High-throughput Oligopaints identifies druggable 3D genome regulators

Daniel S. Park[1,2], **Son C. Nguyen**[1,2], **Randi Isenhart**[1,2], **Parisha P. Shah**[2,3], **Wonho Kim**[2,3], **R. Jordan Barnett**[1,2,4], **Aditi Chandra**[2,5,6], **Jennifer M. Luppino**[1,2], **Jailynn Harke**[1,2], **May Wai**[1,2], **Patrick J. Walsh**[1,2], **Richard J. Abdill**[1,2], **Rachel Yang**[2,3], **Yemin Lan**[2], **Sora Yoon**[2,5,6], **Rebecca Yunker**[1,2], **Masato T. Kanemaki**[7,8,9], **Golnaz Vahedi**[2,5,6], **Jennifer E. Phillips-Cremins**[1,2,4], **Rajan Jain**[2,3], **Eric F. Joyce**[1,2,*]

[1]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[2]Penn Epigenetics Institute, University of Pennsylvania, Philadelphia, PA, USA

[3]Department of Cell and Developmental Biology, Department of Medicine, Institute of Regenerative Medicine, Penn Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[4]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

[5]Institute for Immunology, Institute for Diabetes, Obesity and Metabolism, University of Pennsylvania, Philadelphia, PA, USA

[6]Abramson Family Cancer Center, University of Pennsylvania, Philadelphia, PA, USA

[7]Department of Chromosome Science, National Institute of Genetics, Research Organization of Information and Systems (ROIS), Shizuoka, Japan

[8]Department of Genetics, The Graduate University for Advanced Studies (SOKENDAI), Shizuoka, Japan

[9]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan

Summary

[*]Correspondence: erjoyce@upenn.edu (EJ).

CODE AVAILABILITY

CellProfiler pipeline for image segmentation is available at https://doi.org/10.5281/zenodo.7699078

The human genome functions as a three-dimensional (3D) chromatin polymer, driven by a complex collection of chromosome interactions[1–3]. Although the molecular rules governing these interactions are being quickly elucidated, relatively few proteins regulating this process have been identified. To address this gap, we developed HiDRO (high-throughput DNA or RNA labeling with optimized Oligopaints), an automated imaging pipeline that permits quantitative measurement of chromatin interactions in single cells across thousands of samples. By screening the human druggable genome, we identified over 300 factors that influence genome folding during interphase. Among these, 43 genes were validated as either increasing or decreasing interactions between topological associating domains (TADs). Our findings showed that genetic or chemical inhibition of the ubiquitous kinase GSK3A leads to increased long-range chromatin looping interactions in a genome-wide and cohesin-dependent manner. These results demonstrate the importance of GSK3A signaling in nuclear architecture and the utility of HiDRO for identifying novel mechanisms of spatial genome organization.

Keywords

fluorescence *in situ* hybridization; three-dimensional genome organization; topologically associating domains; cohesin; GSK3A

Introduction

The highly conserved cohesin complex has emerged as a major player in genome folding that can actively extrude DNA into chromatin loops, which further organize into extensive intrachromosomal units termed topologically associated domains (TADs)[4,5]. Despite these advances, relatively few proteins regulating this process have been identified. Indeed, our current abilities to link chromosome arrangement to function are limited. Foremost among our needs are advanced technologies that allow us to visualize chromosome arrangement at the single-nucleus level, in a high-throughput manner, to uncover the underlying mechanisms driving the spatial organization of the genome. This need becomes increasingly important as more evidence emerges of the roles chromosome positioning and long-range interactions play in gene regulation and development in essentially every species[4,6–10].

Here, we describe our development of HiDRO, which represents a novel combination of high-throughput fluorescence *in situ* hybridization (FISH) with Oligopaint probe technology to label single-copy genomic regions and nascent RNA species with high efficiency[11]. Using HiDRO, we performed a screen of 3,083 human genes encoding for enzymes that could potentially regulate the dynamic nature of chromatin folding. After high-content and high-resolution imaging, we measured 21 different parameters of nuclear organization and identified 43 factors in three distinct signaling pathways that affect the frequency and extent of interactions between TADs. In particular, we isolated the highly conserved and ubiquitously expressed kinase GSK3A, whose depletion enhances long-range interactions and reduces the insulation between domains. We find that this occurs in a genome-wide and cohesin-dependent manner to control the length of chromatin loops. Collectively, these results highlight a druggable enzyme underlying signaling cascades that can influence

genome folding, underscoring the broader utility of HiDRO to identify novel mechanisms that drive the spatial organization of the genome.

Results

Development of HiDRO

HiDRO builds on recent innovations in DNA and RNA FISH, combining an optimized version of array-based oligonucleotide probes (Oligopaints[11]) with a high-throughput FISH protocol (Fig. 1a). Increased scale and improved labeling efficiencies at lower concentrations are achieved by expansion of the homology length region and incorporation of amino-allyl nucleotides into each oligo during probe synthesis (Fig. 1a and Extended Data Fig. 1a,b; Methods). These modifications both increase the number of fluorophores per oligo and obviate the need for secondary labeling, streamlining the high-throughput FISH protocol.

The efficiency of our optimized protocol was initially tested against a library designed to target two consecutive TADs spanning ~2 Mb at chromosome 22:33.4–36.5 Mb (hereafter referred to as chr22 D1 and D2) in human HCT-116 cells (Fig. 1a). Following high-content imaging and automated image analysis, we achieved ~90% labeling of both D1 and D2 at a 25-fold lower concentration as compared to conventional Oligopaints (Fig. 1a and Extended Data Fig. 1a,b)[7,11]. We found comparable efficiencies across 40 additional loci representing a range of target sizes from 80 kb to >2 Mb (Extended Data Fig. 1c-p and Supplementary Table 1). Moreover, we observed identical bursting frequencies of three distinct genes when labeling nascent RNA using a modified version of HiDRO (Extended Data Fig. 1q,r), indicating that HiDRO can be used for both large-scale DNA and RNA FISH assays.

As proof of the concept that HiDRO can identify architectural proteins in a screening format, we focused on chr22 D1 and D2 domains that are separated by a relatively strong boundary (Fig. 1a)[4,7]. This boundary is also occupied by well-known architectural proteins cohesin and CTCF and is highly sensitive to their perturbation, as evidenced by both Hi-C and FISH[4,7]. We seeded >160 wells with siRNA against the cohesin loader NIPBL[12], the cohesin unloader WAPL[13,14], or a non-targeting control (Fig. 1b). Following 72 hours of siRNA treatment, we captured >150 nuclei in each well and used automated image analysis to calculate two independent metrics for inter-TAD interactions at each allele: (1) the center-to-center distance (CCD) between domains to measure compaction of the locus and (2) the normalized area of spatial overlap to measure the degree of inter-TAD interactions (Fig. 1b). In 100% of wells targeting either NIPBL or WAPL, we observed opposite and significant shifts ($z > 1.5$) for both CCD and spatial overlap, owing to a loss or gain in cohesin-mediated looping events, respectively[7] (Fig. 1c-e). No control siRNAs produced a significant shift in either direction, indicating that HiDRO can specifically and reliably detect altered chromatin folding in a bidirectional manner across a large number of samples.

Discovery of novel 3D genome regulators

To identify novel factors that alter genome organization, we scaled HiDRO to screen 3,083 genes comprising the human druggable genome[15,16], using an siRNA library consisting of enzymes that can be modulated using small-molecule compounds (Fig. 2a, Extended Data

Fig. 2a and Supplementary Table 3). A total of 7,392 FISH assays were performed in the primary screen with >1.5 million nuclei analyzed and an average of >600 alleles per gene knockdown. We identified 321 unique genes whose knockdown significantly altered either CCD or spatial overlap between the chr22 domains across two replicate plates (Fig. 2b-d). However, a much smaller number of genes mirrored the phenotypic signature of NIPBL and WAPL whose knockdown significantly altered both metrics (Fig. 2d,e and Extended Data Fig. 2b-e). This filtered our hits to 58 genes whose knockdown decreased (25) and increased (33) inter-TAD interactions, representing 1.87% of the druggable genome. Kinases (17, 29%), ubiquitin ligases (12, 21%), and proteases (8, 14%) were among the most common protein classes, indicating that a wide array of enzymes have the potential to regulate the dynamic nature of chromatin folding (Fig. 2e and Extended Data Fig. 2b-e).

In addition to our primary metrics, we also calculated 18 secondary parameters of genome organization (Fig. 2f and Extended Data Fig. 2f). These included measurements related to the size and shape of each domain and the nucleus itself, creating a multimodal dataset of nuclear organization for all 3,083 genes analyzed (Supplementary Table 4). A phenotypic tree indicated that 581 hits distributed across 48 distinct phenotypic profiles. This also showed that a minority of genes influenced nuclear organization with 2,502 genes displaying no phenotype across any parameter. Most hits (70%) that altered inter-TAD interactions exhibited no secondary phenotypes, including NIPBL and WAPL. Indeed, across all genes tested, our primary metrics did not correlate with any specific secondary phenotypes, suggesting inter-TAD interactions were not solely dependent on local or global changes in chromatin compaction, nuclear shape, or the cell cycle when G1 and G2 cells were separately analyzed (Extended Data Fig. 2f-j). However, the remaining 30% of hits harbored complex phenotypic profiles, indicating that altered chromatin folding can be associated with a wide range of nuclear architecture phenotypes.

Three major pathways influence genome folding

To identify the highest-confidence gene targets and further validate our primary hits, we performed a secondary screen with four siRNA duplexes de-pooled to target each hit in separate wells. This screen was conducted in triplicate using the same HiDRO protocol targeting the adjacent domains on chr22 (Fig. 2g and Extended Data Fig. 2k). For 74% (43 of 58) of our primary hits, knockdown with at least one duplex produced a significant shift in both CCD and spatial overlap (Supplementary Table 5). Ten genes validated with three or more duplexes, representing our highest-confidence hits (Fig. 2g,h). This included RNF2/RING1B, a subunit of the polycomb repressive complex 1 (PRC1), which mediates long-range CTCF-independent chromatin interactions in mice[17]. To our knowledge, the remaining nine genes have not been directly linked with chromatin folding prior to this study. Knockdown of each of the ten genes increased inter-TAD interactions, indicating that it might be easier to recover genes that normally antagonize rather than promote long-range interactions. Despite this similarity, comparative analysis of secondary metrics revealed distinct phenotypic signatures among the genes (Fig. 2h). For example, knockdown of several hits increased the size of each individual domain, consistent with local decompaction, whereas depletion of our top hit, glycogen synthase kinase 3 alpha (GSK3A), caused no notable changes in domain size or structure.

We used a protein–protein interaction database to perform functional network analysis[18] on all 43 validated hits and identified three interconnected protein classes: (1) ubiquitin ligases, (2) calcium signaling, and (3) GSK3 kinase signaling (Fig. 2i). Each of these protein classes was represented by one or more of our highest-confidence hits, further validating their involvement in chromatin folding (Fig. 2h,i). We chose one gene from each of the three major protein classes (GSK3A, FBXL14, and CALM1) and performed HiDRO at 13 total domain pairs scattered across the genome (Fig. 2j-l and Extended Data Fig. 2l-o). In addition to WAPL, we found that depletion of GSK3A, CALM1, or FBXL14 each significantly increased inter-TAD interactions at nearly all (12/13) boundaries tested, indicating that these hits likely represent general chromatin-folding factors (Fig. 2l and Extended Data Fig. 2m-o). Taken together, these results show that HiDRO can isolate novel pathways that influence chromatin folding.

GSK3A has noncanonical role in genome folding

We sought to further examine the role of our top hit GSK3A in chromatin folding. GSK3A and its closely related paralog GSK3B are ubiquitously expressed serine/threonine kinases with many, but not all, shared nuclear and cytoplasmic targets[19–23]. We first confirmed that GSK3A and GSK3B were expressed in HCT-116 cells and localized to both the nucleus and cytoplasm, as has been reported in other cell types (Fig. 3a and Extended Data Fig. 3a). Selective depletion of each paralog by siRNA was confirmed by immunofluorescence, and western blot (Fig. 3a and Extended Data Fig. 3a,b). Next, we used a conventional Oligopaint FISH protocol with higher-resolution three-dimensional imaging to confirm that GSK3A depletion increases spatial overlap between the chr22 domain pair (Fig. 3b-d). All four independent siRNA duplexes targeting GSK3A resulted in selective depletion and a significant increase in spatial overlap between domains (Extended Data Fig. 3c-e). However, we observed no significant change in spatial overlap between domains following GSK3B depletion, confirming that the regulation of inter-TAD interactions in HCT-116 cells is a unique property of the GSK3A paralog (Fig. 3c,d).

GSK3 signaling has a well-established role in suppressing WNT signaling through phosphorylation and degradation of beta-catenin[24,25]. However, because these kinases are typically redundant in WNT regulation[19,24,25], our specific isolation of GSK3A suggested this activity may be separate from its role in chromatin folding. To test this hypothesis in HCT-116 cells, we measured total beta-catenin protein levels and found no increase following depletion of either GSK3A or GSK3B alone (Extended Data Fig. 3b). Further, single knockdowns did not decrease the level of phosphorylated beta-catenin marked for degradation and did not activate the expression of two WNT target genes, AXIN2 and LGR5 (Fig. 3e and Extended Data Fig. 3b). Instead, only simultaneous depletion of both GSK3 paralogs led to a decrease in degraded phosphorylated beta-catenin levels and transcriptional activation of WNT target genes, similar to what has been reported for other cell types (Fig. 3e and Extended Data Fig. 3b)[19,26]. Co-depletion of GSK3A and GSK3B did not alter the extent of spatial overlap between domains by FISH as compared with GSK3A depletion alone (Fig. 3c,d). Taken together, these observations suggest that GSK3A has a unique role in chromatin folding separate from its canonical function in WNT signaling.

GSK3i increases inter-TAD interactions

To independently validate the role of GSK3 signaling in chromatin folding, we took advantage of our HiDRO platform to test five different chemical inhibitors that attenuate GSK3 kinase activity: three non-selective inhibitors (CHIR99021, Lithium, and LY2090314), and two recently described selective inhibitors of either GSK3A or GSK3B (BRD0705 or BRD3731, respectively)[26–28]. Each inhibitor was tested at four different concentrations across five replicate wells and administered for 24 hours (Fig. 3f, n=100 FISH assays total).

We found that all three non-selective inhibitors in addition to the GSK3A-selective inhibitor BRD0705 (GSK3Ai) led to an increase in spatial overlap between domains at various concentrations (Fig. 3g-k). In particular, lithium and LY2090314 treatment displayed a dose-dependent effect on inter-TAD interactions, with a similar effect size to that of GSK3A siRNA depletion (Fig. 3i,j). However, the GSK3B-selective inhibitor BRD3731 showed no change in spatial overlap, similar to our results above from selective siRNA knockdown (Fig. 3l). Furthermore, when using a conventional Oligopaint FISH protocol with higher-resolution three-dimensional imaging, we found significantly increased spatial overlap at four different loci across the genome following GSK3Ai treatment at a concentration (20 μM) below the threshold needed to activate WNT signaling[26] (Extended Data Fig. 3f,g). This supports the notion that chromatin folding represents a noncanonical function for GSK3A. These findings also confirm our screen results and demonstrate that the extent of inter-TAD interactions at multiple loci can be altered by acute GSK3A inhibition.

GSK3A restricts chromatin looping

To determine how GSK3A influences chromatin folding genome-wide, we performed *in situ* Hi-C following a 72-hour treatment with siRNA against GSK3A or a non-targeting control. After quality control, we annotated ~450 million valid unique cis-contacts per condition in each of two biological replicates (Extended Data Fig. 4a). When examining relative contact probabilities as a function of genomic distance in GSK3A-depleted cells, we found an overall decrease in short-range contacts (<500 kb) and increase in mid- to long-range contacts (>500 kb) (Fig. 4a and Extended Data Fig. 4b). We identified similar numbers and sizes of TADs in the two conditions (Extended Data Fig. 4c-e), indicating that GSK3A does not remove or redistribute TAD boundaries. However, pileup analysis of TADs across the entire genome showed that GSK3A depletion reduced insulation across their boundaries, validating our FISH data (Fig. 4b,c and Extended Data Fig. 4f). Similar effects were observed across subTADs (Extended Data Fig. 4c,e,g).

We also detected an overall reduction of 37% in the number of chromatin loops following GSK3A depletion (Fig. 4d,e). Specifically, GSK3A-depleted cells lost 6,010 loops, representing 50.5% of control loops, but also gained 1,596 unique loops (Fig. 4e,f). On average, gained loops were significantly longer than lost or retained loops (Fig. 4g). For comparison, we performed Hi-C following depletion of the cohesin unloaders WAPL or PDS5A, which restrict chromatin loop extension (Fig. 4c,e and Extended Data Fig. 4a,h,i)[1,14]. We found a comparable reduction in TAD boundary insulation, gain of longer chromatin loops, and reduction in shorter-range loops between all conditions (Fig. 4c,e-h

and Extended Data Fig. 4h-n). Interestingly, while GSK3A more closely mirrored PDS5A in terms of effect size and number of loops, GSK3A and WAPL gained loops at similar locations distinct from PDS5A (Fig. 4d,e and Extended Data Fig. 4j-n).

The majority (77%, 1229/1596) of gained loops in GSK3A depleted cells utilized at least one control loop anchor, an observation that is consistent with the extension of existing loops rather than formation of *de novo* loop anchors (Fig. 4d,g and Extended Data Fig. 4m,n). We therefore next isolated and analyzed architectural stripes, which are specific features of Hi-C maps indicative of loop-extrusion events[29,30]. Pileup analysis of control stripes across the genome showed that distal ends of stripes (500 kb to 2 Mb) were strengthened, whereas proximal ends (<500 kb) were weakened after GSK3A depletion (Fig. 4h and Extended Data Fig. 4l). A similar, albeit stronger effect was observed following depletion of WAPL and, to a much lesser extent, in PDS5A. This result is consistent with our chromatin loop analysis and further suggests a model in which GSK3A restricts chromatin looping to promote proper insulation between TADs to a similar extent to that of known anti-cohesin factors.

Cohesin-dependent GSK3A regulation

Given the similarities in phenotypes, we reasoned that GSK3A may influence chromatin looping in a cohesin-dependent manner. To test this, we depleted GSK3A before and after cohesin loss using a cell line in which the cohesin ring component RAD21 is fused to an auxin-inducible degron (AID)[4,31]. Following six hours of auxin treatment, we confirmed loss of RAD21 (Extended Data Fig. 5a) and observed decreased spatial overlap between domains at the chr22 locus by FISH, consistent with our previous findings[7] (Extended Data Fig. 5b,c). After GSK3A depletion, we observed increased spatial overlap in the absence of auxin treatment; however, this phenotype was completely lost following degradation of RAD21 (Extended Data Fig. 5b,c). Identical results were also found with chemical inhibition of GSK3A (Extended Data Fig. 5d,e). This indicates that the role of GSK3A in chromatin folding is dependent on cohesin activity.

We next evaluated the levels of two different cohesin ring components: RAD21 and SMC1A. No significant change in mRNA or total protein was observed following GSK3A depletion (Extended Data Fig. 6a,b). We also found no significant change in the recovery dynamics of nuclear RAD21 signal by fluorescence recovery after photobleaching (FRAP) (Extended Data Fig. 6c,d). However, chromatin fractionations followed by quantitative western blots showed >50% increase in chromatin-bound levels of both RAD21 and SMC1A (Extended Data Fig. 6e-g). A similar extent of cohesin accumulation on chromatin was observed following depletion of WAPL or PDS5A (Extended Data Fig. 6f,g). To determine where cohesin was accumulating on chromatin, we performed RAD21 ChIP-seq after GSK3A depletion, which revealed a significant increase in RAD21 levels at 4,069 sites across the genome as compared with control samples (Extended Data Fig. 6h-j). Very few (28) sites showed a significant loss of RAD21 signal (Extended Data Fig. 6h-j). Interestingly, as compared to retained RAD21 sites, sites that gained RAD21 signal exhibited near complete colocalization with CTCF (Extended Data Fig. 6h,i,k,l), suggesting that cohesin was accumulating at CTCF sites following GSK3A depletion.

To further examine the distribution of cohesin on chromatin, we performed RAD21 immunofluorescence. Notably, WAPL knockout cells harbor a distinct, thread-like RAD21 staining pattern referred to as "vermicelli"[1,32]. We quantified this aggregation using a nuclear RAD21 signal granularity index (see Methods) and found increased RAD21 aggregation when WAPL was partially depleted by either siRNA or auxin-mediated degradation using a custom cell line in which WAPL was biallelically fused to AID (Extended Data Fig. 7a,b). Combining WAPL siRNA and auxin-mediated degradation removed all detectable protein from cells by western and further increased RAD21 aggregation, consistent with previous WAPL knockout studies[1,32] (Extended Data Fig. 7b-d). We found that GSK3A depletion increased RAD21 aggregation and further enhanced this phenotype following partial WAPL degradation (Extended Data Fig. 7c,d). No additive effects were observed, however, when GSK3A depletion was combined with both WAPL siRNA and AID. Together, these results show that GSK3A can genetically interact with WAPL to influence cohesin levels on chromatin.

GSK3A promotes WAPL recruitment to chromatin

Our results above suggest that GSK3A was restricting chromatin looping and cohesin aggregation through WAPL regulation. Indeed, we found that chromatin-bound levels of WAPL were reduced by 73% despite unchanged mRNA and total protein levels following GSK3A depletion (Fig. 5a and Extended Data Fig. 8a). Co-immunoprecipitation on chromatin fractions also revealed a significantly weaker association (~50%) between WAPL and SMC1A following GSK3A depletion (Fig. 5b and Extended Data Fig. 8b,c). Thus, GSK3A is important for the recruitment or stability of WAPL on chromatin.

To determine if GSK3A interacts with WAPL or other cohesin subunits, we used a proximity labeling method known as TurboID[33]. We fused biotin ligase to GSK3A (GSK3A-BirA) and expressed the fusion protein in HCT-116 cells (Fig. 5c and Extended Data Fig. 8d). As a control, beta-catenin was significantly enriched in the GSK3A-BirA samples over a construct expressing biotin ligase alone (BirA), indicating that we could readily detect transient interactions with a known substrate of GSK3A. We found similar enrichment for several cohesin components, including core subunits, RAD21 and SMC1A, in addition to the cohesin unloading subcomplex components WAPL and PDS5A (Fig. 5c and Extended Data Fig. 8d). Notably, the cohesin loader NIPBL was not recovered in either replicate. Given the mutually exclusive nature of NIPBL and PDS5A/WAPL attachment to cohesin[34,35], this suggests that GSK3A interacts with the cohesin ring complex in a substrate-specific manner (Fig. 5d).

Previous studies have shown that increased cohesin activity via depletion of WAPL can offset phenotypes associated with partial loss of cohesin on chromatin[7,14,36,37]. We therefore used HiDRO to test whether co-depletion of GSK3A and NIPBL was sufficient to rescue chromatin misfolding by FISH. Depletion of NIPBL resulted in a ~30% loss of chromatin-bound RAD21 levels and a ~50% reduction in spatial overlap between domains by FISH, consistent with defective chromatin looping (Fig. 5e,f and Extended Data Fig. 8e,f)[7]. We found that depletion of GSK3A, WAPL, or PDS5A were each able to partially restore spatial overlap to control levels (Fig. 5f). Excitingly, similar results were observed with GSK3Ai

(Extended Data Fig. 8g,h), suggesting that reduced cohesin levels on chromatin can be balanced by increased cohesin activity driven either by genetic or chemical inhibition of GSK3A.

Discussion

High-throughput FISH-based screening is a powerful method to directly interrogate chromatin structure in a large number of samples[38–40]. However, this approach has been limited in both scale and design due to the technical challenge of generating FISH probes to single-copy regions, compounded with the quantities of probe required for large-scale assays. The associated labor and cost have hindered the use of FISH-based screens for simultaneous testing of thousands of genes. We have addressed this issue through the development of HiDRO.

In this study, we show the utility of HiDRO by screening the 3,083 genes that make up the human druggable genome and measuring 21 parameters of nuclear organization, creating a rich resource of nuclear phenotypic data. We identified >300 genes that have the capacity to influence chromatin compaction. In particular, we isolated 43 factors in three distinct signaling pathways whose depletion alters the frequency and extent of interactions between TADs. These factors include the ubiquitously expressed and highly conserved protein kinase GSK3A, whose depletion increases interactions between TADs. This result was validated by both Hi-C and FISH with four independent siRNA duplexes and four different small-molecule inhibitors at multiple loci across the genome, confirming the ability of HiDRO to identify novel 3D genome regulators.

While GSK3 signaling has not previously been directly associated with genome organization, our study provides a mechanistic explanation for recent genetic interactions between its chemical inhibition and cohesin dysfunction[41,42]. Lithium treatment, in particular, rescued morphological phenotypes associated with mutations in cohesin in both zebrafish and *Drosophila* systems[42,43]. Lithium has also been shown to restore proliferation rates in cells derived from Cornelia de Lange Syndrome (CdLS) patients harboring mutations in NIPBL[42]. We found that small molecule inhibitors of GSK3, including lithium, are agonists of cohesin activity and increase inter-TAD interactions. Notably, this includes a GSK3A-selective inhibitor used at concentrations below the threshold required for its canonical function in WNT signaling[26]. We show that chemical inhibition of GSK3A also has the capacity to counteract chromatin misfolding following partial cohesin loss due to NIPBL knockdown. These results provide proof of the concept that chromatin looping is a druggable feature of the genome with the potential to correct defects associated with cohesin dysfunction.

We found that GSK3A depletion led to a ~70% reduction in WAPL recruitment to cohesin on chromatin. This phenotype is accompanied by an enrichment of longer-range chromatin loops and an aggregation of cohesin at CTCF sites (Fig. 5g). These changes resemble those observed upon depletion of WAPL, in which increased cohesin residence time leads to an extension of chromatin loops and increased violation of TAD boundaries[1,10,14]. Indeed, we show that GSK3A can genetically interact with WAPL to influence the extent of RAD21

aggregation on chromosomes. Thus, our data support a model in which GSK3A is a novel anti-cohesin factor that regulates the switch between the extrusion of cohesin and its removal from chromatin via recruitment of WAPL (Fig. 5d). This model is further supported by co-immunoprecipitation of GSK3A with the cohesin unloading complex, including WAPL and PDS5A, suggesting GSK3A contacts cohesin directly on chromatin to promote its turnover. We therefore propose that this kinase has the potential to connect genome folding to numerous signaling pathways, including cellular responses to growth factors, insulin, receptor tyrosine kinases, and G-protein-coupled receptors[21,44]. Future work will be important to identify the specific target(s) of GSK3A as well as the other enzymes identified in this study that modulate chromatin folding. Indeed, many residues on cohesin components have been identified as being post-translationally modified, though the functions of these modifications are largely unknown[45,46].

Finally, given that Oligopaints for any DNA or RNA target of interest can be bioinformatically designed[47], we anticipate that HiDRO-based screening could be extended to other features of nuclear organization and transcription or combined with multiplexed approaches to better characterize how chromatin is packaged and spatially controlled in the nucleus[48–50].

Materials and Methods

Cell culture.

We obtained the HCT-116 cell line from ATCC (RRID:CVCL_0291), the HCT-116-RAD21-AID line from ref[31], and the HCT-116-WAPL-AID cell line was generated as described below. All 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. Degradation of the AID-tagged RAD21 was induced with addition of 500 μM indole-3-acetic acid (auxin; Sigma-Aldrich). Degradation of AID-tagged WAPL was induced with addition of 1μM 5-Ph-IAA (modified auxin, ref[51]).

Optimized Oligopaint Design and synthesis.

Oligopaints were designed to have 80 bases of homology using OligoMiner design pipeline[47] with an average of 4 probes per kb and were purchased from Twist Bioscience. Probe coordinates can be found in Table S1. Oligopaints were synthesized as described previously. However, in this study, we directly labeled probes by dye-conjugation. Specifically, 60 nmol dye aliquots were made from either Cy3 (Gold Biotechnology, B-230–1) or Alexa 647 (ThermoFisher Scientific, A20006), which were resuspended in DMSO then vacuum desiccated and stored at −20 °C. Probes were resuspended in 10 μl 0.3 M sodium bicarbonate, then mixed with 60 nmol of fluorescent dye. Probe was incubated at room temperature (RT) in the dark for at least 6 hr to conjugate to dye. Probes were then purified using the Zymo DNA Clean & Concentrator-100. 90 μl water was added to the probe, then 200 μl of oligo binding buffer followed by 400 μl of ethanol. The probe mixture was vortexed and then transferred to a spin column, which was spun at 11,000 x g for 30s. The filter column was transferred to a new 1.5 mL tube and 100 μl water was added to elute the bound probe. After 1 minute (min) incubation at RT, the column was spun at 16,000 x

g for 1 min. The probe mixture was mixed again with 200 μl of oligo binding buffer and 400 μl of ethanol and transferred to the same column as before, spinning at 11,000 x g for 30s. Sample was washed twice with 600 μl wash buffer, spinning at 11,000 x g for 30s and discarding flowthrough each time. The column was spun dry at 16,000 x g for 1 min, then transferred to a new 1.5 mL tube. 100 μl of water was added to the column, incubated at RT for 1 min, then spun at 16,000 x g for 1 min to elute probe. The probe and dye concentration was determined with a Nanodrop spectrophotometer to confirm proper dye conjugation.

HiDRO for DNA FISH.

For HiDRO experiments involving RNAi, 384-well plates (Perkin Elmer 6057300) were either manually seeded with siRNA or ordered pre-seeded with siRNA (Dharmacon) from the Harvard Medical School ICCB-L. Additional control siRNA (Dharmacon) were used: non-targeting control #5, NIPBL, and WAPL (Supplementary Table 2). RNAiMAX transfection reagent (Thermo Fisher Scientific) was diluted in Opti-MEM reduced serum medium (Thermo Fisher Scientific) and pipetted onto siRNA using a manual multichannel pipette. Plates were then spun and incubated at RT for 20 min. All spins for HiDRO plates were performed at 1200 rpm for 2 minutes at RT unless otherwise indicated. For most steps, pipetting was performed by a Matrix WellMate (Thermo Fisher Scientific). HCT-116 cells were trypsinized and resuspended in antibiotic-free medium, then $2x10^3$ to $3x10^3$ cells were seeded in each well. Plates were spun and incubated until collection for downstream experiments. For drug treatments, a manual multichannel pipette was used to add drug or DMSO to cells at the appropriate time points prior to cell collection. On the day of collection, media was aspirated, PBS was added to all wells, and plates were spun. PBS was aspirated and cells were fixed in each well with 4% PFA, 0.1% Tween-20 in 1x PBS for 10 minutes at RT. Plates were spun once during fixation. Plates were then rinsed with 1xPBS and washed twice for 5 minutes with 1xPBS with a spin during each wash. 70% ethanol was then added to each well, plates were sealed with foil plate covers (Corning) and stored for at least 20 hours at 4 °C prior to FISH. On the first day of DNA FISH, ethanol was aspirated and plates were washed in 1xPBS for 10 min to reach RT. Plates were then spun, washed briefly again in 1xPBS and spun again. Cells were permeabilized for 15 minutes in 0.5% Triton-X and 5 minutes in 2xSSCT (0.3 M NaCl, 0.03 M sodium citrate and 0.1% Tween 20). Then 2xSSCT/50% formamide was added to all wells, and plates were double sealed with foil covers. Predenaturation was performed at 91 °C for 3 minutes and then 60 °C for 20 minutes on heat blocks (VWR). After plates were spun, foil covers were removed and hybridization mix was added to wells. Hybridization mix consisted of 50% formamide, 10% dextran sulfate, 4% polyvinylsulfonic acid (PVSA), 0.1% Tween-20, 2xSSC, and each probe at 0.1pmol/μL. 2 pmol of probe was used per 20 μl of hybridization mix. After spinning, plates were double sealed with foil covers and denatured at 91 °C for 20 minutes on heat blocks. Heat blocks were covered to block light and preserve primary fluorescently labeled probes. Plates were spun after denaturation and then hybridized overnight (20–24 hours) at 37 °C. The following day, hybridization mix was aspirated, and plates were washed quickly twice with RT 2x SSCT, then with 60 °C 2xSSCT for 5 minutes. Plates were then washed with RT 2x SSCT for 5 minutes. Nuclei were stained by washing for 5 minutes in Hoescht (1:10,000 in 2x SSCT). Plates were spun, washed for 15 minutes with 2x SSC and spun again. Finally, plates were mounted with imaging buffer (2x SSC, 10% glucose, 10 mM

Tris-HCl, 0.1 mg/mL catalase, 0.37 mg/mL glucose oxidase) and imaged within 5 days of FISH. Although we cannot completely rule out that aspects of chromatin structure were altered following this protocol[52], we found identical CCD measurements between the chr22 domain pairs with a reduced denaturing temperature (75 °C), suggesting that chromatin structure at the scale and resolution we were imaging was preserved.

HiDRO for RNA FISH.

Cell seeding was performed as for DNA FISH HiDRO. After media was aspirated, all wells had PBS added to them and plates were spun. All spins were at room temperature, 1200 rpm for 2 minutes unless otherwise noted. PBS was aspirated and cells were fixed in each well with 4% PFA in 1x PBS (no Tween-20) for 10 minutes at room temperature. When possible, pipetting was performed by a Matrix WellMate (Thermo Fisher Scientific). Plates were spun once during fixation. Then plates were rinsed with 1xPBS and washed twice for 5 minutes with 1xPBS with a spin during each wash. Permeabilization solution (70% EtOH, 1% SDS) was then added to each well, plates were sealed with foil plate covers (Corning) and stored for at least 20 hours at 4 °C until used for FISH. Plates were used for RNA FISH within two days of fixation. On the first day of RNA FISH, permeabilization solution was aspirated and 2xSSCT was added to all wells. Plates were spun and this brief 2xSSCT wash was repeated once. After 2xSCCT was aspirated, primary hybridization mix was added to each well by using a manual multichannel pipette due to the viscous nature of the mix (31.25 nM probes, 1% SDS, 50% formamide, 10% dextran sulfate, 4% PVSA, 0.1% Tween-20, 2xSSC). Plates were spun, double sealed with foil covers, and spun again. Plates were then denatured at 60 °C for three minutes on heat blocks. Plates were then spun again and incubated overnight at 42 °C on heat blocks. The next day, 2xSSCT was preheated to 65 °C. Primary hybridization mix was aspirated and plates were briefly washed with RT 2xSSCT twice. Then 65 °C 2xSSCT was added to plates and incubated on a 65 °C heat block for five minutes. Liquid was aspirated, and this 65 °C 2x SSCT wash was repeated three times. A final brief wash with RT 2x SSCT was performed. Then, secondary hybridization mix (0.05 µM probes, 1% SDS, 50% formamide, 10% dextran sulfate, 4% PVSA, 0.1% Tween-20, 2xSSC) was added and plates were double sealed and spun. After 1 hour incubation at 37 °C in the dark, hybridization mix was aspirated. Nuclei were stained by washing for five minutes in Hoescht (1:10,000 in 2x SSCT) and plates were spun. Plates were then washed with 37 °C 2x SSCT for five minutes on a 37 °C heat block and spun. This wash was repeated once. Plates were mounted with imaging buffer (2x SSC, 10% glucose, 10 mM Tris-HCl, 0.1 mg/mL catalase, 0.37 mg/mL glucose oxidase) and imaged within 24 hours of RNA FISH.

3D DNA slide FISH.

Cells were trypsinized and resuspended in fresh culture medium at $1x10^6$ cells/mL, then settled onto poly-L-lysine coated glass slides for 2h. Cells were fixed to the slides for 10 min with 4% paraformaldehyde in phosphate-buffered saline (PBS) with 0.1% Tween-20, followed by three washes in PBS for 5 min each wash. Slides were stored in PBS at 4 °C until use. On the first day of FISH, slides were warmed to room temperature in PBS for 10 minutes then permeabilized in 0.5% Triton-X 100 in PBS for 15 minutes with nutation. After dipping in PBS, cells were dehydrated in an ethanol row, consisting of consecutive 2 minute incubations in 70%, 90%, and 100% ethanol. Slides were then air dried for 1–2

minutes at RT. Slides were then incubated for 5 minutes each in 2× SSCT (0.3 M NaCl, 0.03 M sodium citrate and 0.1% Tween 20) and 2x SSCT/50% formamide at RT, followed by 1 hr incubation in 2x SSCT/50% formamide at 37 °C. Hybridization mix containing primary labeled Oligopaint probes, hybridization buffer (10% dextran sulfate, 2× SSCT, 50% formamide and 4% polyvinylsulfonic acid (PVSA)), 5.6 nM dNTPs and 10 μg RNase A was added to slides, covered with a glass coverslip, and sealed with rubber cement. 2 pmol of primary Oligopaint probe was used per 25 μl of hybridization mix. Slides were then denatured on a metal heat block in a water bath set to 80 °C for 30 min, then transferred to a humidified chamber and incubated overnight at 37 °C.

The next day, coverslips were removed and the slides were briefly dipped in 2xSSC (0.3 M NaCl, 0.03 M sodium citrate, no Tween-20). Then slides were washed in 2xSSCT at 60 °C for 15 min, 2xSSCT at RT for 10 min and 0.2xSSC at RT for 10 min. DNA was stained by washing slides for 5 min in Hoescht (1:10,000 in 2xSSC). Slides were mounted with SlowFade Gold Antifade (Invitrogen) and sealed under glass coverslips with nail polish.

3D RNA slide FISH.

Cells were trypsinized and resuspended in fresh culture medium at 1×10^6 cells/mL, then settled onto poly-L-lysine coated glass slides for 2h. Cells were fixed to the slides for 10 min with 4% paraformaldehyde in phosphate-buffered saline (PBS) (no Tween-20 in contrast to DNA FISH), followed by three washes in PBS for 5 min each wash. Slides were permeabilized overnight in permeabilization solution (70% ethanol, 1% SDS), and RNA FISH was started within 48 hours of cell fixation. On the first day of FISH, slides were washed twice in 2xSSCT for 5 min, then hybridization mix containing 31.25 nM primary unlabeled Oligopaint probes, hybridization buffer (10% dextran sulfate, 2xSSCT, 50% formamide and 4% PVSA), and 1% SDS was added to slides, covered with a glass coverslip, and sealed with rubber cement. Slides were then denatured for 3 min at 60 °C on a heat block in a water bath and incubated overnight at 37 °C in humidified chamber. The following day, coverslips were removed and the slides were dipped twice in 2xSSCT. Then slides were washed four times for 5 min each in pre-warmed 65 °C 2xSSCT in a plastic coplin jar to avoid glass coplin jars cracking. Slides were washed twice for 5 min each in RT 2xSSCT, then secondary hybridization mix containing 50 nM secondary fluorescently-labeled Oligopaint probes, hybridization buffer, and 1% SDS was added to slides, covered with glass coverslips, and sealed with rubber cement. After 1 hr incubation at 37 °C in a dark humidified chamber, coverslips were removed and DNA was stained in Hoescht (1:10,000 in 37 °C 2xSSCT) for 5 min. Slides were then washed twice for 2 min each in 37 °C 2xSSCT, mounted with SlowFade Gold Antifade, and sealed under glass coverslips with nail polish. Slides were imaged within 24 hours of completion of FISH.

RNA interference.

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% CO2. RNAiMAX transfection reagent (Thermo Fisher Scientific) was diluted in Opti-MEM reduced serum medium (Thermo Fisher Scientific), and pipetted onto siRNA. Plates were incubated at room temperature for 20 minutes. HCT-116 cells were trypsinized

and resuspended in antibiotic-free medium, then plated on siRNA duplexes to obtain a final siRNA concentration of 25 nM. After 72 h of RNAi treatment, cells were collected for experiments. Available siRNA sequences can be found in Table S2.

Microscopy.

Images for HiDRO experiments were acquired on a Molecular Devices Image Xpress Micro 4 high-content microscope using a 0.95 NA 60X air objective (Nikon) or a Molecular Devices Image Xpress Micro 4 Confocal high-content microscope with 0.42 μm pinhole and 1.4 NA 60X water immersion objective. Max projections of z-series (6 images, 0.5 μM spacing) were generated automatically in MetaXpress and used for downstream analyses.

Images of slides were acquired on either a Leica widefield fluorescence microscope using a 1.40 NA 63X oil immersion objective with Andor iXon Ultra emCCD camera or a Leica SP8 confocal microscope using a 1.40 NA 63X oil immersion objective (Nikon).

Immunofluorescence.

Slides were brought to room temperature in 1xPBS for 10 minutes then permeabilized in Triton-X 100 in 1xPBS (0.1% final, Sigma) at room temperature for 15 minutes. Slides were then washed three times with Tween20 in 1xPBS (0.02% final, Sigma; room temperature, 5 minutes each wash), and blocked with BSA solution (2% BSA, 0.02% Tween20, 1xPBS) for 1 hr. Primary antibody was diluted in BSA solution and applied on glass coverslips. The following concentrations of primary antibodies were used: RAD21 (Santa Cruz sc-166973, 1:100), GSK3A (Cell Signaling Technology 4337, 1:50), GSK3B (BD Biosciences 610201, 1:50). Each slide was mounted onto a coverslip and sealed with rubber cement (Elmer's) before overnight incubation at 4 °C in a humidified chamber. On the following day, coverslips were removed and slides were washed three times in Tween20 in 1xPBS (room temperature, five minutes each wash). Secondary antibody was diluted in BSA solution then added to glass coverslips. The following concentrations of secondary antibodies were used: anti-Mouse Alexa Fluor 488 (Jackson ImmunoResearch Labs 115–545-003, 1: 100), anti-Rabbit Alexa Fluor 647 (Jackson ImmunoResearch Labs 111–605-003, 1: 100). Slides were mounted on the coverslips, sealed with rubber cement, and incubated for 1 hr RT in a humidified chamber in the dark. Coverslips were removed and slides were washed twice in Tween 20 in 1xPBS (room temperature, 5 minutes each). DNA was stained by washing in Hoechst once in DAPI stain (1:10,000 in 0.02% Tween-20, 1xPBS) for five minutes. Slides were then mounted in SlowFade Gold Antifade Mountant (Invitrogen).

ChIP-Seq.

ChIP was performed as described previously[53]. HCT-116 cells were cross-linked in culture medium by addition of methanol-free formaldehyde (ThermoFisher, final 1% v/v) and incubated at room temperature for 10 minutes with gentle nutation. Crosslinking was quenched by addition of glycine (final 125 mM) and incubated at room temperature for 5 minutes with gentle nutation. Media was aspirated and replaced with cold DPBS. Cells were scraped and transferred to conical tubes, then pelleted by centrifugation (1500 rpm, 3 minutes, room temperature). Pellets were flash frozen on liquid nitrogen and stored at −80 °C. For ChIP, 30 μl protein G magnetic beads (per sample; ThermoFisher) were washed

three times in blocking solution (0.5% BSA in DPBS). Beads were then resuspended in 250 μl blocking solution and 2 μg antibody (RAD21, Abcam ab992; CTCF, Cell Signaling Technology 3418) was added. Beads and antibody were rotated at 4 °C for at least six hours. Nuclei were isolated from frozen cell pellets as follows: pellet was resuspended in 10 mL cold lysis buffer 1 (50mM HEPES-KOH pH7.5, 140mM NaCl, 1mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100, and protease inhibitors) and rotated at 4 °C for 10 minutes, followed by centrifugation (1500 rpm, 3 minutes, 4 °C). Supernatant was aspirated and the pellet was resuspended in 10 mL cold lysis buffer 2 (10mM Tris-HCl pH 8.0, 200mM NaCl, 1mM EDTA, 0.5mM EGTA, and protease inhibitors) and rotated at room temperature for 10 minutes, followed by centrifugation (1500 rpm, 3 minutes, 4 °C). Supernatant was discarded and nuclei were resuspended in 1 mL cold lysis buffer 3 (10mM Tris-HCl, pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% Na-Deoxycholate, and protease inhibitors) and transferred to pre-chilled 1mL Covaris AFA tubes (Covaris). Nuclei were sonicated using a Covaris S220 sonicator (high cell chromatin shearing for 15 minutes; Covaris). Sonicated chromatin was transferred to 1.5mL microcentrifuge tubes and Triton-X 100 was added (1% final v/v) followed by centrifugation (top speed, 10 minutes, 4 °C). Supernatant was transferred to a new tube. Antibody-conjugated protein G beads were washed three times in blocking solution, resuspended in 50 μl blocking buffer, and added to 500 μg sonicated chromatin. Chromatin was rotated overnight at 4 °C. 50 μg lysate was reserved in a separate tube at −20°C for input. On day 2, beads were washed five times in 1 mL RIPA buffer (50mM HEPES-KOH pH 7.5, 500mM LiCl, 1mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate). Beads were then washed in 1 mL final wash buffer (1xTE, 50mM NaCl) for 2 minutes. Beads were finally resuspended in 210 μl elution buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 1% SDS), and chromatin was eluted from beads by agitation at 65 °C for 30 minutes. 200 μl eluate was transferred to a new tube, and all samples (ChIP and input) were reverse cross-linked overnight at 65 °C with agitation for between 12 and 18 hours. 200 μl 1xTE was added to all samples, and samples were treated with RNase A (final 0.2mg/mL RNase; 37 °C for 2 hours) and proteinase K (final 0.2mg/mL Proteinase K; 55C for 2 hours). DNA was purified using phenol:chloroform extraction and resuspension in 10mM Tris-HCl pH 8.0.

ChIP and input DNA were quantified by Qubit (ThermoFisher) before library preparation using the NEBNext Ultra II DNA library prep kit (NEB). Samples were indexed for either single or dual-index sequencing. Library quality was assessed on Bioanalyzer (Agilent) and quantified by qPCR (Kapa Biosystems). Libraries were pooled, re-quantified, and sequenced on the Illumina NextSeq 500 platform (Illumina, single-end 75bp) or the Illumina NovaSeq 6000 platform (Illumina, paired-end 100bp).

In situ Hi-C.

Hi-C libraries were generated from $1x10^6$ cells using the Arima-HiC+ kit (Arima Genomics) and the Kapa Hyper Prep Kit with KAPA Library Amplification Primer Mix (KK8502), according to manufacturer's recommendations. Libraries were validated for quality and size distribution using the Qubit dsDNA HS Assay Kit (Invitrogen, cat# Q32851), KAPA Library Quantification kit (Roche, Cat# KK4824) and TapeStation 2200 (Agilent). Hi-C libraries were paired-end sequenced (61bp+61bp) on a NovaSeq 6000 (Illumina).

Subcellular Fractionation and Western Blot.

Cells were trypsinized and resuspended in fresh culture medium, then washed once in cold Dulbecco's PBS. Cells were pelleted by centrifugation at 1,200g for 5 min at 4 °C. The pellet was then either processed with the Subcellular Protein Fractionation Kit for Cultured Cells kit (Thermo Fisher Scientific, 78840) or resuspended in 1x RIPA buffer with protease inhibitors to extract whole cell lysate.

For subcellular fractionation, manufacturer's instructions were followed. For whole cell lysate, the sample was nutated at 4 °C for 30 min, and then centrifuged at 16,000g for 20 min at 4 °C. Supernatant containing protein was extracted and quantified using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific, 23225). Protein was stored at −80 °C until needed.

For blotting, protein was mixed with NuPAGE LDS sample buffer and sample reducing reagent (ThermoFisher Scientific), denatured at 70C for 10 min, then cooled on ice. Benzonase (final concentration 8.3 U μl-1) was added to the sample followed by 15 min incubation at 37 °C. Then 10–15 μg of sample was run on a Mini-PROTEAN TGX Stain-free precast gels. Samples were transferred to a 0.2 μm nitrocellulose membranes for 1 hr at 4 °C then washed once in TBS buffer and twice in TBS buffer + 0.05% Tween-20 (TBST). Membranes were blocked for 30 min at RT in 5% milk in TBST, then incubated with primary antibodies diluted in 5% milk overnight at 4 °C. The following primary antibody concentrations were used: β-catenin (Cell Signaling Technology 8480, 1:1000), p-β-catenin Ser33/37/Thr41 (Cell Signaling Technology 9561, 1:500), p-β-catenin Ser675 (Cell Signaling Technology 9567, 1:750), CTCF (Santa Cruz sc-271474, 1:1000), GAPDH (Cell Signaling Technology 5174, 1:2000), GSK3A (Cell Signaling Technology 4337, 1:1000), GSK3B (BD Biosciences 610201, 1:1000), HDAC2 (Cell Signaling Technology 5113, 1:1000), H3 (Abcam 1791, 1:10,000), NIPBL (Santa Cruz sc-374625, 1:400), PDS5A (Bethyl A300–089A, 1:1000), RAD21 (Abcam ab992, 1:1000), SMC1A (Thermo Fisher Scientific PA5–29122, 1:1000), WAPL (Cell Signaling Technology 77428, 1:1000). The following day, membranes were washed twice in TBST for 5 min each. Membranes were then incubated with secondary antibodies diluted in 5% milk TBST in the dark at RT for 1 hr. The following secondary antibody concentrations were used: anti-Rabbit HRP-linked (Cell Signaling Technology 7074, 1:5000), anti-Rat HRP-linked (Cell Signaling Technology 7077, 1:5000), anti-Mouse HRP-linked (Cell Signaling Technology 7076, 1:5000), anti-Mouse IRDye 800CW (LiCOR 926–32210, 1:3333), anti-Rabbit Cy3 (Jackson ImmunoResearch Labs 111–165-003, 1:6666). Membranes were then washed twice in TBST for 15 min each and once in TBS. Blots with HRP-linked secondary antibodies were additionally incubated for 1–5 minutes with Clarity or Clarity Max ECL Western Blotting substrates (Bio-Rad). Blots were then imaged on a ChemiDoc MP Imaging System and analyzed with Bio-Rad Image Lab software (v5.2.1).

Fluorescence recovery after photobleach (FRAP).

FRAP was performed similar to previous studies[1,14,54]. We used a HCT-116 cell line with RAD21 biallelically tagged with mClover. Images were acquired with a Leica Stellaris 5 Confocal microscope with 63X/1.4 NA oil objective and an Okolab incubation enclosure.

Five prebleach images were acquired, then half of each nucleus was bleached with ten pulses of an 488-nm laser. After bleaching, five frames were acquired every two seconds, followed by 120 frames every 10 seconds. Fluorescence intensity was measured using ImageJ. Background intensity was subtracted from bleached and unbleached intensity measurements. Then the relative intensity between bleached and unbleached regions was calculated by dividing the background-corrected bleached intensity by the background-corrected unbleached intensity. The mean of the relative intensity of the prebleach images was determined and all values were normalized such that the relative intensity before bleaching had a mean of 1. The median normalized intensity +/- SEM was plotted for all time points.

Proximity labeling by Turbo-ID.

Plasmid vectors with the CMV promoter were designed to express either GSK3A fused to BirA and V5 tag or BirA-V5 alone. The TurboID protocol was performed based on the published protocol[55] with some modifications. Cells were seeded in 6 well plates with 3 ml of media containing 0.5×10^6 cells/ml. The next day, media was aspirated from each well and replaced with 2 ml of fresh media prior to transfecting with a mixture of 2.5 μg of plasmid vectors, 250 μl Opti-MEM reduced serum medium (Thermo Fisher Scientific), 7.5 μl Lipofectamine 3000 reagent, and 5 μl of P3000 (Thermo Fisher Scientific) reagent. After 24 hours, media was aspirated from each well and replaced with media containing 500 μM of biotin to allow for proximity labeling. After another 24 hrs, the plate was moved to ice, media was aspirated, and wells were washed five times with 500 μl ice-cold PBS. Cells were pelleted at 300 g at 4°C for 3 min, supernatant was aspirated, and the cell pellet was resuspended in ice-cold RIPA (50 mM Tris pH 8, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) supplemented with protease inhibitors and 1 mM PMSF and incubated for at least 10 minutes on ice. Lysates were clarified by centrifugation at 13,000 g at 4 °C for 10 min, and the clarified supernatant was transferred to a new tube. Protein concentration was quantified using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific, 23225).

To perform the immunoprecipitation (IP), 25 ul of MyOne Streptavidin T1 magnetic beads (Thermo Fisher Scientific) were washed in RIPA buffer before being combined with 75 μg of lysate in a total volume of 1 mL of RIPA buffer and incubated overnight in a rotator. The next day, beads were collected on a magnet and washed and incubated in RIPA buffer twice for 2 min each time, 1 M KCl for 2 min, 0.1 M Na_2CO_3 for 15 s, 8M urea for 15 s, and twice more in RIPA buffer for 2 min each time. Beads were then incubated at 95 °C for 15 min with 2 mM biotin, NuPAGE LDS sample buffer and sample reducing reagent (ThermoFisher Scientific), and supernatant was collected for blotting using the same Western procedure described earlier. Antibody concentrations are as follows: RAD21 (Abcam 992, 1:1000), SMC1A (Bethyl A300–055A, 1:1000), PDS5A (Bethyl A300–089A, 1:1000), β-catenin (Cell Signaling Technology 8480S, 1:1000), V5 (Invitrogen R96025, 1:1000), WAPL (Cell Signaling Technology D9J1U, 1:1000), and NIPBL (Santa Cruz sc-374625, 1:500).

Real-time quantitative polymerase chain reaction.

RNA was isolated from HCT-116 cells using the RNeasy Plus kit (Qiagen) per manufacturer's instructions. cDNA was synthesized in a 50 μl reaction containing 20 μl RNA, 0.7μl Maxima RT (Thermo Scientific), 1500pmol Oligo dT primer (IDT), 1.6mM dNTPs, 1x RT Buffer (Thermo Scientific) and 0.5μl RNase OUT (Invitrogen). This reaction was incubated at 50°C for two hours then 85°C for 5 min. Samples were stored at −20°C until use. RT-PCR was performed using PowerUP Sybr (ThermoFisher, #A25741). Briefly, cDNA was diluted to a working concentration of 6μg and HCT-116 genomic DNA (gDNA) was diluted in a 1:10 serial dilution. A 6μl reaction was prepared per well, with 1x PowerUP Sybr and 0.2μM of the forward and reverse primers and combined with 4μl diluted DNA. Each reaction was performed in triplicate. qPCR was performed on the QuantStudio7 Flex System. UBC and TBP were used as reference control genes.

Generation of HCT-116-WAPL-AID2 cell line.

Plasmids were generated as previously described[51]. Briefly, the CRISPR plasmid (WAPL-C CRISPR) to induce a double strand break in C-terminus of WAPL was modified from pX330-U6-Chimeric_BBCBh-hSpCas9[56] by inserting sequence for a gRNA targeting the C-terminus of WAPL (IDT, target site: 5'TTACCGAGCACCTGAAGCAA3'). The donor plasmids for the mAID2 cassette were created by first amplifying the C-terminus of WAPL from HCT-116 genomic DNA and inserting into pBluescript II cloning vector. The WAPL-mAID2-mClover3-Neomycin (WAPL-mAC-Neo, modified from Addgene 72827) and WAPL-mAID2-mClover3-Hygromycin (WAPL-mAC-Hygro, modified from Addgene 72828) were inserted into the homology site. The previously described parental cell line HCT-116-OsTir(F74G)[57] was transfected with the three plasmids: WAPL-C CRISPR, WAPL-mAC-Neo and WAPL-mAC-Hygro. The cells were selected with G418 and hygromycin for 10–13 days. Colonies were isolated and genotyped by PCR as shown in Extended Data Fig. 5. WAPL was depleted by addition of 1μM 5-Ph-IAA (modified auxin)[51].

QUANTIFICATION AND STATISTICAL ANALYSIS

HiDRO Image Analysis.

To reduce computing time, we did not apply a deconvolution algorithm prior to analysis of our primary screen high-content images. Our data from Figure 1 strongly indicated that our segmentation of raw 2D max-projected images via the CellProfiler software package was incredibly efficient at detecting changes in spatial overlap and CCD. However, images for all follow-up experiments, were deconvolved with Huygens Essential v20.04 (Scientific Volume Imaging) using the Classic Maximum Likelihood Estimation algorithm and theoretical PSFs. Unprocessed images from HiDRO plates were segmented and measured using CellProfiler (version 3.1.8, ref. [58]). Edges of nuclei were detected using minimum cross entropy thresholding[59] and edges of FISH spots were detected using three-class Otsu thresholding[60]. For spot segmentation, only spots within segmented nuclei were kept for measurement. The parameters for Otsu thresholding were tuned by visually inspecting the segmentation results in CellProfiler. Spot overlap measurements were determined as follows. Spots were assigned to nuclei using "RelateObjects" in CellProfiler, and any spots not contained in

a nucleus were excluded. Overlap regions between two spots A and B were identified and paired to each spot using the "MaskObjects" and "RelateObjects" modules. Then, to identify A and B spots on the same allele, for spots A1, …, Ai and B1,…, Bj in a single nucleus, a list of all possible pairings of spots (Ai, Bj) was generated. An empirical center-center distance (CCD) cutoff was applied to exclude pairs that had a CCD greater than the cutoff. An empirical cutoff of 2.85 μm for TADs and 1.37 μm for subTADs was determined from examining the distribution of distances for the range of TAD and subTAD pairs evaluated in Figure 1. Spot pairs with the minimum CCD between them were matched by destructively iterating through the list of all possible TAD pairs. To compare HiDRO measurements across different plates and experiments, robust z-scores for each measurement were generated comparing each individual well against all of the wells on the same replicate plate, excluding the positive control wells seeded with NIPBL and WAPL siRNA. Negative control wells with non-targeting siRNA were included for comparison. The mean of each measurement was calculated per well, and the formula for robust z-score was applied: (mean measurement for well – median of mean measurements for all wells on plate excluding positive controls)/ (1.486*(median absolute deviation for the mean measurements across the plate, excluding positive controls)). Full data is reported in Supplementary Table 3. For the validation de-pooled screen, robust z-scores were calculated from a null distribution of 62 wells seeded with non-targeting control siRNA #5. Full data is reported in Supplementary Table 5.

To correct for chromatic aberrations, we recorded the chromatic offset in x,y (for 2D analysis of HiDRO images) with 0.1 micron TetraSpeck microspheres (Thermo Fisher Scientific, T7279). Microscope calibrations were then performed through mechanical and software modifications to correct for the chromatic offset during subsequent acquisitions using MetaXpress software. Note that our spatial overlap metric requires images to be shifted prior to segmentation as any shift may change the amount of overlap in a nonlinear fashion (domains are larger and more amorphous than diffraction-limited foci). We continually took measurements with Tetraspeck beads and recalibrated our system prior to each experimental run if we recorded an offset greater than the size of a pixel of that system. We note that we cannot completely rule out that some chromatic shifts transiently occurred throughout our long HiDRO imaging sessions (24-hours per plate x 6 plates per week). Therefore, to account for additional plate-to-plate variation, we prepared replicate plates on separate weeks and restricted our analysis of screen results to robust z-scores (Supplementary Table 3) rather than the absolute values of CCD or spatial overlap.

Primary Screen and Validation Screen hit selection.

Hits were selected from the primary screen data by meeting the following criteria based on nuclei count and overlap measures (CCD, D1 overlap, D2 overlap): Both replicates had z-scores with absolute value   1.5. Both replicates should have z-scores of the same sign (positive or negative). The list of 58 primary hits was manually selected from genes that altered CCD and at least one overlap measure. In the primary screen, each gene was targeted with a pool of four siRNA duplexes in a single well.

For our validation screen, we ordered from ICCB-L 384-well plates seeded with the four siRNA duplexes for each primary gene hit de-pooled into four separate wells, arranged randomly on one 384-well plate. We performed the validation screen in triplicate. Robust z-scores for overlap measurements were calculated for each well versus a null distribution of 62 wells seeded with non-targeting control siRNA #5 (Supplementary Table 4). We set a nucleus count cutoff of 100 nuclei. Individual siRNA duplexes were considered validated hits if they had z-scores with absolute values 2 in at least 2/3 replicates for at least 2/3 overlap measurements. This resulted in 106/232 duplexes validating corresponding to 52/58 primary hit genes. To further filter validated genes from these duplexes, genes that had a mix of NIPBL-like and WAPL-like duplexes were excluded. This left us with 43/58 validated genes, which are listed in Supplementary Table 5.

For analysis of multiple TAD boundaries in Figure 2, we measured four biological replicates for each gene KD at each boundary tested. We found CCD to be the most reliable measurement across boundaries of different insulation scores, so we scored a gene KD as a hit for a boundary if it had at least 2/4 replicates with CCD robust z-score vs. negative control wells absolute value 2.

Measurement correlation and tree visualization.

Primary screen measurements across all replicates were converted to z-scores. Replicates with fewer than 100 nuclei per well were excluded. We then converted the z-scores into "hits" by setting all values between -1.5 and 1.5 (exclusive) to zero. Values of -1.5 or lower were set to -1, and values of 1.5 or greater were set to 1, resulting in a new matrix of the same shape that indicated whether a significant shift was observed (and in which direction) for a measurement within a single replicate of a single gene knockdown. If more than 50 percent of replicates for a single gene were significant in the same direction, that measure was recorded as either a positive or negative hit for that gene. (If a gene had replicates that were hits in opposite directions, that gene was recorded as a non-hit.) This resulted in a deduplicated list of all genes, with a single entry for each measure indicating whether that knockdown had a positive, negative, or null effect on the measure. We used this to extract a vector of hits for each measurement that was a combination of -1, 0 and 1, one entry per gene—we then calculated pairwise relationships between each measurement using the polychoric correlation of these vectors (Polycor, https://cran.r-project.org/web/packages/polycor/index.html)[61]. The rho values for each pair were used to generate a heatmap (Extended Data Fig. 2f) that we used to build five groups of related measurements, referred to as measurement categories, that were also highly correlated: overlap metrics, domain area, domain shape, nuclear area and nuclear shape. In several cases, hits within a group were negatively correlated with the others in the group—that is, increases in one measure were strongly correlated with decreases in another measure. In these cases (i.e. measures of compactness and eccentricity), the correlation scores were multiplied by -1, which maintained the separate classifications but aligned the values with the other measures in each group. We then selected a single measure from each group to represent all measures in that group. These group-level hits were used to visualize the genes in Fig. 2f.

DNA and RNA slide FISH Image Analysis.

Widefield images were deconvolved with Huygens Essential v20.04 (Scientific Volume Imaging) using the Classic Maximum Likelihood Estimation algorithm with theoretical PSFs. Additional settings include signal-to-noise ratio (SNR) of 20 and 50 iterations for FISH spots, SNR 10 and 50 iterations for IF signal, or SNR of 40 and 2 iterations for DNA stain. The deconvolved images were segmented and measured using a modified version of the TANGO 3D-segmentation plug-in for ImageJ[62,63]. Nuclei and FISH spots were segmented using a hysteresis-based algorithm. To correct for chromatic aberrations, we recorded the chromatic offset in x,y,z with 0.1 micron TetraSpeck microspheres (Thermo Fisher Scientific, T7279). Microscope calibrations were then performed through mechanical and software modifications to correct for the chromatic offset during subsequent acquisitions using Leica software. In some cases, corrections were applied to images immediately after acquisition. Note that our spatial overlap metric requires images to be shifted prior to segmentation as any shift may change the amount of overlap in a nonlinear fashion (domains are larger and more amorphous than diffraction-limited foci). We continually took measurements with Tetraspeck beads and recalibrated our system prior to each experimental run if we recorded an offset greater than the size of a pixel of that system.

RAD21 IF and Signal Granularity Quantification.

From a z-series, the plane of best focus in the RAD21 channel was identified using the Normalized Variance algorithm in ImageJ. The corresponding z-plane was pulled from the DAPI channel for concurrent analysis of nuclei. Images were automatically processed in CellProfiler (v4.2.1)[64] for nuclear segmentation and granularity measurement with a structuring element radius of 5 pixels and a range of granular spectrum of 10. The granularity measurement for the second instance of the granularity spectrum was reported. Nuclei that were out of focus or for which segmentation was inaccurate were manually filtered and removed from analysis. The statistical analysis of the cell counts were performed in GraphPad Prism v.9.2.0.

ChIP-seq Analysis.

Sequencing quality was examined using FastQC (v0.11.9) to ensure that the library GC% and duplication rate were within expected range and adapters were effectively trimmed[65]. Trimming was performed with Trimmomatic (v0.38) using a sliding window of four bases and a minimum average quality score of 20; reads shorter than 20 base pairs were discarded. Remaining reads were aligned against the GRCh38 reference genome[66] using Bowtie2 (v2.4.1)[67] allowing soft clipping ("--local") with the "sensitive local" preset options. Alignments were then filtered for duplicates using the Picard (v2.27.1) "MarkDuplicates" tool. Alignments that mapped to ENCODE blacklisted regions[14] were also removed for downstream analysis using BEDtools v2.30.0[68]. Visualization tracks were generated using the deepTools (v3.3.1) "bamCompare" utility[69],in which corresponding input controls were subtracted from IP and reads with a MAPQ score lower than 20 were discarded. Peaks were called for IP libraries against their corresponding input controls using MACS2 (v2.1.1)[70], with default parameters and a 0.01 q-value cutoff. Lost, retained, and gained sites in GSK3A KD were determined with DiffBind v3.6.2[71] using the DESeq2 analysis method[72], a false-

discovery cutoff of 0.1, and only reads with a minimum MAPQ score of 20. Example ChIP-seq tracks were generated in PyGenomeTracks v3.6[73].

Binding heatmap visualization.

ChIP-seq signal heatmaps illustrate results based on merged samples. Samples were grouped by condition (GSK3A knockdown, control) and ChIP target (RAD21, CTCF, input); within each group, alignment files were downsampled so that each sample had an approximately equal number of aligned reads. Each sample was then merged together, resulting in a single file for RAD21 ChIP-seq in the GSK3A knockdown, another for input sequencing in the GSK3A knockdown, another for RAD21 ChIP-seq in the control condition, and so on. IP and input samples were then compared using the same bamCompare parameters described above. Once these tracks were generated, they were used as the input into the deepTools "computeMatrix" command, using a bin size of 10 and the "center" reference point option. The result was used as the input to the deepTools "plotHeatmap" command[69].

Hi-C Analysis

Data processing:

Raw reads for each Hi-C sample were processed with HiC-Pro (version 2.11.1) to obtain putative interactions with default parameters except LIGATION_SITE = GATCGATC and GENOME_FRAGMENT generated for MboI restriction enzyme[74]. For downstream analyses, ValidPairs were converted to hic files using the "hicpro2juicebox.sh" in utils of HiC-Pro. We recovered the following number of contacts in each condition: Control replicate 1 (507359205), Control replicate 2 (680457594), GSK3A KD replicate 1 (445835605), GSK3A KD replicate 2 (533783209), WAPL KD replicate 1 (375716000), WAPL KD replicate 2 (280660878), PDS5A replicate 1 (278576035), PDS5A replicate 2 (275103778). To match sequencing depth achieved in each GSK3A replicate, PDS5A and WAPL replicates were merged with HiCExplorer by using the "merge" function. Replicates were assessed by using the Python implementation of HiCRep package[75] using the parameters --bDownSample --h 10 --dBPMax 1000000 to calculate a stratum-adjusted correlation coefficient.

Domain calling with 3DNetMod:

We assembled binned Hi-C contact matrices for each chromosome in each condition at 10 kb resolution. We performed a simple scalar normalization. We computed a simple scalar size factor for each pixel based on the genomic distance between pairs of bins in the contact matrix[76]. Each pixel value in the contact matrix was divided by the size factor as a way of normalizing the counts distributions across multiple conditions. Each pixel was then rounded to the nearest whole number. We applied a filter to the scaled matrices to remove regions with sparse row counts and high spatial outliers. Sparse row count regions were identified as rows with less than 35 nonzero entries in the first 750 kb in either direction from the diagonal. We considered pixels to be an outlier if their value exceeded four times the local median in a surrounding 5x5 pixel square neighborhood of adjacent pixels, or greater than four if the local median was less than one. Finally, we performed Knight-Ruiz matrix balancing on the filtered matrices as previously described[77,78].

We used our previously published code 3DNetMod[79], to identify TADs and subTADs genome-wide on normalized and balanced 10 kb Hi-C matrices. As previously described[76,77,79,80], we log transformed genome-wide counts data. We chunked counts data into 6 Mb regions with 4 Mb overlap as well as 3 Mb regions with 2 Mb overlap between adjacent regions. In our analysis, we filtered out sparse regions with consecutive zero counts on the diagonal for distances greater than or equal to 500 kb. Sparse regions with zero counts for 1/3 of all pixels on diagonal of a chunked region were also filtered. We used a plateau size of 16 and 8 for 6 Mb and 3 Mb chunked regions, respectively, as the minimum consecutive gamma plateau size required for analysis. Gamma steps of 0.01 were used to identify plateaus that result in equal number of domains, determined as a mean per 20 partitions. We filtered domains smaller than 60 kb as well as removed domains within 20 bins from the edges of chunked regions. We merged the 6 Mb and 3 MB chunked regions genome-wide. Finally, domains that co-localized within +/- 70 kb to both boundary edges were considered redundant calls and merged into a single domain. We determined a final set of unique boundary locations by adjusting domains whose gap between adjacent close boundaries was less than 7.5% of the domain size for all adjacent domains or less than 70 kb to share a single consistent boundary.

To calculate insulation scores, .hic files were converted to either .cool format using HiCExplorer "convert" function (Release 3.7.2)[81] and Knight-Ruiz normalized using the "correct" function. Insulation scores were called for each 5 kb bin using the "cooltools insulation" function from the cooltools package[82] using a window of 1 Mb and the parameter --ignore-diags 0. Then, insulation scores were retrieved for each domain boundary and its neighboring upstream and downstream flanks for a combined 500 kb window, and the median insulation score was graphed for each condition.

Loop calling:

Loops were called using Mustache (version 1.2.4)[83]. At 5 kb and 10 kb resolutions, loops were called using the parameters -pt 0.01 (q value) -p 10 -st 0.88. Loops called at both resolutions were merged into one list by using HiCExplorer's "hicMergeLoops" function with parameter --lowestResolution 10000. Loop anchors were deemed overlapping if they had >=1 bases of overlap. Lost, retained, and gained loop anchors were determined using bedtools. Lost loops were defined as control loops with anchors A1 and A2 where when compared against all GSK3A KD loops with anchors B1 and B2, (A1 had zero overlap with B1) OR (A2 had zero overlap with B2). Retained loops were defined as control loops with both loop anchors A1 and A2 where when compared against all GSK3A loops with anchors B1 and B2, (A1 had >=1 base overlap with B1) AND (A2 had >=1 base overlap with B2). Gained loops were defined as GSK3A KD loops with anchors B1 and B2 where when compared against all control loops with anchors A1 and A2, (B1 had zero overlap with A1) OR (B2 had zero overlap with A2).

Pileup of Mustache loops and domain boundaries:

The aggregate pileup analysis of loops in each condition and the log2fold change was done with the coolpup.py function from the coolpuppy package (version 0.9.5)[84] using the 5 kb resolution .cool files. For loops, the parameters --mindist 0 and --expected were used.

For domain boundaries, the parameters used were --local --ignore_diags 2 --pad 250 and --expected. The --expected parameter for both pileups utilizes the output from cooltools "compute-expected" function and parameter -t cis. These pileups were then plotted using plotpup.py function.

Stripe detection and analysis:

Stripes were detected using Stripenn version 1.1.50[85] with settings resolution=5000, –norm=weight, –canny=1.5, –minL=10, maxW=8, –maxpixel=0.95,0.96,0.97,0.98,0.99 and –pvalue=0.05. The pileup analysis of stripes in each condition and the log2fold change of stripes was done with the coolpup.py function from the coolpuppy package (version 0.9.5)[84] using the 5 kb resolution .cool files and parameters --mindist 0 --local --pad 2000 --ignore_diags 2 and --expected, where the last parameter utilizes the output from cooltools "compute-expected" function and parameter -t cis. These pileups were then plotted using plotpup.py.

STATISTICS AND REPRODUCIBILITY

The number of samples per condition (n) is indicated in the respective figure legends of all graphs. For HiDRO experiments, each well was a biological replicate. Each 3D slide FISH experiment reported in this study was repeated in at least one biological replicate (Fig. 2k/Extended Data Fig. 2l; Fig. 3c/d; Fig. 5a/b, Extended Data Fig 3c/d, 3f/g; Extended Data Fig.5b/c, Extended Data Fig. 6g/h). Immunofluorescence experiments were repeated in at least one biological replicate (Fig. 3a, Fig. 5i/Extended Data Fig. 5l). Co-immunoprecipitation experiment in Fig. 6b was repeated in three additional biological replicates (Extended Data Fig. 6b). GSK3A Turbo-ID experiment in Extended Data Fig. 6c was repeated in one additional biological replicate (Extended Data Fig. 6d). Western blot experiments were repeated in at least one additional biological replicate (Extended Data Fig. 3b,e; Extended Data Fig. 5a). FRAP experiment data (Extended Data Fig. 5f,g) is from two biological replicates pooled together. Biological replicates involved an independent collection of cells including relevant treatments. Protein class designations were determined by PANTHER and manual curation of unassigned proteins[30]. All other statistical tests are discussed in the corresponding methods subsection or figure legend. Statistical analyses were performed using either python or Prism 9 software by GraphPad (v9.2.0).
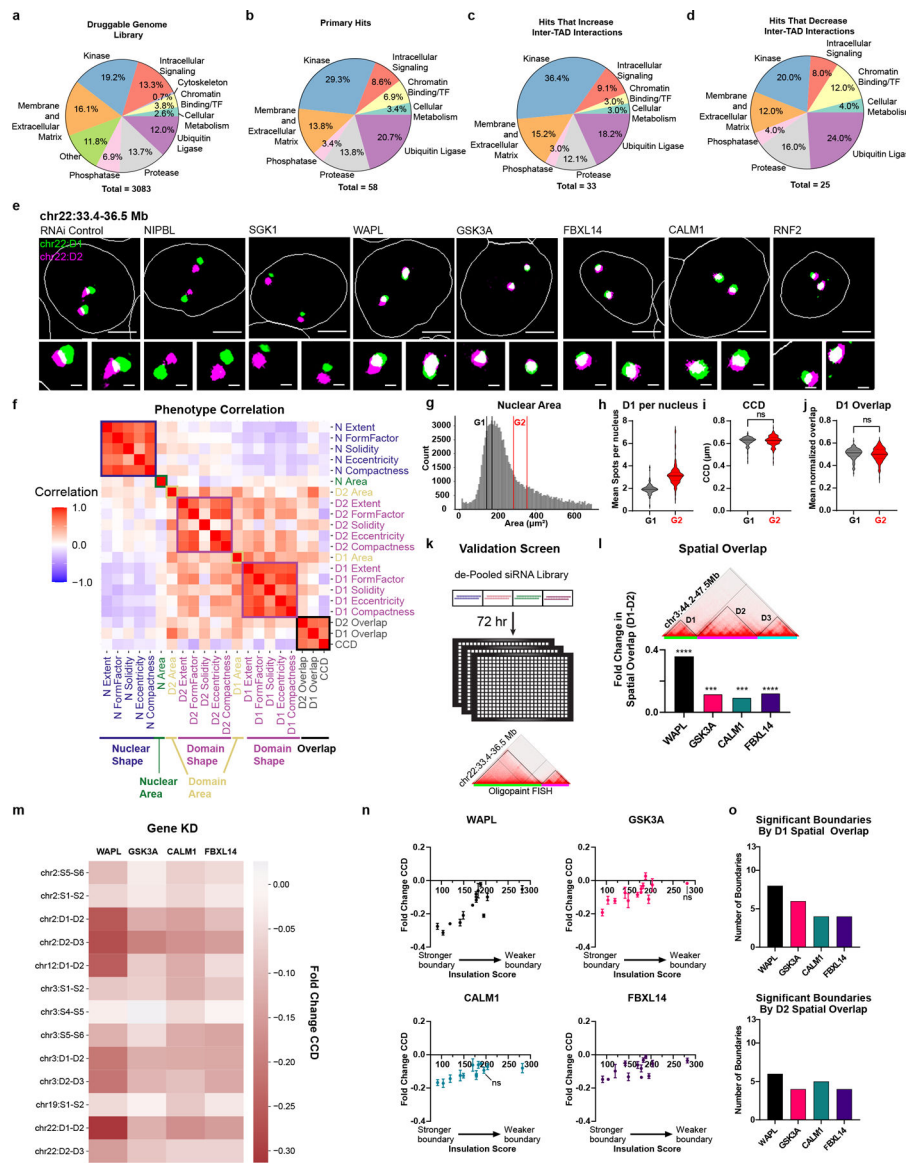
Extended Data
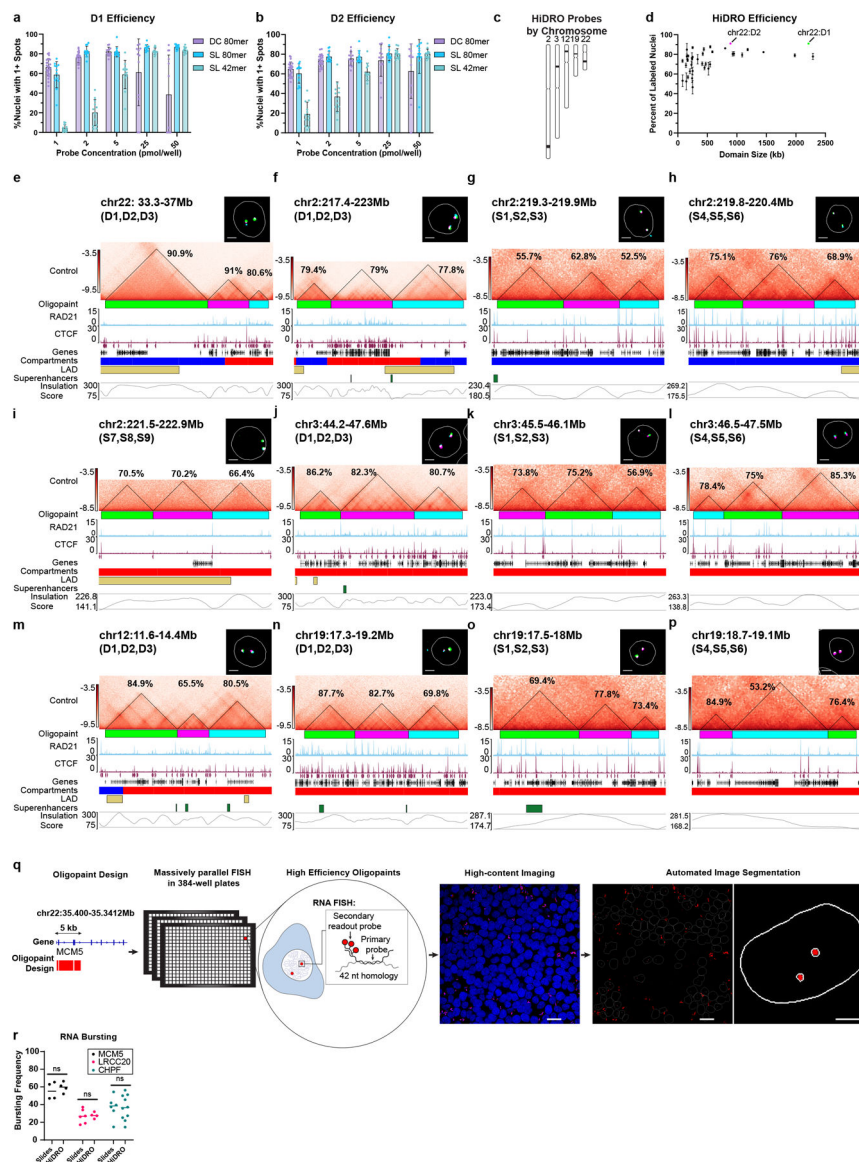
Extended Data Figure 1. Additional DNA HiDRO spot efficiency and RNA HiDRO workflow, Related to Figure 1.

a) Labeling efficiency for D1 locus as measured by percent of nuclei with at least one signal for different Oligopaint probe designs including dye-conjugated 80-mers (DC 80), secondary 80-mers (SL 80) and secondary labeled 42-mers (SL 42, conventional Oligopaints used in ref[7]). Each bar is mean +/- SD. [n = 12 biological replicate wells for all conditions except DC80 1pmol/well and 2pmol/well (n = 24) and UL80 25 pmol/well (n = 11)].

b) Labeling efficiency for D2 locus for different Oligopaint probe designs. Each bar is mean +/- SD. [n = 12 biological replicate wells for all conditions except DC80 1pmol/well and 2pmol/well (n = 24) and UL80 25 pmol/well (n = 11)].

c) Ideograms showing chromosomal locations of Oligopaint probes to 42 DNA regions tested by HiDRO.

d) Labeling efficiency as measured by percent of nuclei in a well with one or more signals detected. Chr22 D1 (green) and D2 (magenta) highlighted. Each data point represents mean +/- SEM of six biological replicates.

e) – (p) Hi-C contact matrices for boundaries tested by DNA HiDRO. Hi-C from ref[4]; tracks below are Oligopaint design, RAD21 ChIP-seq peaks (ENCODE ENCFF001UEG), CTCF ChIP-seq peaks with directionality (GEO GSM1022652), Genes, Compartment designation by eigenvector[4], lamina associated domains (LADs) (4DN Data Portal : 4DNFI2BGIZ5F), superenhancers[86], and insulation scores[7]. Percentages above each domain represent the probe efficiency for that domain as measured by the percentage of nuclei with at least one spot detected.

q) Schematic for RNA HiDRO. Probes can be designed to introns and/or exons and RNA FISH is performed in 384-well plates. Wells are imaged on a high-content microscope, and nascent signals in the nucleus are segmented and measured computationally. Solid white line indicates nuclear edge. Scale bar field, 10 μm; Scale bar nucleus, 5 μm.

r) Bursting frequency of three genes *MCM5*, *LRCC20*, and *CHPF* as measured by 3D RNA FISH on slides and RNA HiDRO shown. Each dot represents one biological replicate of bursting frequency calculated from greater than 100 nuclei. *ns* = P>0.05, two-tailed t-test. [Biological replicate wells for MCM5 Slides and HiDRO: n = 4; LRCC20 Slides n = 6, HiDRO n = 4; CHPF Slides n = 6, HiDRO n = 10].

Extended Data Figure 2. HiDRO screen validated hits are region non-specific regulators of TAD boundaries, Related to Figure 2.

a) Protein classes of genes in the Druggable Genome library. Targeted genes encode proteins across diverse classes including kinases, membrane and extracellular matrix proteins, and proteases. n = 3083.

b) Protein class designations of all primary hits from Druggable Genome HiDRO screen (n = 58).

c) Protein class designations of primary hits that increase inter-TAD interactions (n = 33).

d) Protein class designations of primary hits that decrease inter-TAD interactions (n = 25).

e) Representative images of chr22:D1-D2 for hits altering inter-TAD interactions. Solid white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm. Each gene was tested in two biological replicates.

f) Correlation heatmap of 21 image-based phenotypes with color-coded squares outlining the five measurement categories used for the phenotypic tree in Fig. 2f. The five categories are overlap metrics, domain area, domain shape, nuclear area and nuclear shape.

g) Histogram of nuclear area for non-targeting control wells from replicate 1 of Fig. 1b data. Black lines denote the 20th ($142\ \mu m^2$) and 30th percentiles ($170\ \mu m^2$) of nuclear area, representing G1 nuclei. Red lines denote the 70th ($283\ \mu m^2$) and 80th percentiles ($354\ \mu m^2$) of nuclear area, representing G2 nuclei. n = 128 wells.

h) Violin plot of D1 spots detected per nucleus per well in G1 and G2 nuclei. Solid line is median, dotted lines are 25th and 75th percentiles. [Data from n = 128 wells for each bar.]

i) Violin plot of mean CCD per well in G1 and G2 nuclei. Solid line is median, dotted lines are 25th and 75th percentiles. [Data from n = 128 wells for each bar.] *ns* P-value > 0.05, two-tailed t-test.

j) Violin plot of mean D1 overlap per well in G1 and G2 nuclei. Solid line is median, dotted lines are 25th and 75th percentiles. [Data from n = 128 wells for each bar.] *ns* P-value > 0.05, two-tailed t-test.

k) Validation HiDRO screen workflow tests each primary hit with four independent siRNA duplexes in separate wells, then applies DNA FISH to chr22 domains D1 and D2.

l) Hi-C contact matrix and Oligopaint design for three adjacent TADs on chr3. Hi-C data from ref[4]. Fold change in spatial overlap between chr3 D1 and D2 for top hits. [Data from one well per condition of HiDRO experiment; Number of Alleles for Control (n = 1,075), WAPL (n = 426), GSK3A (n = 855), CALM1 (n = 1,321), FBXL14 (n = 1,184)]. *** P-value < 0.001, **** P-value < 0.0001, two-tailed Mann-Whitney U test.

m) Heatmap displaying fold change in CCD at 13 boundaries across the human genome for WAPL, GSK3A, CALM1 and FBXL14 KD. Each boundary was tested with 3–4 biological replicates per gene KD.

n) Fold change in CCD versus insulation score at boundary for top siRNA KD. For each graph, x-axis is insulation score of the boundary. Insulations scores from ref[4]. Y-axis is fold change in center-center distance between domains relative to control. Each point is the mean of 4 biological replicates except WAPL KD at insulation scores 91, 119, 142, 149, 195 (n = 3 replicates) error bars are +/- SEM.

o) Number of significantly altered boundaries as measured by D1 overlap and D2 overlap for different gene KD.
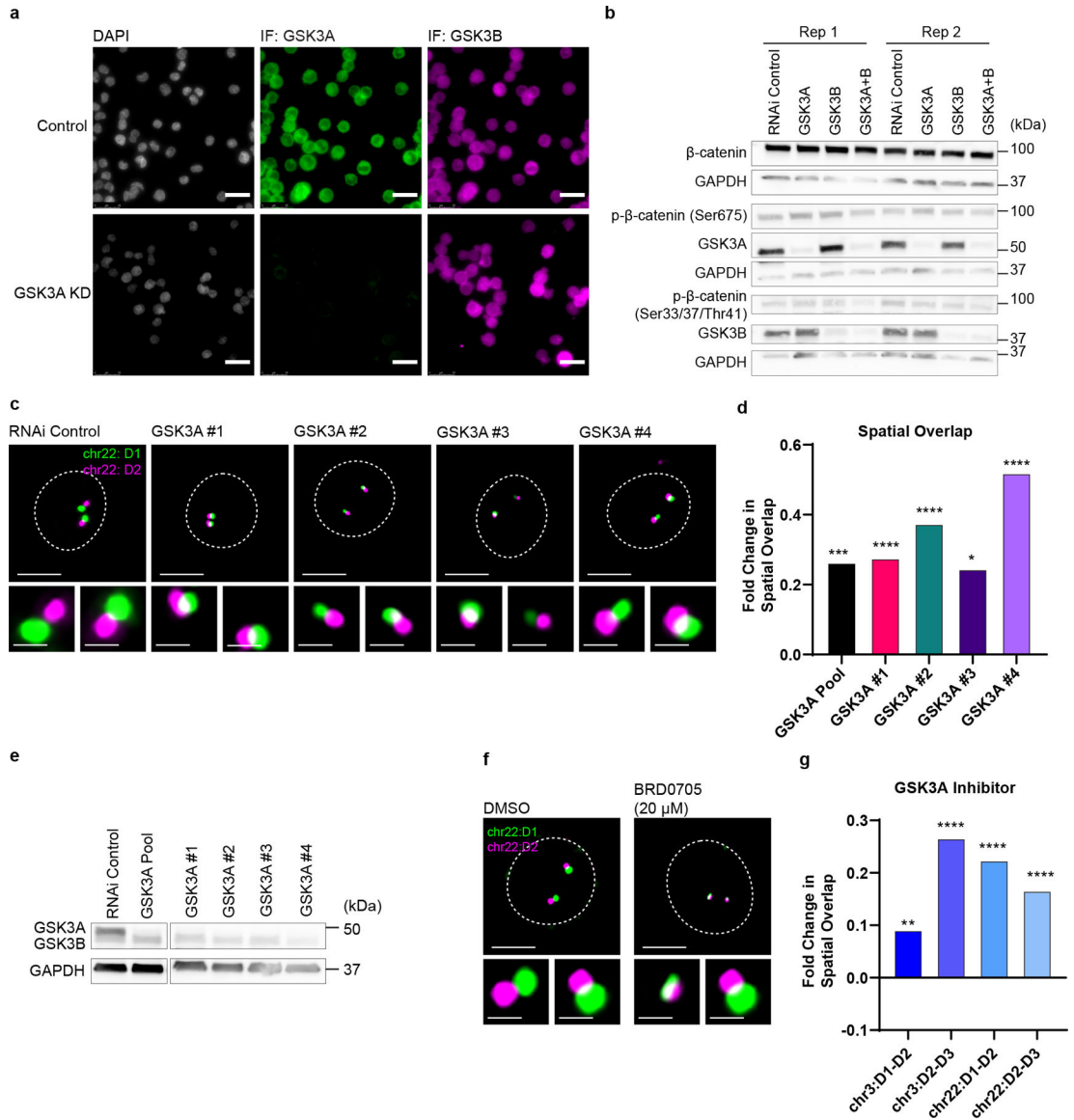
Extended Data Figure 3. Validation of a non-canonical role for GSK3A in inter-TAD interactions, Related to Figure 3.

a) Representative fields of control and GSK3A KD HCT-116 cells after IF to GSK3A and GSK3B. Scale bar, 25 μm.

b) Western blot of whole cell lysate after KD of GSK3A, GSK3B, or GSK3A+GSK3B. Proteins were labeled with HRP-linked antibodies. [Images in order top to bottom. Blot 1: Beta-catenin total, GAPDH; Blot 2: phospho-Beta-catenin Ser675 (activated), GSK3A, GAPDH; Blot 3: phospho-Beta-catenin Ser33/37/Thr41 (marked for degradation), GSK3B, GAPDH].

c) Representative 3D DNA FISH images of chr22 domains after KD of GSK3A using four independent siRNA constructs. Dotted white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

d) Fold change in spatial overlap between chr22:D1 and D2 after KD of GSK3A using four independent siRNA constructs and a pool of all four constructs. * P-value < 0.05, ***

P-value < 0.001, *** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Alleles for RNAi Control (n = 961), GSK3A Pool (n = 383), GSK3A #1 (n = 807), GSK3A #2 (n = 586), GSK3A #3 (n = 314), GSK3A #4 (n = 571)].

e) Western blot of whole cell lysate after GSK3A KD using four independent siRNA constructs and pool of all four constructs leads to selective depletion of GSK3A. Proteins were labeled with HRP-linked antibodies. All lanes cropped from same blot to show only relevant lanes.

f) Representative 3D DNA FISH images of chr22 domains after 24 hour DMSO- and BRD0705- (GSK3Ai, 20 μM) treated HCT-116 cells. Dotted white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

g) Fold change in spatial overlap after 24 hour 20 μM BRD0705 treatment for four TAD boundaries of varying insulation scores. ** P-value < 0.01, **** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Number of alleles for chr3:D1-D2: DMSO Control (n = 556), BRD0705 (n = 502); chr3:D2-D3: Control (n = 542), BRD0705 (n = 491); chr22:D1-D2: Control (n = 732), BRD0705 (n = 440); chr22:D2-D3: Control (n = 687), BRD0705 (n = 427)].
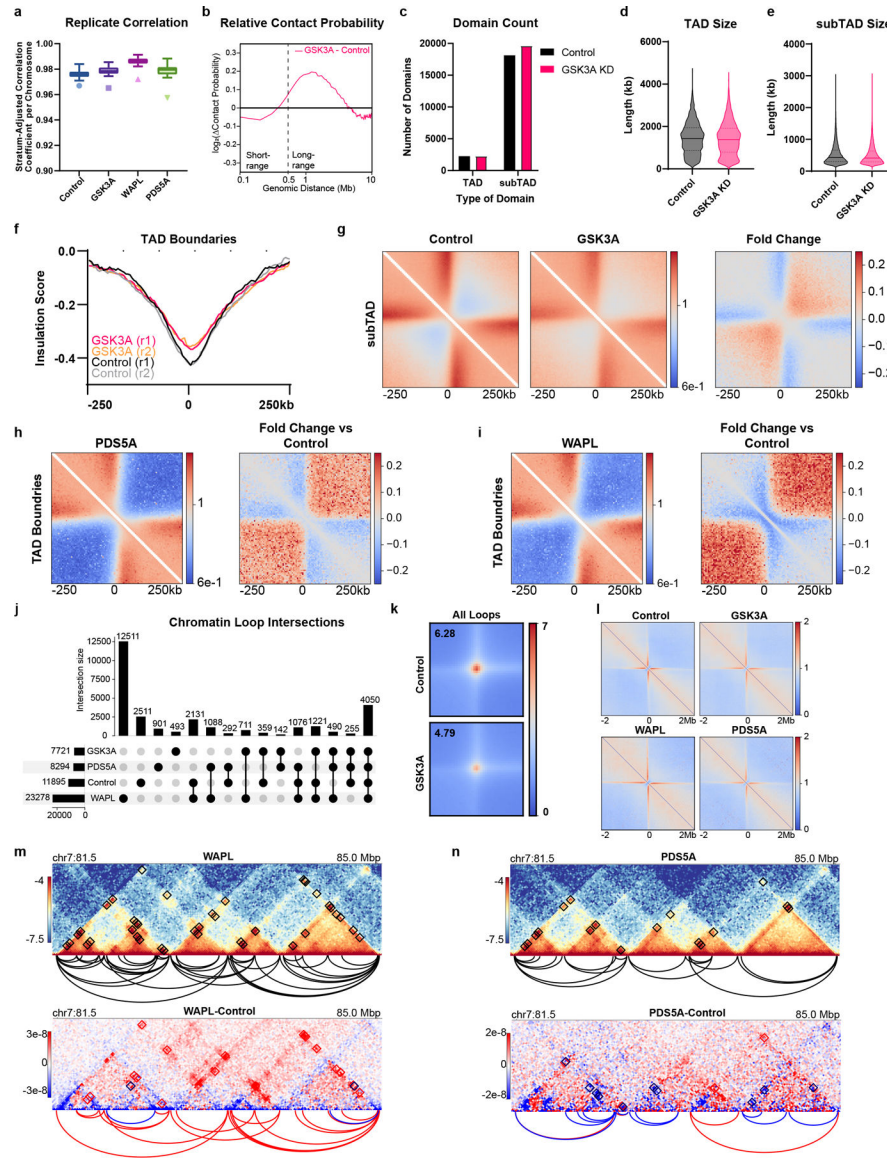
Extended Data Figure 4. Hi-C reveals long-range looping interactions are gained in GSK3A KD, Related to Figure 4.

a) Tukey box plots for stratum-adjusted correlation coefficient per chromosome for replicates of Hi-C in control, GSK3A KD, WAPL KD and PDS5A KD. Lower whisker = 25th percentile – 1.5*IQR, lower box bound = 25th percentile, middle of box = median, upper box bound = 75th percentile, upper whisker = 75th percentile + 1.5*IQR. [n = 23 chromosomes for each condition].

b) Log2 of difference (GSK3A KD replicate 2 – control replicate 2) in contact probability as a function of genomic distance (log scale). Dotted line at 500 kb indicates divide between short-range intra-TAD and long-range inter-TAD interactions.

c) Domain counts for TAD [Control (n = 2,253), GSK3A KD (n = 2,271)] and subTADs [Control (n = 18,117), GSK3A KD (n = 19,632)].

d) Violin plot of TAD length in kb. Solid line is median, dotted lines are 25th and 75th percentiles. [Control (n = 2,253), GSK3A KD (n = 2,271)].

e) Violin plot of subTAD length in kb. [Control (n = 18,117), GSK3A KD (n = 19,632)].

f) Insulation score pileup at control TAD boundaries in control replicate 1, control replicate 2, GSK3A KD replicate 1 and GSK3A KD replicate 2 Hi-C.

g) 3D pileup plots of Hi-C interactions in control and GSK3A KD at control subTAD boundaries, as well as log2 fold change in interactions across those boundaries.

h) 3D pileup plots of Hi-C interactions in control and WAPL KD at control TAD boundaries, as well as log2 fold change in interactions across those boundaries.

i) 3D pileup plots of Hi-C interactions in control and PDS5A KD at control TAD boundaries, as well as log2 fold change in interactions across those boundaries.

j) Upset plot of chromatin loop intersections between control, GSK3A KD, PDS5A KD and WAPL KD. [Control (n = 11,895), GSK3A KD (n = 7,721), PDS5A KD (n = 8,294), WAPL (n = 23,278)].

k) Aggregate peak analysis of control and GSK3A KD at union set of loops across control and GSK3A KD (n = 13,491).

l) 3D pileup plots of Hi-C interactions in control, GSK3A KD, WAPL KD and PDS5A KD at stripes detected in control.

m) Hi-C contact matrices for control, WAPL KD, and WAPL KD – Control subtraction at chr7:81.5–85 Mbp. Looping interactions highlighted with squares on contact map and arcs below contact map. For subtraction map, red loops are gained in WAPL KD, and blue loops are lost.

n) Hi-C contact matrices for control, PDS5A KD, and PDS5A KD – Control subtraction at chr7:81.5–85 Mbp. Looping interactions highlighted with squares on contact map and arcs below contact map. For subtraction map, red loops are gained in PDS5A KD, and blue loops are lost.
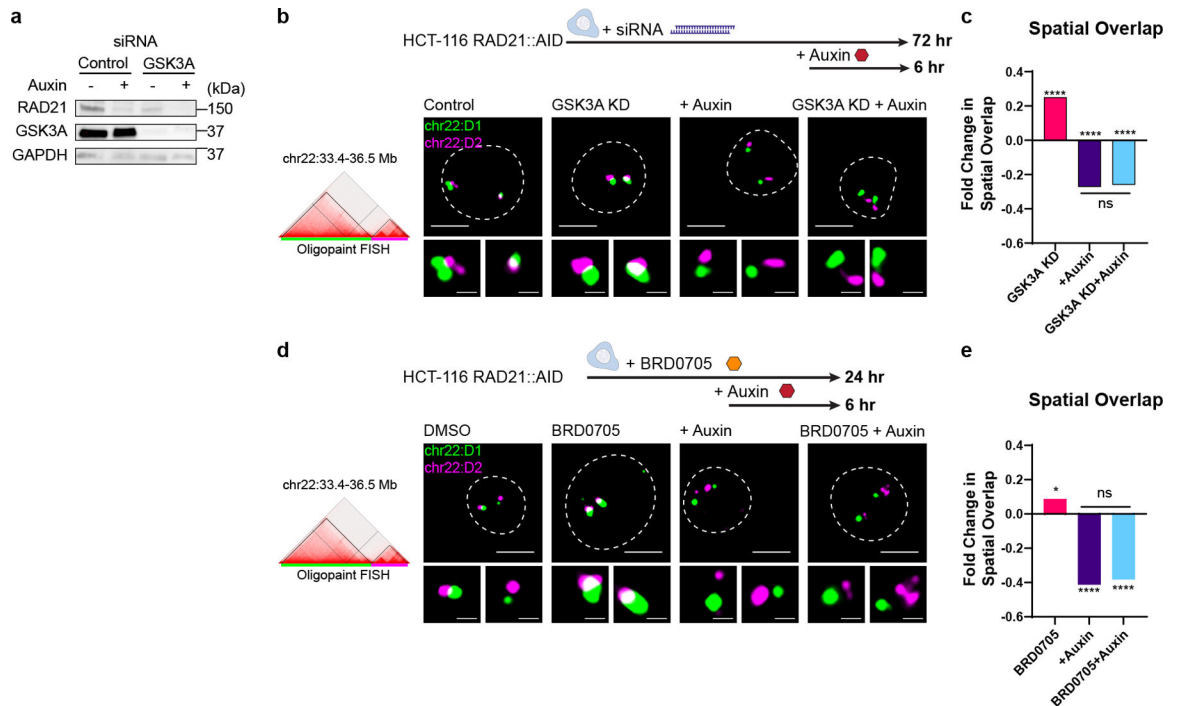
Extended Data Figure 5: GSK3A regulates genome folding in a cohesin-dependent manner.

(a) Western blot of whole cell lysate after six hour auxin treatment in HCT-116-RAD21-mClover-AID cells and depletion of GSK3A after 72 hour RNAi KD. Proteins are labeled with HRP-linked antibodies. All bands and lanes from same blot.

(a) Representative 3D DNA FISH images of chr22 domains following six-hour RAD21 auxin-inducible degradation, 72 hour GSK3A KD or both. Dotted white line indicates nuclear edge. Scale bar for nucleus, 5 μm; scale bar for spots, 1 μm.

(b) Fold change in spatial overlap between chr22:D1-D2 after GSK3A KD, RAD21 degradation or both. *ns* P-value > 0.05, **** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Alleles for Control (n = 669); GSK3A KD (n = 712); +Auxin (n = 680); GSK3A KD + Auxin (n = 560)].

(b) Representative 3D DNA FISH images of chr22 domains after auxin-inducible degradation of RAD21, 24 hour treatment with BRD0705 (GSK3Ai) or both. Dotted white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

(c) Fold change in mean spatial overlap at chr22:D1-D2. *ns* P-value > 0.05, * P-value < 0.05, **** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Alleles for control (n = 350), BRD0705 (n = 302), +Auxin (n = 525), BRD0705+Auxin (n = 280)].
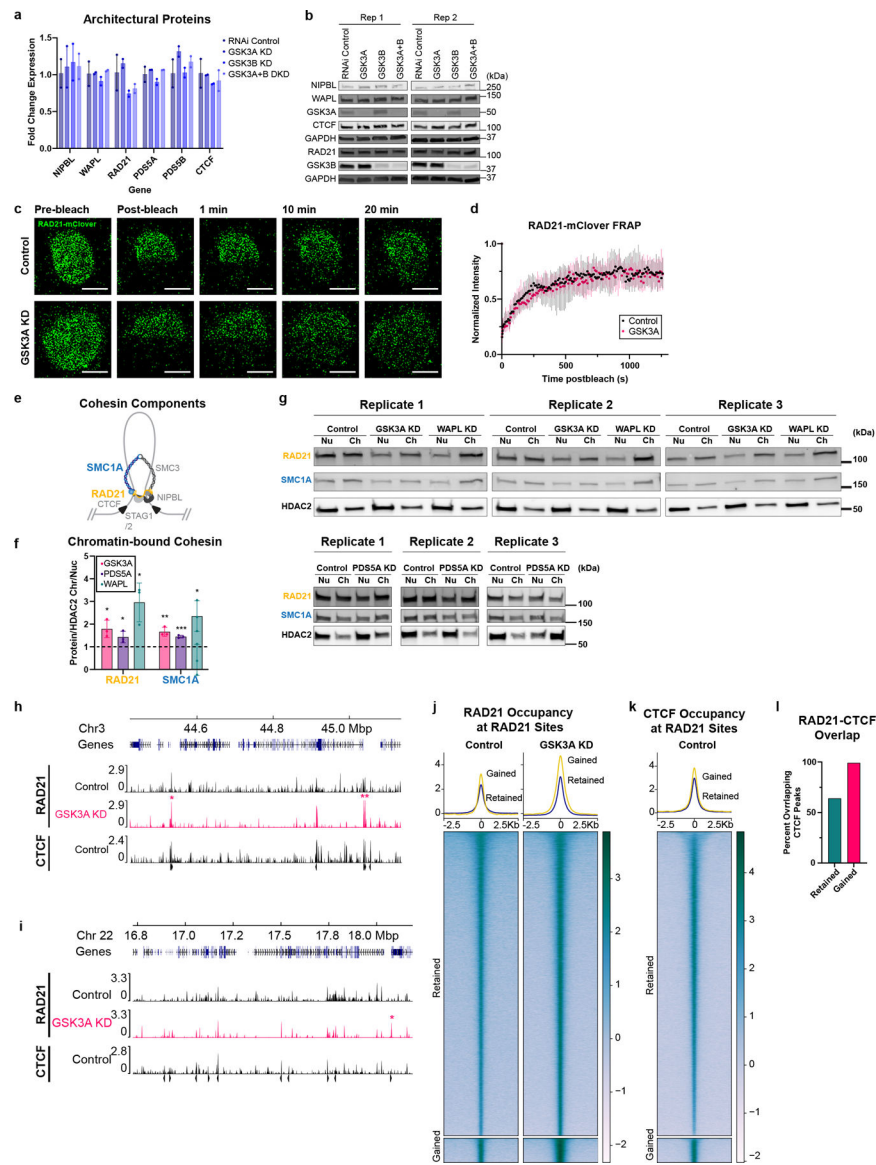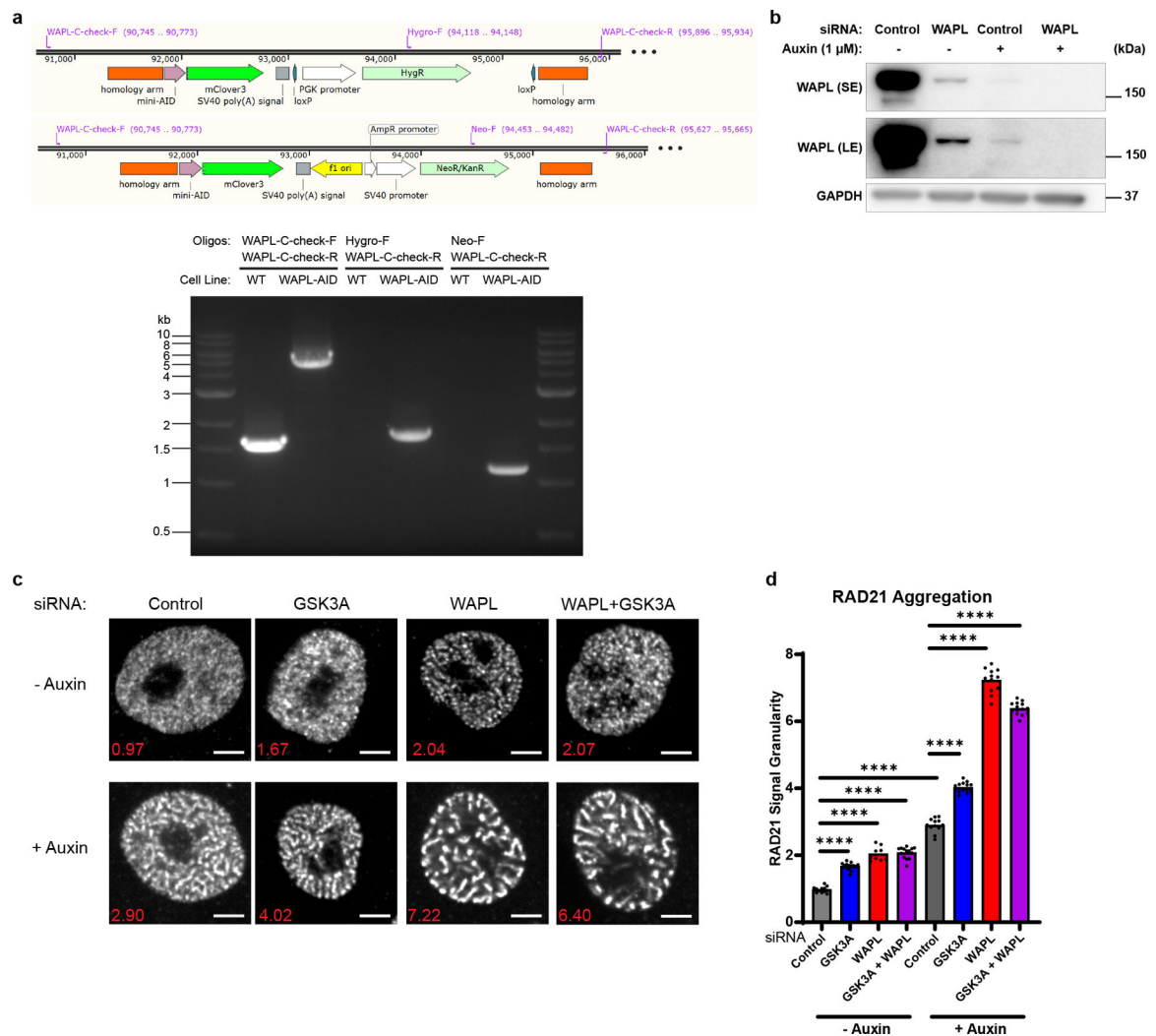
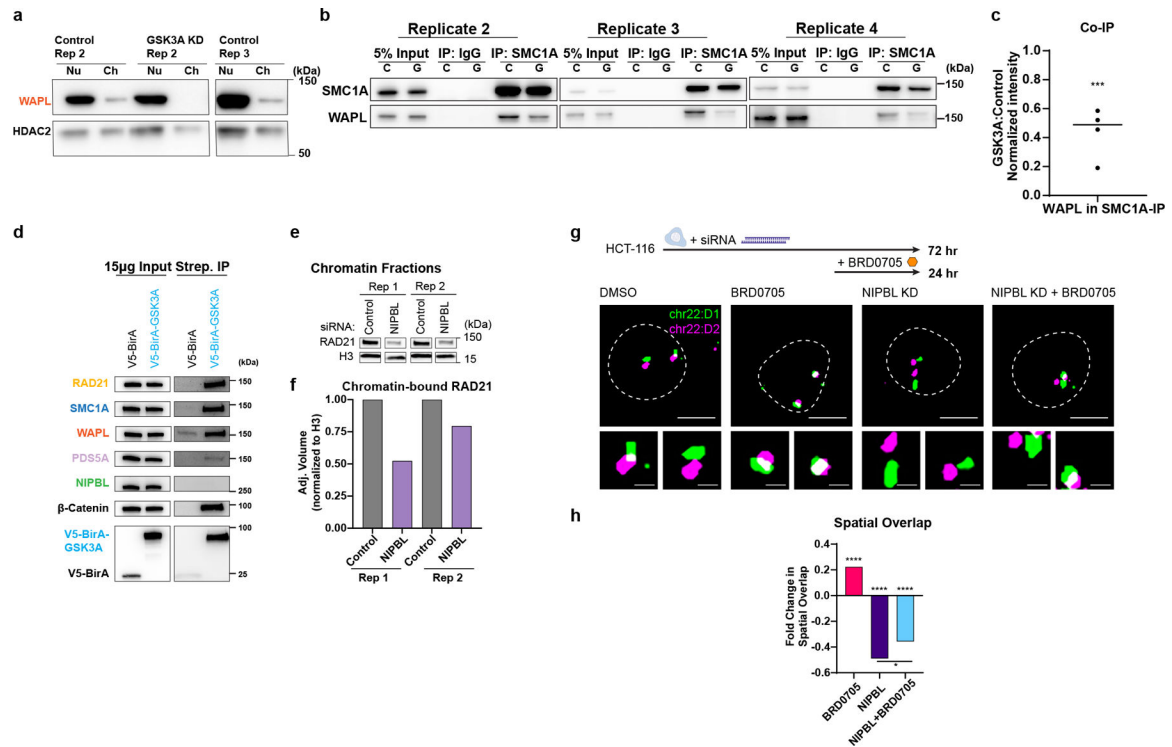**Extended Data Figure 6: GSK3A regulates levels of cohesin on chromatin.**

a) RT-qPCR of architectural proteins after GSK3A, GSK3B or GSK3A+B KD. Bar is mean of two biological replicates +/- SD. Each biological replicate data point is the average of three technical replicates.

b) Western blots to cohesin components and CTCF in whole cell lysate after GSK3A, GSK3B or GSK3A+B KD. Proteins are labeled with fluorescent antibodies. Two biological replicates shown, run on the same gel with ladder lanes cropped out. [Images in order top to bottom. From Blot 1: NIPBL, WAPL, GSK3A, CTCF, GAPDH; Blot 2: RAD21, GSK3B, GAPDH].

c) Representative images from half-nuclear FRAP of RAD21-mClover in control and GSK3A KD HCT-116-RAD21-mClover-AID cells. Scale bar, 5 μm.

d) Half-nuclear FRAP curves for RAD21-mClover in control and GSK3A KD HCT-116-RAD21-mClover-AID cells. [Control (n = 22 nuclei), GSK3A KD (n = 21 nuclei)]. Each point is median of all nuclei for that condition +/- 95% CI.

e) Model of core cohesin components bound to chromatin: RAD21 (yellow) and SMC1A (blue).

f) Chromatin-bound to nucleoplasmic ratio of RAD21 or SMC1A, normalized to loading control (HDAC2). Dotted line represents control. Bar is mean +/- SD of three biological replicates. Western blots in Extended Data Fig. 6g. * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001, two-tailed t-test of each condition vs negative control. P-values left to right: 0.0236, 0.0425, 0.0167, 0.0045, 0.0002, 0.0263.

g) Western blots of nuclear and chromatin-bound fractions of cohesin components following GSK3A KD, PDS5A KD and WAPL KD. Proteins are labeled using fluorescent antibodies. Corresponds to Extended Data Fig. 6f. Three total biological replicates.

h) Representative genome browser image of chr3:44.5–45.2Mbp with RAD21 ChIP signal for control and GSK3A KD and CTCF ChIP signal for control. * = significantly increased RAD21 sites.

i) Genome browser track chr22:16.8–18.1 Mbp with RAD21 and CTCF ChIP-seq signal shown. * = significantly gained RAD21 site.

j) RAD21 occupancy for control and GSK3A KD at retained (n = 50,874) and significantly gained (n = 4,069) RAD21 sites.

k) CTCF occupancy for control at retained (n = 50,874) and gained (n = 4,069) RAD21 sites.

l) Percentage of retained and gained RAD21 sites that overlap with at least one CTCF peak.

Extended Data Figure 7: GSK3A genetically interacts with WAPL to regulate cohesin levels.
a) Plasmid maps for inserting WAPL-mAID2-mClover3-Hygromycin and WAPL-mAID2-mClover3-Neomycin into parental HCT-116-OsTir1(F74G) cell line to create HCT-116-WAPL-AID2 cell line. DNA electrophoresis below confirms insert of both cassettes.
b) Western blot of WAPL from HCT116-WAPL-AID lysate +/- WAPL KD and +/- auxin-inducible degradation of WAPL. LE = long exposure, SE = short exposure.
c) Representative images of RAD21 immunofluorescence after auxin-induced degradation of WAPL and/or siRNA treatment. Median RAD21 signal granularity noted in red. Quantification in Extended Data Fig. 5. Scale bar nucleus, 5 μm.
d) Median RAD21 signal granularity for +/- siRNA treatment and +/- degradation of WAPL. Each bar is median. [Replicate wells for -Auxin, Control (n = 13), GSK3A (n = 14), WAPL (n = 9), GSK3A + WAPL (n = 14). For +Auxin, Control (n = 12), GSK3A (n = 14), WAPL (n = 12), GSK3A +WAPL (n = 14)] **** P-value < 0.0001, two-tailed t-test.

Extended Data Figure 8. Additional data supporting role of GSK3A in regulating cohesin unloading through WAPL, Related to Figure 5.

a) Additional biological replicates of western blot of nucleoplasmic and chromatin-bound WAPL in control and GSK3A KD. Proteins labeled with HRP-linked antibodies.

b) Additional biological replicates of co-immunoprecipitation of chromatin fractions in control and GSK3A KD with blotting of WAPL in SMC1A-IP.

c) Quantification of WAPL normalized intensity in SMC1A-IP. *** P-value = 0.0007, two-tailed t-test.

d) Additional biological replicate of GSK3A Turbo-ID. Western blots for cohesin components input lysate or lysate from with 24 hour biotin incubation with either control construct (V5-BirA) or GSK3A TurboID (GSK3A-V5-BirA). Proteins labeled with HRP-linked antibodies.

e) Western blot of RAD21 in chromatin fractions of control and 72 hour NIPBL KD. Proteins are labeled with fluorescent antibodies. All bands from same blot, cropped for clarity. Two biological replicates represented.

f) Quantification of western blot in Extended Data Figure 6e. Protein intensity was adjusted for background, normalized to H3 volume and then normalized to the control lane for two biological replicates shown.

g) Representative 3D DNA FISH images of chr22 domains after 72 hour NIPBL KD, 24 hour 20 μM BRD0705 treatment, or both. Dotted white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

h) Fold change in mean spatial overlap between chr22:D1-D2. * P-value < 0.05, **** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Alleles for Control+DMSO (n = 732); Control+BRD0705 (n = 440); NIPBL+DMSO (n = 543); NIPBL+BRD0705 (n = 373)].

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

DATA AVAILABILITY

Datasets reported in this paper are available at the Gene Expression Omnibus with accession number GEO: GSE199607.

References

1. 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). [PubMed: 29217591]

2. Beagan JA & Phillips-Cremins JE On the existence and functionality of topologically associating domains. Nat. Genet 1–9 (2020) doi:10.1038/s41588-019-0561-1. [PubMed: 31911675]

3. Davidson IF & Peters J-M Genome folding through loop extrusion by SMC complexes. Nat. Rev. Mol. Cell Biol 1–20 (2021) doi:10.1038/s41580-021-00349-7. [PubMed: 33244181]

4. Rao SSP et al. Cohesin Loss Eliminates All Loop Domains. Cell 171, 305–320.e24 (2017). [PubMed: 28985562]

5. Schwarzer W et al. Two independent modes of chromatin organization revealed by cohesin removal. Nature 551, 51–56 (2017). [PubMed: 29094699]

6. Merkenschlager M & Nora EP CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annu. Rev. Genomics Hum. Genet 17, 17–43 (2016). [PubMed: 27089971]

7. Luppino JM et al. Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes. Nat. Genet 1–9 (2020) doi:10.1038/s41588-020-0647-9. [PubMed: 31911675]

8. Kriz AJ, Colognori D, Sunwoo H, Nabet B & Lee JT Balancing cohesin eviction and retention prevents aberrant chromosomal interactions, Polycomb-mediated repression, and X-inactivation. Mol. Cell 81, 1970–1987.e9 (2021). [PubMed: 33725485]

9. Linares-Saldana R et al. BRD4 orchestrates genome folding to promote neural crest differentiation. Nat. Genet 53, 1480–1492 (2021). [PubMed: 34611363]

10. Liu NQ et al. Rapid depletion of CTCF and cohesin proteins reveals dynamic features of chromosome architecture. 2021.08.27.457977 https://www.biorxiv.org/content/10.1101/2021.08.27.457977v1 (2021) doi:10.1101/2021.08.27.457977.

11. Beliveau BJ et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. Proc. Natl. Acad. Sci 109, 21301–21306 (2012). [PubMed: 23236188]

12. Ciosk R et al. Cohesin's Binding to Chromosomes Depends on a Separate Complex Consisting of Scc2 and Scc4 Proteins. Mol. Cell 5, 243–254 (2000). [PubMed: 10882066]

13. Kueng S et al. Wapl Controls the Dynamic Association of Cohesin with Chromatin. Cell 127, 955–967 (2006). [PubMed: 17113138]

14. Haarhuis JHI et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell 169, 693–707.e14 (2017). [PubMed: 28475897]

15. Overington JP, Al-Lazikani B & Hopkins AL How many drug targets are there? Nat. Rev. Drug Discov 5, 993–996 (2006). [PubMed: 17139284]

16. Sakharkar MK & Sakharkar KR Targetability of Human Disease Genes. Curr. Drug Discov. Technol 4, 48–58 (2007). [PubMed: 17630928]

17. Boyle S et al. A central role for canonical PRC1 in shaping the 3D nuclear landscape. Genes Dev 34, 931–949 (2020). [PubMed: 32439634]

18. Szklarczyk D et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47, D607–D613 (2019). [PubMed: 30476243]

19. Doble BW, Patel S, Wood GA, Kockeritz LK & Woodgett JR Functional Redundancy of GSK-3α and GSK-3β in Wnt/β-Catenin Signaling Shown by Using an Allelic Series of Embryonic Stem Cell Lines. Dev. Cell 12, 957–971 (2007). [PubMed: 17543867]

20. Sutherland C What Are the bona fide GSK3 Substrates? Int. J. Alzheimerx2019s Dis 2011, e505607 (2011).

21. Beurel E, Grieco SF & Jope RS Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. Pharmacol. Ther 0, 114–131 (2015).

22. Chen X et al. A Chemical-Genetic Approach Reveals the Distinct Roles of GSK3α and GSK3β in Regulating Embryonic Stem Cell Fate. Dev. Cell 43, 563–576.e4 (2017). [PubMed: 29207259]

23. Shinde MY et al. Phosphoproteomics reveals that glycogen synthase kinase-3 phosphorylates multiple splicing factors and is associated with alternative splicing. J. Biol. Chem 292, 18240–18255 (2017). [PubMed: 28916722]

24. Peifer M, Pai L-M & Casey M Phosphorylation of the Drosophila Adherens Junction Protein Armadillo: Roles for Wingless Signal and Zeste-white 3 Kinase. Dev. Biol 166, 543–556 (1994). [PubMed: 7529201]

25. Yost C et al. The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. Genes Dev 10, 1443–1454 (1996). [PubMed: 8666229]

26. Wagner FF et al. Exploiting an Asp-Glu "switch" in glycogen synthase kinase 3 to design paralog-selective inhibitors for use in acute myeloid leukemia. Sci. Transl. Med 10, (2018).

27. Engler TA et al. Substituted 3-imidazo[1,2-a]pyridin-3-yl- 4-(1,2,3,4-tetrahydro-[1,4]diazepino-[6,7,1-hi]indol-7-yl)pyrrole-2,5-diones as highly selective and potent inhibitors of glycogen synthase kinase-3. J. Med. Chem 47, 3934–3937 (2004). [PubMed: 15267232]

28. An WF et al. Discovery of Potent and Highly Selective Inhibitors of GSK3b. in Probe Reports from the NIH Molecular Libraries Program (National Center for Biotechnology Information (US), 2010).

29. Vian L et al. The Energetics and Physiological Impact of Cohesin Extrusion. Cell 173, 1165–1178.e20 (2018). [PubMed: 29706548]

30. Barrington C et al. Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. Nat. Commun 10, 2908 (2019). [PubMed: 31266948]

31. Natsume T, Kiyomitsu T, Saga Y & Kanemaki MT Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. Cell Rep 15, 210–218 (2016). [PubMed: 27052166]

32. Tedeschi A et al. Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature 501, 564–568 (2013). [PubMed: 23975099]

33. Branon TC et al. Efficient proximity labeling in living cells and organisms with TurboID. Nat. Biotechnol 36, 880–887 (2018). [PubMed: 30125270]

34. Kikuchi S, Borek DM, Otwinowski Z, Tomchick DR & Yu H Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy. Proc. Natl. Acad. Sci 113, 12444–12449 (2016). [PubMed: 27791135]

35. Petela NJ et al. Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. Mol. Cell 70, 1134–1148.e7 (2018). [PubMed: 29932904]

36. Kean CM et al. Decreasing Wapl dosage partially corrects embryonic growth and brain transcriptome phenotypes in Nipbl+/− embryos. Sci. Adv 8, eadd4136 (2022). [PubMed: 36449618]

37. Luppino JM et al. Co-depletion of NIPBL and WAPL balance cohesin activity to correct gene misexpression. PLOS Genet 18, e1010528 (2022). [PubMed: 36449519]

38. Joyce EF, Williams BR, Xie T & Wu C. -ting. Identification of Genes That Promote or Antagonize Somatic Homolog Pairing Using a High-Throughput FISH–Based Screen. PLoS Genet 8, e1002667 (2012). [PubMed: 22589731]

39. Shachar S, Voss TC, Pegoraro G, Sciascia N & Misteli T Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. Cell 162, 911–923 (2015). [PubMed: 26276637]

40. Finn EH et al. Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. Cell 0, (2019).

41. Chin CV et al. Cohesin mutations are synthetic lethal with stimulation of WNT signaling. eLife 9, e61405 (2020). [PubMed: 33284104]

42. Grazioli P et al. Lithium as a possible therapeutic strategy for Cornelia de Lange syndrome. Cell Death Discov 7, 1–11 (2021).

43. Bottai D et al. Modeling Cornelia de Lange syndrome in vitro and in vivo reveals a role for cohesin complex in neuronal survival and differentiation. Hum. Mol. Genet 28, 64–73 (2019). [PubMed: 30239720]

44. Kaidanovich-Beilin O & Woodgett J GSK-3: Functional Insights from Cell Biology and Animal Models. Front. Mol. Neurosci 4, (2011).

45. Hegemann B et al. Systematic Phosphorylation Analysis of Human Mitotic Protein Complexes. Sci. Signal (2011) doi:10.1126/scisignal.2001993.

46. Liang C et al. A kinase-dependent role for Haspin in antagonizing Wapl and protecting mitotic centromere cohesion. EMBO Rep 19, 43–56 (2018). [PubMed: 29138236]

47. Beliveau BJ et al. OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. Proc. Natl. Acad. Sci 115, E2183–E2192 (2018). [PubMed: 29463736]

48. Bintu B et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science 362, eaau1783 (2018). [PubMed: 30361340]

49. Chen KH, Boettiger AN, Moffitt JR, Wang S & Zhuang X Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, aaa6090 (2015). [PubMed: 25858977]

50. Mateo LJ et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature 568, 49 (2019). [PubMed: 30886393]

SUPPLEMENTARY REFERENCES

51. Saito Y & Kanemaki MT Targeted Protein Depletion Using the Auxin-Inducible Degron 2 (AID2) System. Curr. Protoc 1, e219 (2021). [PubMed: 34370399]

52. Beckwith KS et al. Visualization of loop extrusion by nanoscale 3D DNA tracing in single human cells. 2021.04.12.439407 Preprint at 10.1101/2021.04.12.439407 (2022).

53. Shah PP et al. Pathogenic LMNA variants disrupt cardiac lamina-chromatin interactions and de-repress alternative fate genes. Cell Stem Cell 28, 938–954.e9 (2021). [PubMed: 33529599]

54. Rhodes J, Mazza D, Nasmyth K & Uphoff S Scc2/Nipbl hops between chromosomal cohesin rings after loading. eLife 6, e30000 (2017). [PubMed: 28914604]

55. Cho KF et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. Nat. Protoc 15, 3971–3999 (2020). [PubMed: 33139955]

56. Ran FA et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc 8, 2281–2308 (2013). [PubMed: 24157548]

57. Yesbolatova A et al. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. Nat. Commun 11, 5701 (2020). [PubMed: 33177522]

58. McQuin C et al. CellProfiler 3.0: Next-generation image processing for biology. PLOS Biol 16, e2005970 (2018). [PubMed: 29969450]

59. Li CH & Lee CK Minimum cross entropy thresholding. Pattern Recognit 26, 617–625 (1993).

60. Otsu N A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man Cybern 9, 62–66 (1979).

61. Drasgow F Polychoric and Polyserial Correlations. in Encyclopedia of Statistical Sciences (John Wiley & Sons, Ltd, 2006). doi:10.1002/0471667196.ess2014.pub2.

62. Ollion J, Cochennec J, Loll F, Escudé C & Boudier T TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics 29, 1840–1841 (2013). [PubMed: 23681123]

63. Ollion J, Cochennec J, Loll F, Escudé C & Boudier T Analysis of Nuclear Organization with TANGO, Software for High-Throughput Quantitative Analysis of 3D Fluorescence Microscopy Images. in The Nucleus (ed. Hancock R) 203–222 (Springer, 2015). doi:10.1007/978-1-4939-1680-1_16.

64. Stirling DR et al. CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics 22, 433 (2021). [PubMed: 34507520]

65. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

66. Schneider VA et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res 27, 849–864 (2017). [PubMed: 28396521]

67. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012). [PubMed: 22388286]

68. Quinlan AR & Hall IM BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010). [PubMed: 20110278]

69. Ramírez F et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–W165 (2016). [PubMed: 27079975]

70. Zhang Y et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol 9, R137 (2008). [PubMed: 18798982]

71. Wu D-Y, Bittencourt D, Stallcup MR & Siegmund KD Identifying differential transcription factor binding in ChIP-seq. Front. Genet 6, (2015).

72. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550 (2014). [PubMed: 25516281]

73. Lopez-Delisle L et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. Bioinformatics 37, 422–423 (2021). [PubMed: 32745185]

74. Servant N et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 16, 259 (2015). [PubMed: 26619908]

75. Yang T et al. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. Genome Res 27, 1939–1949 (2017). [PubMed: 28855260]

76. Fernandez LR, Gilgenast TG & Phillips-Cremins JE 3DeFDR: statistical methods for identifying cell type-specific looping interactions in 5C and Hi-C data. Genome Biol 21, 219 (2020). [PubMed: 32859248]

77. Emerson DJ et al. Cohesin-mediated loop anchors confine the locations of human replication origins. Nature 606, 812–819 (2022). [PubMed: 35676475]

78. Knight PA & Ruiz D A fast algorithm for matrix balancing. IMA J. Numer. Anal 33, 1029–1047 (2013).

79. Norton HK et al. Detecting hierarchical genome folding with network modularity. Nat. Methods 15, 119–122 (2018). [PubMed: 29334377]

80. Zhang H et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. Nature 1–5 (2019) doi:10.1038/s41586-019-1778-y.

81. Wolff J et al. Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res 48, W177–W184 (2020). [PubMed: 32301980]

82. Open2C et al. Cooltools: enabling high-resolution Hi-C analysis in Python. 2022.10.31.514564 Preprint at 10.1101/2022.10.31.514564 (2022).

83. Roayaei Ardakany A, Gezer HT, Lonardi S & Ay F Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. Genome Biol 21, 256 (2020). [PubMed: 32998764]

84. Flyamer IM, Illingworth RS & Bickmore WA Coolpup.py: versatile pile-up analysis of Hi-C data. Bioinformatics 36, 2980–2985 (2020). [PubMed: 32003791]

85. Yoon S, Chandra A & Vahedi G Stripenn detects architectural stripes from chromatin conformation data using computer vision. Nat. Commun 13, 1602 (2022). [PubMed: 35332165]

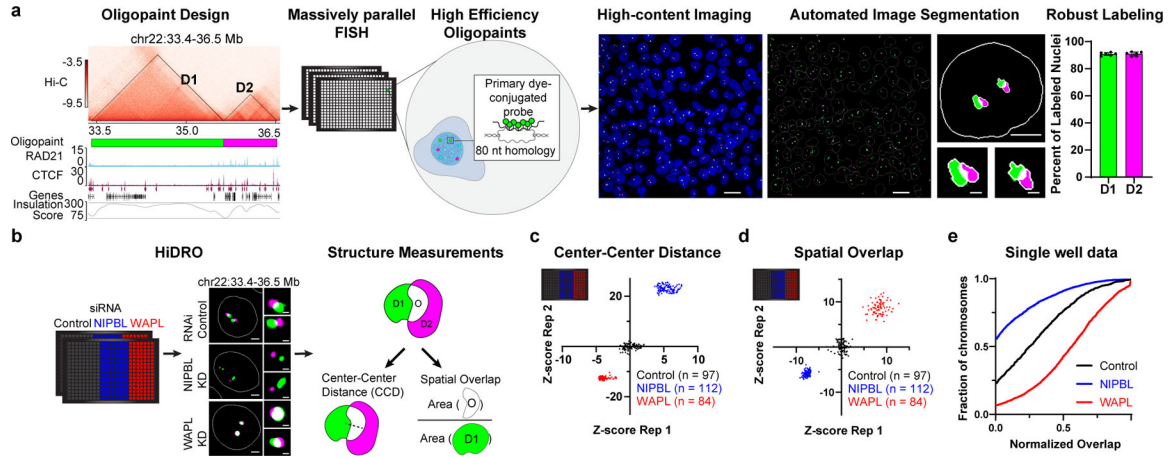86. Hnisz D Super-Enhancers in the Control of Cell Identity and Disease. 28.

Figure 1: Development of HiDRO.

(a) Workflow for DNA HiDRO at two consecutive TADs (D1, D2) on chromosome 22:33.4–36.5Mb. Solid white line following segmentation indicates nuclear edge. Scale bar field, 10 μm; Scale bar nucleus, 5 μm; Scale bar spots, 1 μm. Far right graph: the efficiency of each chr22 spot is measured by the percentage of nuclei with one or more spots. Bars for spot efficiency are mean of six biological replicate wells +/- SD and each replicate well measured > 2,000 alleles.

(b) HiDRO with siRNA perturbations at chr22 D1 & D2. Black wells are seeded with non-targeting control, red wells are seeded with NIPBL siRNA, and blue wells are seeded with WAPL siRNA. Structure measurements quantify inter-domain interactions including center-center distance (CCD) and spatial overlap normalized to area of either D1 or D2. Solid white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

(c) Robust z-scores for D1-D2 center-center distance of control (n = 97 wells), NIPBL KD (n = 112 wells) and WAPL KD (n = 84 wells) across two biological replicates.

(d) Spatial overlap (normalized to D1 area) z-scores shown for two biological replicates.

(e) Cumulative frequency plot for spatial overlap from one well of DNA HiDRO data for each condition. X-axis is area of spatial overlap between D1 and D2 normalized to D1 area and Y-axis is fraction of alleles measured with that amount of overlap. [Alleles for Control (n = 1137); NIPBL (n = 3,775); WAPL (n = 2,648)]. **** P<0.0001, two-tailed Mann-Whitney U-test.
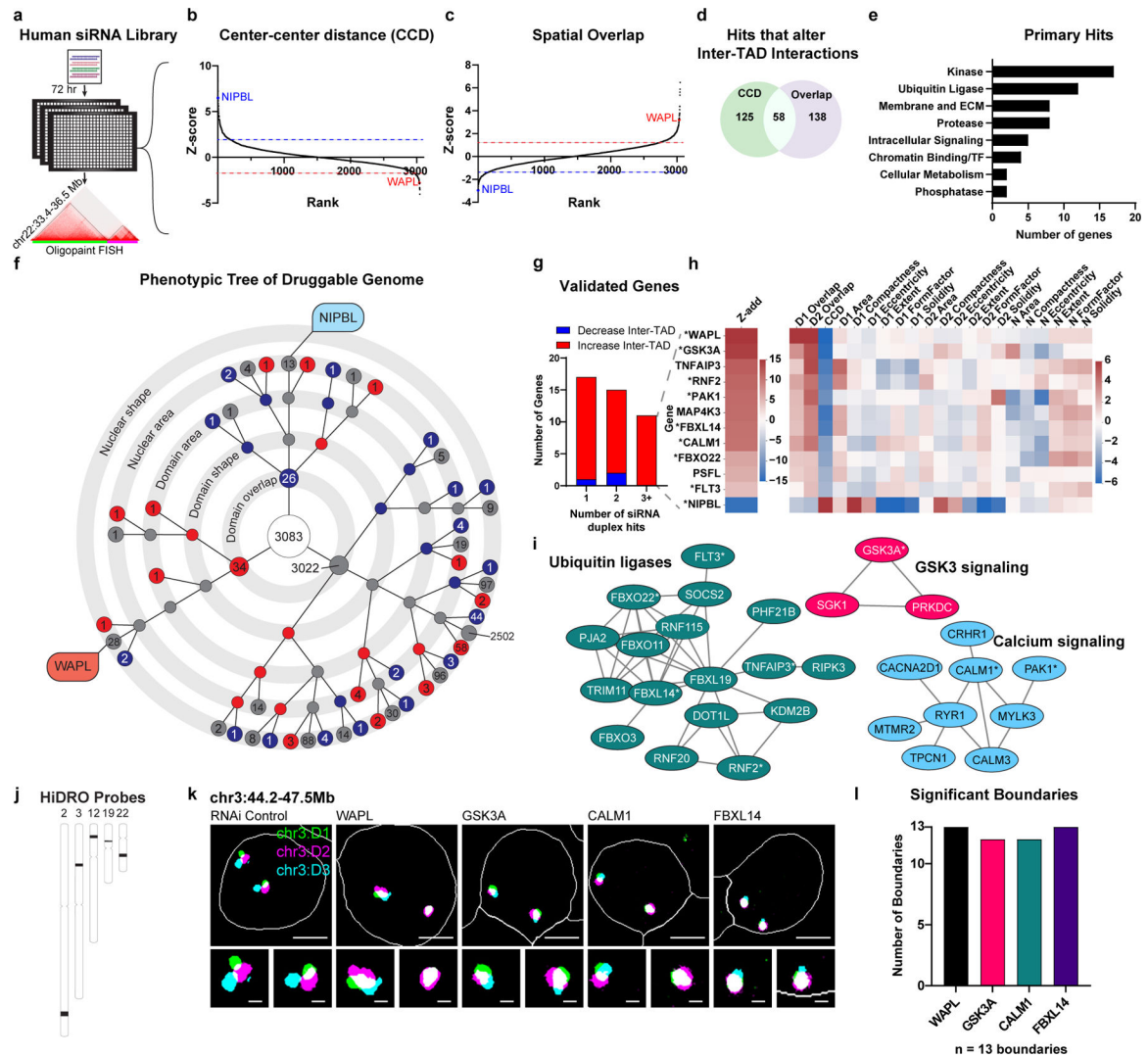
Figure 2: HiDRO identifies novel regulators of genome folding.

a) Workflow for primary HiDRO Screen of the human Druggable Genome siRNA library.

b) Rank ordered robust z-scores for center-center distance (CCD) for all genes tested. Z-scores are average of two biological replicates. Dotted lines represent cutoffs of 1.5 (blue) or −1.5 (red).

c) Rank ordered robust z-scores for normalized spatial overlap. Z-scores are average of two biological replicates. Cutoffs of 1.5 (blue) or −1.5 (red).

d) Venn diagram of genes that alteredCCD or spatial overlap. n = 337.

e) Protein classes of primary hits that altered both CCD and spatial overlap. n = 58.

f) Phenotypic tree of 3,083 genes in primary screen scored across five measurement categories. Red = increase, blue = decrease, gray = no significant change. Change is z-score 1.5 or −1.5 across both replicates for any measurement in that category.

g) Number of validation screen genes in that replicated overlap phenotype for individual siRNA duplexes.Blue genes decrease and red genes increase inter-TAD interactions.

h) Heat map of image phenotypes for genes with 3 validated duplexes, rank-ordered by additive z-score (z-add = sum of CCD, D1 overlap, and D2 overlap z-scores). * = Top 10 rank ordered validated hit, seesee Fig. 2i.

i) STRING network clusters of validated hits exhibiting connections with 2 validated genes. Green: Ubiquitin ligases and chromatin modifiers, Red: GSK3 signaling, Blue: calcium signaling. * = Top 10 rank ordered validated hit.

j) Ideograms of Oligopaint probes to thirteen TAD and subTAD boundaries.

k) Representative DNA HiDRO images of chr3 domains after RNAi KD of top druggable genome hits. Solid white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

l) Number of boundaries with significantly altered CCD after KD of top hits. N = 13 boundaries.
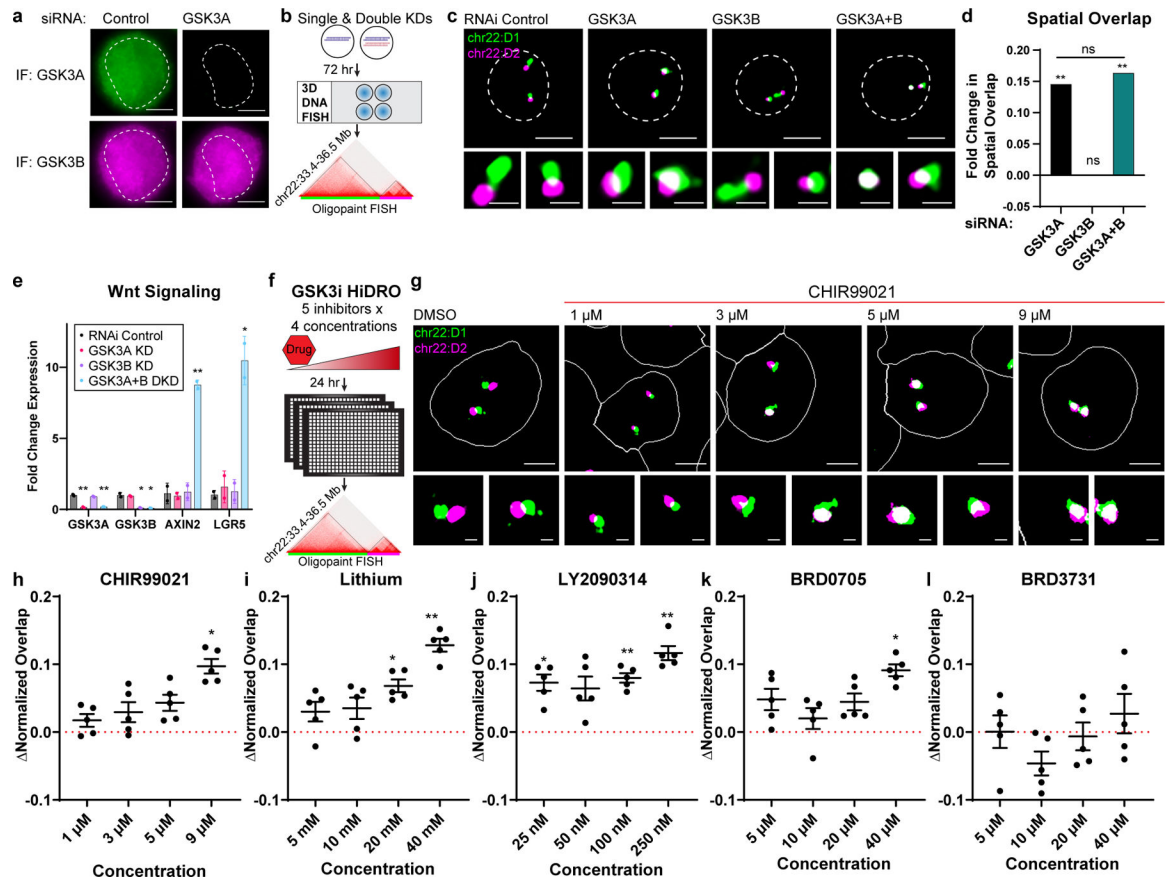
Figure 3: GSK3A has noncanonical role in genome folding.

(a) Immunofluorescence (IF) of GSK3A and GSK3B after GSK3A KD. Dotted white line indicates nuclear edge determined by DAPI staining (not shown). Scale bar, 5 μm.

(b) Validation of HiDRO by high-resolution 3D DNA FISH at chr22 domains.

(c) Representative 3D DNA FISH images of chr22 domains after GSK3A and/or GSK3B KD. Dotted white line indicates nuclear edge. Scale bar nucleus, 5 μm; scale bar spots, 1 μm.

(d) Fold change in mean spatial overlap versus control [Alleles for RNAi Control (n = 696); GSK3A (n = 461); GSK3B (n = 636); GSK3A+B (n = 596)]. ns P-value > 0.05, ** P-value < 0.01, two-tailed Mann-Whitney U-test. P-values vs Control: GSK3A = 0.0015, GSK3B = 0.7979, GSK3A+B = 0.0015; GSK3A vs GSK3A+B = 0.8901.

(e) Reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) to WNT targets AXIN2 and LGR5 after GSK3A +/- GSK3B KD. Bar is mean of two biological replicates. Each biological replicate is average of three technical replicates. * P-value < 0.05, ** P-Value < 0.01, two-tailed t-test. P-values left to right: 0.0053, 0.0047, 0.0200, 0.0186, 0.0055, 0.0165.

(f) DNA HiDRO testing varying concentrations of fiveGSK3i for 24 hours at chr22 domains.

(g) Representative DNA HiDRO images of chr22 domains after 24 hour GSK3A inhibition by CHIR99021 (Data in Fig. 3h). Solid white line indicates nuclear edge. Scale bar nucleus, 5 μm; scale bar spots, 1 μm.

(h) – (l) Difference in mean D1 overlap (Treatment – Control) +/- SEM for non-selective GSK3 inhibitors (CHIR99021, Lithium, LY2090314), GSK3A-selective inhibitor

(BRD0705), and GSK3B-selective inhibitor (BRD3731). Four concentrations tested for 24 hours, five biological replicates perconcentration. * P-value < 0.05, ** P-Value < 0.01, two-tailed t-test. P-values left to right: CHIR99021 (0.0239); Lithium (0.0191, 0.0016); LY2090314 (0.0284, 0.0031, 0.0044); BRD0705 (0.0220).
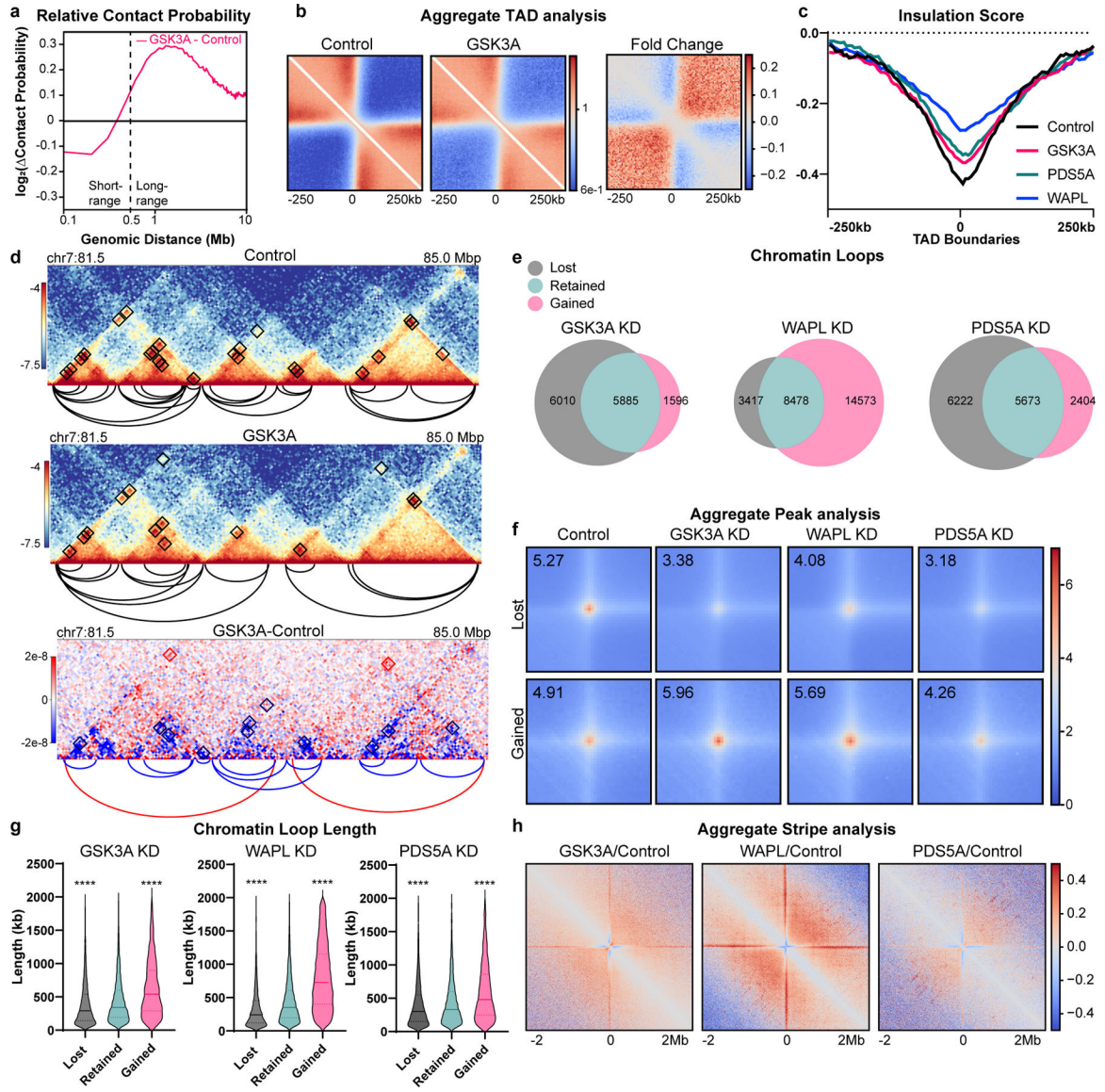
Figure 4: GSK3A restricts chromatin looping to promote TAD insulation.

(a) Log2 of difference (GSK3A KD – control) in contact probability at 100 kb resolution as a function of genomic distance (log scale). Dotted line denotes 500 kb to show phenotypic switch between short-range and long-range interactions.

(b) 3D pileup plots of Hi-C interactions in control and GSK3A KD at control TAD boundaries. Right-most is log2 fold change in interactions across TAD boundaries.

(c) Insulation score pileup at control TAD boundaries in control, GSK3A KD, PDS5A KD and WAPL KD.

(d) Hi-C contact matrices for control, GSK3A KD, and GSK3A – Control at chr7:81.5–85 Mbp. Looping interactions highlighted with black diamonds on contact map and arcs below contact map. For subtraction map, red loops are gained in GSK3A KD, and blue loops are lost.

(e) Number of loops lost, retained, and gained in GSK3A KD, PDS5A KD and WAPL KD vs. Control.

(f) Aggregate peak analysis for Hi-C signal in Control, GSK3A KD, PDS5A KD and WAPL KD at chromatin loops that are gained in GSK3A KD (n = 1,596) or lost in GSK3A KD (n = 6,010).

(g) Violin plots of loop length in kb of different classes of loops. Solid line is median, dotted lines are 25th and 75th percentiles.** P-value < 0.01, **** P-value < 0.0001, two-tailed Mann-Whitney U-test. [Number of loops lost, gained, retained in Fig. 4e].

(h) 3D pileup plot of Control Hi-C, GSK3A KD Hi-C and log2 fold change of GSK3A/Control at architectural stripes detected in control sample.
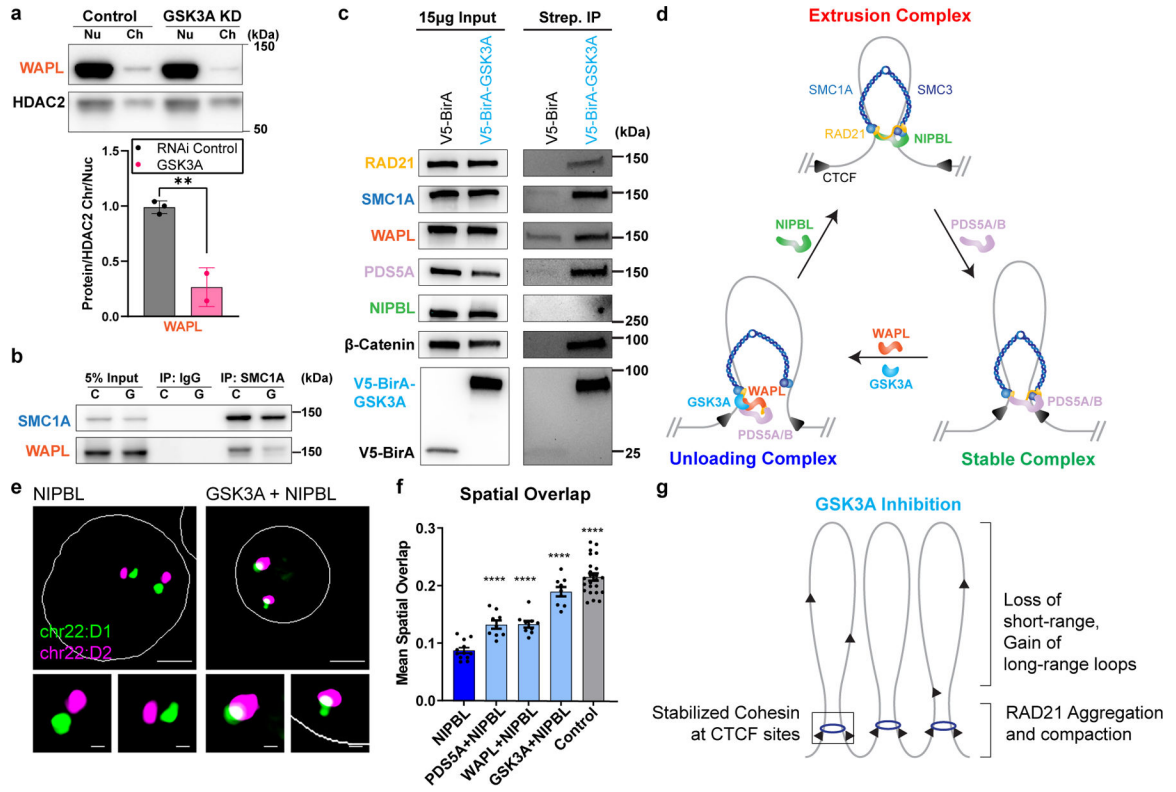
Figure 5: GSK3A promotes WAPL recruitment to chromatin.

(a) Western blot of nucleoplasmic and chromatin-bound fractions of WAPL for control and GSK3A KD. WAPL and HDAC2 run on the same blot. Additional replicates in Extended Data Fig. 6. Quantification is chromatin-bound to nucleoplasmic ratio of WAPL, normalized to loading control (HDAC2). Bars are mean of three biological replicates for control and two biological replicates for GSK3A KD. ** P-value = 0.0058, two-tailed t-test.

(b) Co-immunoprecipitation on chromatin fraction of control and GSK3A KD lysate. C = Control, G = GSK3A KD. Proteins labeled with HRP-linked antibodies. Biological replicates and quantification in Extended Data Fig. 8.

(c) Western blots following Turbo-ID proximity labeling for cohesin components input lysate or lysate from biotin incubation with either a control construct (V5-BirA) or GSK3A-V5-BirA. Proteins labeled with HRP-linked antibodies. Biological replicate in Extended Data Fig. 8.

(d) Model for three states of cohesin including (1) an actively extruding and NIPBL-bound complex, (2) a PDS5-bound looping complex paused at CTCF sites, and (3) a WAPL-bound complex unloading from chromatin. GSK3A facilitates the transition between the stable and unloading complexes.

(e) Representative DNA HiDRO images at chr22:D1 & D2 of NIPBL KD and double KD NIPBL + GSK3A. Solid white line indicates nuclear edge. Scale bar nucleus, 5 μm; Scale bar spots, 1 μm.

(f) Mean D1 overlap +/- SD at chr22 domains for Control, NIPBL and double KD of NIPBL and cohesin negative regulators. **** P-value < 0.0001, two-tailed t-test. [Biological

replicate wells for control (n = 24); NIPBL (n = 12); PDS5A+NIPBL, WAPL+NIPBL, GSK3A+NIPBL (n=9)]

(g) Loss of GSK3A activity leads to overactive cohesin phenotypes including accumulation of cohesin at CTCF sites, RAD21 aggregation on chromatin (vermicelli) and gain of long-range loops at the expense of short-range loops.