either 20 or 40, and 40 iterations (DNA FISH) or signal to noise ratio of 40 and 2 iterations (DNA stain). The deconvolved images were segmented and measured using a modified version of the TANGO 3D-segmentation plug-in for ImageJ28,29. Edges of nuclei and FISH signals were segmented using a Hysteresis-based algorithm. Contact between signals was defined by two objects with ≥500 nm voxel colocalization.

3D-stochastic optical reconstruction microscopy imaging. 3D-STORM images were acquired on a Bruker Vutara 352 super-resolution microscope with an Olympus 60x/1.2 NA WD objective and Hamamatsu ORCA Flash 4.0 v3 sCMOS camera. The 640 nm and 561 nm lasers were used to acquire images for TADs labeled with Alexa Fluor 647 and CF568, respectively. 3D-STORM imaging buffer contained 10% glucose, 2x SSC, 0.05 M Tris, 2% glucose oxidase solution, and 1% mercaptatoethanol. The glucose oxidase solution consisted of 20 mg ml\(^{-1}\) glucose oxidase and 200 µM potassium dichromate from bovine liver dissolved in buffer (50 mM Tris and 10 mM NaCl).

Fields of view were selected by widefield microscopy, such that each nucleus contained two distinct pairs of TADs. Z-stacks were determined, such that both homologs were within the imaged space, and ranged from 3.6–9.6 µm. Localization was then recorded in 0.1 µm steps; 150 frames were recorded per z-step, and the z-stack was cycled through 3–4 times. Imaging of the second channel was carried out sequentially, with the Alexa Fluor 647 probe image first. Imaging of the CF568 fluorophore was supplemented with 0.5% power of the 405 nm laser at the second to last cycle. Localization of the fluorophores was carried out using the B-SPLINE PSF interpolation spline.

Images were further filtered for localizations with <20 nm and ≤30 nm radial precision for Alexa Fluor 647 and CF568, respectively, and <60 nm and ≤80 nm axial precision for Alexa Fluor 647 and CF568, respectively. To define the largest clusters of signal, we applied the DBScan algorithm with 0.250 µm maximum particle distance, 45 minimum particles to form cluster and 0.250 µm hull alpha shape radius.

RNA FISH. RNA FISH probes were designed as either custom Stellaris FISH probes or as Oligopaint probes using oligo pools (OPOols) from Integrated DNA Technologies. RNA FISH to the CHFP gene was performed using both probe designs and these provided comparable results. Custom Stellaris FISH probes were designed by utilizing the Stellaris DNA FISH Probe Designer 4.2 (Research Technologies). Probes to introns in the CREBL2, CHFP and GOF15 genes were synthesized with Qsuar 670. The remaining transcripts were probed with RNA FISH Oligopaint probes designed with a similar OligoMiner pipeline used for DNA FISH, with the exception of using the default 36 to 41 nucleotide length
range. Cells were seeded and fixed onto Lab-Tek II 8-well chambered coverglass dishes (Thermo Fisher Scientific), using the same fixation procedures as are used in DNA FISH. After fixation, cells were permeabilized overnight in 300 μl of 70% ethanol containing 2% SDS. With Stellaris FISH probes, cells were washed the next day in 2x SSC containing 10% formamide for 5 min, and then probes were hybridized with cells in a 200 μl mixture containing 10% dextran sulfate, 2x SSC, 10% formamide, 4% PVSA, 2% SDS, 2.8 mM dNTPs and 15.6 mM of probe. Next, 300 μl of mineral oil was added to each well to prevent evaporation, and the dishes were placed in a humidified chamber at 37°C overnight. The next day, cells were washed in 2x SSC containing 10% formamide twice for 30 min each wash, with the last wash containing 0.1 μg/ml of Hoechst 33342 stain. Cells were then washed with 2x SSC (no formamide) for 5 min and mounted in 300 μl of buffer containing glucose oxidase (37 μg ml⁻¹), catalase (100 μg ml⁻¹), 2x SSC, 0.4% glucose, and 10 mM Tris-HCl before imaging.

For RNA FISH with Oligopaint probes, cells were treated with the same overnight permeabilization step, followed by washes the next day in 2x SSC containing 10% formamide for 5 min. Probes were hybridized with cells in a 200 μl mixture containing 10% dextran sulfate, 2x SSC, 50% formamide, 4% PVSA, 2% SDS, 2.8 mM dNTPs and 31.2 nM of probe. An incubation step at 60°C for 3 min was done before the overnight incubation at 37°C, as suggested by Kishi et al.54. The next day, cells were washed once with RT 2x SSC containing 10% formamide for 5 min. Probes were hybridized with cells in a 200 μl mixture containing 10% dextran sulfate, 2x SSC, 10% formamide, 4% PVSA, 2% SDS, 2.8 mM dNTPs and 50 nM of secondary probe. Cells were washed with the hybridization mix at 37°C for 1 h, then washed three times for 5 minutes each with 2x SSC pre-warmed at 37°C, with the first wash containing 0.1 μg/ml of Hoechst 33342 stain. Cells were then mounted with 100 μl of SlowFade Gold Antifade Mountant before imaging.

Subcellular protein fractionation and western blots. Cells were trypsinized and resuspended in fresh medium, washed once in cold Dulbecco’s PBS, and then centrifuged at 1,200 × g for 5 min at 4°C. The cell pellet was then either processed to extract whole cell lysate. The sample was nutated for 30 min at 4 °C, and then centrifuged at 16,000 × g for 20 min at 4 °C. Supernatant containing protein was collected and stored at −80°C.

For western blots, protein was mixed with NuPAGE LDS sample buffer and sample reducing agent (Thermo Fisher Scientific), denatured at 70°C for 10 min, then cooled on ice. Benzamidine was added to the sample (final concentration 8.3 U ml⁻¹), followed by a 15-min incubation at 37°C. Then, 30 μl of each sample was run on a Mini-PROTEAN TGX Stain-Free Precast Gels (Bio-Rad) for 25–40 min at 35 mA per gel. The gel was then activated on a ChemiDoc MP Imaging System (Bio-Rad) for 5 min. Protein was then transferred to 0.2 μm nitrocellulose filter at 100 V for 45 min. The nitrocellulose filter was then washed twice in TBS (150 mM NaCl, 20 mM Tris) for 5 min, and blocked in 5% milk in TBS-T (TBS with 0.05% Tween 20) for 30 min. The nitrocellulose filter was then incubated again in TBS-T, then incubated with primary antibody diluted in 5% milk in TBS-T overnight at 4°C. The following day, the nitrocellulose filter was washed twice in TBS-T for 5 min each wash, then incubated with secondary antibodies conjugated in 5% milk in TBS-T for 1 h at RT. The nitrocellulose filter was then washed twice in TBS-T for 15 min each wash, followed by a final 15-min wash in TBS. For blotting with secondary antibodies conjugated to horseradish peroxidase (HRP), a stain-free image was acquired then the blot was incubated in a 1:1 mixture of Clarity Western ECL Substrate reagents (Bio-Rad). Blots were then imaged on a ChemiDoc MP Imaging System and analyzed with Bio-Rad Image Lab software (v5.2.1).

Antibodies. Immunofluorescence was performed using the following primary antibodies: RAD21 (Santa Cruz sc-166973, 1:100), PCNA (Santa Cruz sc-56, 1:100) and CENPF (Novus Biologicals NB500-101, 1:100). Secondary antibodies used were goat anti-rabbit (Jackson ImmunoResearch 111-165-003, 1:200) and sheep anti-mouse (Jackson ImmunoResearch 505-605-003, 1:100).

Western blots were performed with the following primary antibodies: RAD21 (Abcam ab992, 1:500 or 1:1000), NIPBL (Santa Cruz sc-374625, 1:4000), WAPL (Santa Cruz sc-365189, 1:250), alpha tubulin (Sigma T6074, 1:1000), histone H3 (Abcam ab1791, 1:40,000) and CTCF (Santa Cruz sc-271474, 1:500). Secondary antibodies used were as follows: goat anti-rabbit (Jackson ImmunoResearch 111-165-003, 1:3,700 – 1:3,100), goat anti-mouse (Jackson ImmunoResearch 115-545-003, 1:3,700 – 1:3,100), anti-mouse IgG HRP-linked antibody (Cell Signaling Technologies #7076, 1:5,000) and anti-rabbit IgG HRP-linked antibody (Cell Signaling Technologies #7074, 1:5,000).
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Genomic landscapes at Oligopaint target regions. Genomic profiles of loci imaged by FISH. Hi-C contact matrices visualized by Juicebox (v1.9.0)\textsuperscript{56}. Data from HCT-116-RAD21-AID\textsuperscript{1} cells. Solid and dashed lines indicate domains and subdomains, respectively. The gene density, eigenvectors, and insulation score (computed by the TopDom) are noted below. Insulation score computed prior to (solid) and following 6 hours of auxin treatment (dashed). Published ChIP-seq tracks\textsuperscript{8} depict protein binding and histone modifications in the HCT-116-RAD21-AID cell line prior to and following 6 hours of auxin treatment (-/+ Auxin). Genomic tracks visualized using Integrative Genomics Viewer\textsuperscript{57}.

\textsuperscript{56} Durand, N. C. et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. 3, 99–101 (2016).
\textsuperscript{57} Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26 (2011).
Extended Data Fig. 2 | Additional information related to Fig. 1. **a**, HCT-116-RAD21-AID cells were synchronized at the G1/S transition. Immunofluorescence for CENPF (green) to indicate cells in G2 and PCNA (gray) to mark cells in S phase. DNA (Hoescht stain) is shown in gray in first column. Dashed lines represent nuclear edges. Scale bar equals 5 μm. **b**, Cumulative frequency distribution of spatial overlap between neighboring domains on chr12:11.6Mb-13.6Mb (n = 716 chromosomes) and chr22:33.2Mb-36.8Mb (n = 1410 in asynchronous HCT-116 cells. Overlap normalized to the volume of the upstream domain. n > 716 chromosomes. **c**, Hi-C contact matrix and Oligopaint designs corresponding to (c-e). **d**, Representative FISH images of each subdomain and the downstream D2. Scale bar equals 1 μm. Corresponding 3D segmentation of FISH signals below each image. **e**, Cumulative distribution plot of spatial overlap between the subdomains (S1 (n = 1932); S2 (n = 2283); S3 (n = 1977) and D2, normalized to the volume of D2. ***P < 0.001, two-tailed Mann-Whitney test. **f**, Frequency of contact between each subdomain and D2 from data in **e**. Contact defined as > 500 nm^3 overlap. ****P < 0.0001, two-tailed Fisher’s exact test. **g**, Scatterplot of spatial overlap volume across the strong and weak boundaries on the same allele at the chr22:33.2-36.8 Mb locus. n = 1060 chromosomes. See Fig. 2e for corresponding Oligopaint design.
Extended Data Fig. 3 | See next page for caption.
Extended Data Fig. 3 | Boundary permissibility is a widespread feature of the human genome (additional loci).

a, Distribution of spatial overlap between neighboring domains (D1 & D2) at chr1:36.5-40.3 Mb. Overlap normalized to the volume of the upstream domain. n = 2294 chromosomes.

b, Distribution of spatial overlap between neighboring domains (D2 & D3) at chr1:36.5-40.3 Mb. Overlap normalized to the volume of the upstream domain. n = 2553 chromosomes.

c, Distribution of spatial overlap between neighboring domains (D3 & D4) at chr1:36.5-40.3 Mb. Overlap normalized to the volume of the upstream domain. n = 2502 chromosomes.

d, Distribution of spatial overlap between neighboring domains (D4 & D5) at chr1:36.5-40.3 Mb. Overlap normalized to the volume of the upstream domain. n = 9443 chromosomes.

e, Distribution of spatial overlap between neighboring domains (D1 & D2) at chr2:217.45-223 Mb. Overlap normalized to the volume of the upstream domain. n = 1850 chromosomes.

f, Distribution of spatial overlap between neighboring domains (D2 & D3) at chr2:217.45-223 Mb. Overlap normalized to the volume of the upstream domain. n = 1903 chromosomes.

g, Distribution of spatial overlap between neighboring domains (D2 & D3) at chr3:44.2-47.55 Mb. Overlap normalized to the volume of the upstream domain. n = 1657 chromosomes.

h, Distribution of spatial overlap between neighboring domains (D2 & D3) at chr12:11.6-13.6 Mb. Overlap normalized to the volume of the upstream domain. n = 1606 chromosomes.

i, Distribution of spatial overlap between neighboring domains (S1 & S2) at chr12:11.6-13.6 Mb. Overlap normalized to the volume of the upstream domain. n = 1479 chromosomes.

j, Distribution of spatial overlap between neighboring domains (S2 & S3) at chr12:11.6-13.6 Mb. Overlap normalized to the volume of the upstream domain. n = 1912 chromosomes.

k, Distribution of spatial overlap between neighboring domains (D1 & D2) at chr19:17.35-18.6 Mb. Overlap normalized to the volume of the upstream domain. n = 1406 chromosomes.

l, Distribution of spatial overlap between neighboring domains (S1 & S2) at chr22:33.2-36.8 Mb. Overlap normalized to the volume of the upstream domain. n = 1634 chromosomes.

m, Distribution of spatial overlap between neighboring domains (S2 & S3) at chr22:33.2-36.8 Mb. Overlap normalized to the volume of the upstream domain. n = 1640 chromosomes.

n, Distribution of spatial overlap between neighboring domains (S4 & S5) at chr22:33.2-36.8 Mb. Overlap normalized to the volume of the upstream domain. n = 1494 chromosomes.
Extended Data Fig. 4 | Additional information related to Fig. 2. a, Localization density per domain quantified from 3D-STORM images. Chr12.D1 (median = 0.0002707, n = 91); Chr12.D2 (median = 0.0002076, n = 91); Chr22.D1 (median = 0.0002376, n = 95); Chr22.D2 (median = 0.0002245, n = 95). b, Scatterplot depicting the relationship between domain volumes on the same chromosome by 3D-STORM on chr12:11.6Mb-13.6 Mb. n = 91 chromosomes. c, Scatterplot depicting the relationship between domain volumes between homologs by 3D-STORM on chr12:11.6Mb-13.6 Mb. n = 82 chromosomes. d, Scatterplot depicting the relationship between domain volumes on the same chromosome by 3D-STORM on chr22:33.2Mb-36.8 Mb. n = 95 chromosomes. e, Scatterplot depicting the relationship between domain volumes between homologs by 3D-STORM on chr22:33.2Mb-36.8 Mb. n = 86 chromosomes. f, Scatterplot of overlap volume (x-axis) versus domain density (y-axis) by 3D-STORM for the chr22:33.2Mb-36.8 Mb locus. n = 95 chromosomes.
Extended Data Fig. 5 | Confirmation and quantification of RaD21 degradation. a, Immunofluorescence for RaD21 (cyan). DNA (Hoescht stain) is shown in gray. Dashed lines represent nuclear edges. Scale bar equals 5 μm. b, Western blot to RaD21 protein in the chromatin-bound fraction of HCT-116-RaD21-AID cells with no auxin treatment (−) or following 6 hours of auxin treatment (+). Histone H3 as loading control. Protein was labeled using fluorescent secondary antibodies. c, Fluorescence quantification of RaD21 and H3 isolated from the chromatin-bound fraction of protein corresponding to blot in b using Image Lab v5.2.1. Protein intensity normalized to total protein per well (via stain-free technology) and presented as fraction of protein observed in untreated (− auxin) conditions; we observe a 96% reduction in chromatin-bound RaD21 following auxin treatment.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Cohesin promotes interactions across domain boundaries (additional loci). Hi-C contact matrix and corresponding cumulative distribution of spatial overlap between the neighboring domains prior to or after auxin treatment. Overlap normalized to the volume of the upstream domain. a, Map of chr1:36.5-40.3 Mb for b-e. b, D1 & D2 (-aux n = 2294, +aux n = 2175). c, D2 & D3 (-aux n = 2553, +aux n = 2353). d, D3 & D4 (-aux n = 2502, +aux n = 2488). e, D4 & D5 (-aux n = 9443, +aux n = 8695). f, Map of chr2:217.45-223 Mb for g-h. g, D1 & D2 (-aux n = 1850, +aux n = 1665). h, D2 & D3 (-aux n = 1903, +aux n = 1717). i, Map of chr3:44.2-47.55 Mb for j-k. j, D1 & D2 (-aux n = 1643, +aux n = 2186). k, D2 & D3 (-aux n = 1657, +aux n = 2181). l, Map of chr12:11.6-13.6 Mb for m-o. m, D2 & D3 (-aux n = 1606, +aux n = 1570). n, S1 & S2 (-aux n = 1479, +aux n = 1220). o, S2 & S3 (-aux n = 731, +aux n = 678). p, Map of chr19:17.35-18.6 Mb for q-s. q, D1 & D2 (-aux n = 1406, +aux n = 1419). r, Map of chr22:33.2-36.8 Mb for s-u. s, S1 & S2 (-aux n = 1634, +aux n = 1703). t, S2 & S3 (-aux n = 1640, +aux n = 1702). u, S4 & S5 (-aux n = 1494, +aux n = 1718). P < 0.001, two-tailed Mann-Whitney test for all domain pairs.
Extended Data Fig. 7 | Additional information related to Fig. 3.  

**a,** Minimum distances between localizations contained within each domain on chr12:11.6-13.6 Mb as measured from 3D-STORM data (prior to auxin: median = 0.008292 μm, n = 91; after auxin: median = 0.01234 μm, n = 76). P < 0.001, two-tailed Mann-Whitney test.  

**b,** Minimum distances between localizations contained within each domain on chr22:33.2Mb-36.8 Mb as measured from 3D-STORM data (prior to auxin: median = 0.009470 μm, n = 95; after auxin: median = 0.01563 μm, n = 105). P < 0.001, two-tailed Mann-Whitney test.  

**c,** Violin plots of domain volumes as measured from 3D-STORM data prior to [Chr12.D1 (median = 0.2629 μm³, n = 91); Chr12.D2 (median = 0.1830 μm³, n = 91); Chr22.D1 (median = 0.2749 μm³, n = 95); Chr22.D2 (median = 0.2320 μm³, n = 95)] or after auxin treatment [Chr12.D1 (median = 0.1671 μm³, n = 76); Chr22.D1 (median = 0.2284 μm³, n = 105); Chr22.D2 (median = 0.2213μm³, n = 105)]. Chr12.D1 P = 0.008; Chr12.D2 P = 0.368; Chr22.D1 P = 0.076, Chr22.D2 P = 0.907; two-tailed Mann-Whitney test.  

**d,** Violin plots of domain densities as measured from 3D-STORM data prior to [Chr12.D1 (median = 0.0002707, n = 91); Chr12.D2 (median = 0.0002076, n = 91); Chr22.D1 (median = 0.0002376, n = 95); Chr22.D2 (median = 0.0002245, n = 95)] or after auxin treatment [Chr12.D1 (median = 0.0003080, n = 76); Chr12.D2 (median = 0.0002385, n = 76); Chr22.D1 (median = 0.000265, n = 105); Chr22.D2 (median = 0.0002353, n = 105)]. Chr12.D1 P = 0.247; Chr12.D2 P = 0.014; Chr22.D1 P = 0.025, Chr22. D2 P = 0.602; two-tailed Mann-Whitney test.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | Additional information related to Fig. 4. a, Western blot of NIPBL, WAPL, and CTCF protein of HCT-116 cells following RNAi. H3 as loading control. b, Representative FISH images of neighboring domains on chr12:11.6Mb-13.6 Mb in RNAi control, NIPBL-, WAPL-, or CTCF-depleted cells. Dashed lines represent nuclear edges. Scale bar equals 5 μm (left) or 1 μm (zoomed images, below). c, Spatial overlap between neighboring domains on chr12:11.6Mb-13.6 Mb in control (n = 643), NIPBL- (n = 636), WAPL- (n = 819) or both NIPBL and WAPL- (n = 922) depleted cells. ***P < 0.001, two-tailed Mann-Whitney test. d, Spatial overlap between neighboring domains on chr22:33.2Mb-36.8 Mb in control (n = 1722), NIPBL- (n = 1440), WAPL- (n = 1769) or both NIPBL and WAPL- (n = 1762) depleted cells. Overlap normalized to the volume of the upstream domain. ***P < 0.001, two-tailed Mann-Whitney test. e, Western blot to RAD21 and CTCF in whole cell lysate of HCT-116 cells following RNAi to CTCF and/or 6 hours auxin to degrade RAD21. Alpha tubulin as loading control. Spatial overlap between neighboring domains across six loci (f-k). f, D4 and D5 on chr1:36.5-40.3 Mb in control (n = 2362) or CTCF- (n = 1611) depleted cells. ***P < 0.001, two-tailed Mann-Whitney test. g, D1 and D2 on chr2:217.45-223 Mb in control (n = 1484) or CTCF- (n = 1466) depleted cells. P = 0.604 (ns). two-tailed Mann-Whitney test. h, D2 and D3 on chr2:217.45-223 Mb in control or CTCF-depleted cells prior to (control n = 1145; CTCF n = 1133) or after auxin treatment (control n = 1039; CTCF n = 877). P < 0.001 (***) or P = 0.008 (**). two-tailed Mann-Whitney test. i, D1 and D2 on chr3:44.2-47.55 Mb in control or CTCF-depleted cells prior to (control n = 1272; CTCF n = 1215) or after auxin treatment (control n = 1254; CTCF n = 1090). P < 0.001 (***) or P = 0.996. two-tailed Mann-Whitney test. j, D1 and D2 on chr12:11.6Mb-13.6 Mb in control or CTCF-depleted cells prior to (control n = 1236; CTCF n = 1250) or after auxin treatment (control n = 1174; CTCF n = 1231). P < 0.001 (***) or P = 0.314, two-tailed Mann-Whitney test. k, D1 and D2 on chr22:33.2-36.8 Mb in control or CTCF-depleted cells prior to (control n = 1258; CTCF n = 1370) or after auxin treatment (control n = 1170; CTCF n = 1157). P < 0.001 (***) or P = 0.89, two-tailed Mann-Whitney test.
Extended Data Fig. 9 | Additional information related to Fig. 5. 

a. Hi-C contact matrix chr12:11.6-13.6 Mb and corresponding Oligopaint design for (b-c).

Blue represents the boundary proximal gene, CREBL2. 

b. Cartoon representations for three possible interactions between boundary-proximal genes and their neighboring domains: domain maintenance, switching, and exclusion (top). Representative images of three-color FISH to the chr12:11.6-13.6 Mb locus illustrating the three domain configurations (middle); scale bar equals 1 μm. Corresponding 3D segmentation of FISH signals below each image.

c. Frequencies of domain configurations at chr12:11.6-13.6 Mb prior to (n = 1074) or after auxin treatment (n = 979). P < 0.0001 (*** or P = 0.0046 (**), two-sided Fisher’s exact test.

---

**Extended Data Fig. 9** Additional information related to Fig. 5. 

**a.** Hi-C contact matrix chr12:11.6-13.6 Mb and corresponding Oligopaint design for (b-c).

Blue represents the boundary proximal gene, CREBL2. 

**b.** Cartoon representations for three possible interactions between boundary-proximal genes and their neighboring domains: domain maintenance, switching, and exclusion (top). Representative images of three-color FISH to the chr12:11.6-13.6 Mb locus illustrating the three domain configurations (middle); scale bar equals 1 μm. Corresponding 3D segmentation of FISH signals below each image.

**c.** Frequencies of domain configurations at chr12:11.6-13.6 Mb prior to (n = 1074) or after auxin treatment (n = 979). P < 0.0001 (*** or P = 0.0046 (**), two-sided Fisher’s exact test.
Extended Data Fig 10 | Additional information related to Fig. 6. a, Distribution of domain sizes that harbor either a significantly differentially expressed (median = 183734, n = 4196) or non-differentially expressed genes (median = 194447, n = 8026) in the HCt-116 cell line following auxin treatment. P = 0.102, two-tailed Mann-Whitney test. b, Scatterplot of log2(fold change) of HCt-116 differentially expressed genes (DEGs) versus the distance between their TSS and the center of the nearest domain boundary. c, Distance to nearest super enhancer defined by H3K27ac signal58 in HCt-116 cells for DEGs with >30% fold change in expression following auxin treatment. Genes were categorized by their proximity to a domain boundary (< 5 kb or > 5 kb away) and whether they were up or down regulated following auxin treatment. d, Gene ontology enrichment analysis for HCt-116 differentially expressed genes (> 30% fold change, n = 68) within 5 kb of a domain boundary. e, Gene ontology enrichment analysis for HCt-116 differentially expressed genes (> 30% fold change, n = 1593) not within 5 kb of a domain boundary. f, Scatterplot of [fold change] of differentially expressed genes (DEGs) associated with CdLS versus the distance between their TSS and the center of the nearest domain boundary in GM12878 cells. n = 1569, P = 0.0017, Spearman correlation. g, Fraction of genes that are up or down regulated in CdLS at binned distances from the nearest domain boundary. h, Fold change in burst volume by RNA FISH in HCt-116 cells following auxin treatment. Each dot represents the fold change in average burst volume per biological replicate; horizontal line indicates mean.

58. Hnisz, D., Day, D. S. & Young, R. A. Insulated neighborhoods: structural and functional units of mammalian gene control. Cell 167, 1188–1200 (2016).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Leica LAS-X version 3.3 was used for widefield image data collection and Vutara SRX version 6.02.05 was used for single molecule localization collection for super-resolution imaging.

Data analysis

Widefield image data were deconvolved with Huygens Essential (version 18.10) and analyzed using the TANGO 3D-segmentation plug-in for ImageJ (version 0.97). Super-resolution data was analyzed with Vutara SRX (version 6.02.05). Statistical analyses were performed with GraphPad Prism (version 8.3.0). Additional software include Juicebox (v1.9.8), TopDom (v0.0.2), and Bio-Rad Image Lab software (v5.2.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All relevant data are within the paper and its supporting information files. A data availability statement has been included.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

The number of cells scored were automatically counted. For each replicate, a minimum of 500 cells (DNA FISH) or 200 cells (RNA FISH) were imaged for diffraction-limited analysis and 50 for super-resolution STORM. These sample sizes were chosen to ensure the full spread of chromatin configurations and transcriptional state was sampled per experiment. The exact number of cells analyzed per experiment differed due to variable cell densities in the imaged fields. Figure legends detail the exact number of chromosomes analyzed per experiment.

Data exclusions

No data were excluded.

Replication

All experiments were repeated for at least two biological replicates. Biological replicates involved an independent isolation of cells including any relevant treatment whereas technical replicates were defined as independent FISH reactions to different slides from the same cell prep. All replicates of the same experiment yielded the same results and representative data and images for each experiment appear in the manuscript.

Randomization

Samples were allocated into groups based on genomic perturbations. Auxin-induced degradation was always accompanied by a untreated control. RNAi experiments were accompanied by treatment with a non-targeting control. Unless otherwise noted, these experiments were also performed on synchronized and arrested cells that were confirmed by IF and FACs to avoid heterogeneity due to the cell cycle. IF and Westerns were also performed for each experiment to confirm protein degradation.

Blinding

The data were analyzed simultaneously and computationally using identical pipelines, therefore blinding was not relevant. No manual scoring was used in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | Antibodies            |
| ☒   | Eukaryotic cell lines |
| ☒   | Palaeontology         |
| ☒   | Animals and other organisms |
| ☒   | Human research participants |
| ☒   | Clinical data         |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☐   | ChIP-seq              |
| ☒   | Flow cytometry        |
| ☒   | MRI-based neuroimaging |

Antibodies

Immunofluorescence was performed using the following primary antibodies: RAD21 (Santa Cruz sc-166973, 1:100), PCNA (Santa Cruz sc-56, 1:100), CENPF (Novus Biologicals NB500-101, 1:100). Secondary antibodies used: Goat anti-Rabbit (Jackson ImmunoResearch 111-165-003, 1:200), Sheep anti-Mouse (Jackson ImmunoResearch 505-605-003, 1:100). Western blots were performed with the following primary antibodies: RAD21 (Abcam ab992, lot GR3235714, 1:500 or 1:1000), NIPBL (Santa Cruz sc-374625, 1:400), WAPL (Santa Cruz sc-365189, 1:250), alpha tubulin (Sigma T6074, 1:1000), Histone H3 (Abcam ab1791, lot GR3181935, 1:40000), and CTCF (Santa Cruz sc-271474, 1:500). Secondary antibodies used: Goat anti-Rabbit (Jackson ImmunoResearch 111-165-003, 1:3:7000 – 1:3:10000), Goat anti-Mouse (Jackson ImmunoResearch 115-545-003, 1:3:7000 – 1:3:10000), Anti-mouse IgG HRP-linked Antibody (Cell Signaling Technologies #7076, lot 35, 1:5000), Anti-rabbit IgG HRP-linked Antibody (Cell Signaling Technologies #7074, lot 28, 1:5000).

Validation

All commercial antibodies were validated according to the manufacturer. All antibodies produced a clear band of the expected size by Western Blot and cellular localization by Immunofluorescence as expected. Furthermore, bands corresponding to the RAD21, NIPBL, CTCF, and WAPL antibodies disappeared by Western Blot and following RNAi to each target.
| Eukaryotic cell lines |
|----------------------|
| Policy information about cell lines |
| Cell line source(s) | HCT-116-RAD21-AID cells were obtained from Natsume et al. 2016. Parental HCT-116 cells were obtained by ATCC CCL-247 Colon Carcinoma; Human; Lot 70009735 |
| Authentication | Prior to all auxin-degron experiments, newly thawed cells were selected with 100 µg/ml G418 and 100 µg/ml HygroGold to ensure pure population. Further confirmation by IF to RAD21 was performed before and after treatment. |
| Mycoplasma contamination | Regular testing was performed using MycoAlert (Cambrex); all cell lines tested were negative. |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used. |