Communications Biology

https://doi.org/10.1038/s42003-026-09838-x

Article in Press

Impaired cohesin loading disrupts pancreatic differentiation by Polycomb-driven chromatin rewiring and loop collapse

Received: 28 April 2025

Accepted: 27 February 2026

Cite this article as: Yu, L., Liu, Y., Zhang, J. *et al.* Impaired cohesin loading disrupts pancreatic differentiation by Polycomb-driven chromatin rewiring and loop collapse. *Commun Biol* (2026). https://doi.org/10.1038/s42003-026-09838-x

Longtao Yu, Yayu Liu, Jie Zhang, Chaofan Huang, Huanyu Feng, XueXue Zhao, Shiyu Feng, Wenqian Gong & Xiaowen Lyu

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Impaired Cohesin Loading Disrupts Pancreatic Differentiation by Polycomb-Driven Chromatin Rewiring and Loop Collapse

Longtao Yu[1,2], Yayu Liu[1,2], Jie Zhang[1,2], Chaofan Huang[1,2], Huanyu Feng[1,2], XueXue Zhao[1,2], Shiyu Feng[1,2], Wenqian Gong[1,2], and Xiaowen Lyu[1,2]

[1]State Key Laboratory of Cellular Stress Biology, Fujian Provincial Key Laboratory of Reproductive Health Research, School of Medicine, Faculty of Medicine and Life Sciences, Xiamen University, Xiamen 361102, China

[2]Fujian Provincial Key Laboratory of Organ and Tissue Regeneration, School of Medicine, Faculty of Medicine and Life Sciences, Xiamen University, Xiamen 361102, China

Running Title: NIPBL-Dependent Cohesin Dynamics Regulate Enhancer-Promoter Interactions in Pancreatic Differentiation

Corresponding Author:

Xiaowen Lyu

Email: xiaowenlyu@xmu.edu.cn

Abstract

Cell differentiation is a complex process characterized by specific gene expression patterns regulated through enhancer-promoter interactions within the three-dimensional architecture of the nucleus. The precise role of cohesin loading dynamics in restructuring chromatin during pancreatic lineage commitment remains unclear. Here we show that knockdown of cohesin loader NIPBL disrupts enhancer-promoter interactions and CTCF-mediated loops, leading to widespread transcriptional dysregulation. Furthermore, the loss of cohesin-mediated loops is accompanied by increased contacts between Polycomb Repressive Complex (PRC) domains, highlighting the interplay between cohesin dynamics and PRC-mediated compartmentalization. Although RAD21 and SA1 cohesin levels remain stable at CTCF loop anchors, NIPBL is essential for maintaining long-range chromatin interactions at later differentiation stages. These findings establish cohesin loading as a critical regulator of 3D genome reorganization during cell fate determination, providing a mechanistic framework for understanding cohesinopathy-related developmental disorders.

KEY WORDS: Loop extrusion; Cohesin; 3D Organization; Chromatin; Transcription; Diabetes

Introduction

Cell differentiation involves the establishment of cell-specific gene expression patterns through changes in enhancer-promoter (E-P) interactions, which exist in the three-dimensional (3D) nuclear space [1-5]. Enhancers and promoters can form compartments, which can be identified in Hi-C heatmaps and may correspond to membraneless organelles [6-11]. Either transcription activation or repression complexes can form macro molecular condensates that bring distal elements together, resulting in activation and repression of genes, respectively [7,12-15]. The establishment of compartments usually competes with another genome organizing process, i.e. the continuous cohesin extrusion taking place in the nucleus [2,14,16-20]. Cohesin loads on and continually extrudes the chromatin fiber to form loops [21-24]. Although loop anchors are highly enriched in cohesin, they are not necessarily the location where cohesin is loaded [25-32]. Cohesin can be loaded by either topologically embracing the DNA randomly or by its known loading factor, NIPBL, to specific regions such as enhancers [23,30,33,34]. In addition to its role in loading cohesin, NIPBL is also implicated to be required for efficient loop extrusion [35-38].

The cohesin extrusion process is stopped by CTCF-anchored loops, stabilizing E-P interactions within the loops but destabilizing those interactions across loops anchors, forming insulated neighborhoods [39-43]. This relationship between transcription and chromatin 3D architecture has been investigated in several organisms and cell types by depleting NIPBL, WAPL, or other cohesin subunits [27,41,43-51]. In general, depletion of these proteins leads to global loss or redistribution of cohesin loops, while impacting compartments to different degrees. For example, depleting NIPBL in mouse liver has minor effects on A/B compartments, whereas depleting RAD21 results in increased compartmental interactions between super enhancers (SEs) [41,50,52-55]. Despite these structural changes, transcription levels of most genes stay constant in cells depleted of cohesin components[24,56-58]. However, mutations in genes encoding cohesin proteins or CTCF result in complex adult phenotypes resulting from alteration of various developmental processes [59,60]. This prompted us to further investigate the mechanisms by which the loading and chromatin extrusion processes of cohesin contribute to transcriptional regulation in the context of cell differentiation.

Chromatin architecture undergo a complex remodeling during the differentiation of human embryonic stem cells (hESCs) into pancreatic cells. The differentiation of pancreatic cells requires substantial epigenetic remodeling, including the loss of H3K9me3, 5mC and H3K27me3, and gain of H3K4me3 at pancreas-specific enhancers [61-66]. These changes in transcription are accompanied by changes in 3D chromatin organization [67]. Additionally, our recent work suggests that the differentiation process also involves dynamic changes in cohesin mediated loops anchored at CTCF sites, enhancers and promoters [67].

Here we explore the contribution of cohesin extrusion to pancreatic cell differentiation by knocking down the cohesin loading factor NIPBL in hESCs and their in vitro differentiated pancreatic lineages. We systematically compare structural changes caused directly by defects in cohesin loading and/or extrusion versus compartmentalization of active and repressive chromatin during the differentiation process. The results provide insights into how cohesin extrusion and the formation or disassembly of CTCF loops regulate E-P interactions required for differentiation of pancreatic cells.

Results

Redistribution of CTCF loops and enhancer-promoter interactions in hESCs following NIPBL knockdown

To understand how cohesin extrusion contributes to distinct types of chromatin interactions, we knocked down the cohesin loading/processive factor NIPBL, required for extrusion of chromatin by cohesin[23,24,32-35,37,38], in hESCs. NIPBL protein are efficiently depleted in si-NIPBL cells (Supplementary Figs. 1a, 2e and 9). Quantification of cell growth curves and cell cycle phase distribution (G1/S/G2) confirmed that NIPBL depletion causes no interference to hESCs expansion or cell cycle progression (Supplementary Figs. 2f and 2g). To analyze changes in cohesin occupancy upon NIPBL loss, we performed ChIP-seq against NIPBL and RAD21. We find a global depletion of NIPBL peaks and a loss of 9,577 RAD21 peaks in si-NIPBL hESCs at RAD21 sites in Ctrl cells (Fig. 1a, Supplementary Figs. 1b,1c and 1d). To examine whether the loss of RAD21 is site-specific and sensitive to NIPBL depletion, we separated RAD21 sites into four groups based on their properties, including promoters, enhancers, and CTCF sites. Although NIPBL is highly enriched at promoters (defined as TSS ±1 kb), present at enhancers at intermediate levels (see Methods for definition), and modestly enriched at CTCF sites (Fig. 1b), RAD21 is less enriched at promoters and enhancers than at CTCF sites. Importantly, the expression levels of CTCF and RAD21 were not affected in si-NIPBL(Supplementary Fig. 1e). In contrast, the gained RAD21 peaks were relatively fewer in number and were characterized by lower baseline levels of H3K27ac and reduced NIPBL binding (Supplementary Figs. 1f and 1g).

To analyze changes in 3D organization caused by depletion of NIPBL and loss of the cohesin complex from the genome, we performed *in situ* Hi-C in si-NIPBL hESCs and obtained 198,256,133 valid contacts from multiple replicates after quality filtering (Supplementary Data 1). We then analyzed the Hi-C data to identify changes in CTCF loops and E-P loops. We identified a total of 27403 loops from Ctrl and si-NIPBL hESCs combined. Of these, 21,334 loops persist unchanged and 6,069 are dynamic between conditions. An example of a series of nested loops that increase, decrease and sustain in strength after knockdown (KD) is shown (Fig. 1c). Among these nested loops, the largest one formed between a pair of convergent CTCF motifs shows relatively stable levels of RAD21 and SA1, but decreases in frequency of interactions when NIPBL is depleted (Fig. 1c). Multiple NIPBL peaks embedded within this loop, also containing H3K27ac, are almost depleted of NIPBL after KD, suggesting that failure in forming CTCF loops even when the cohesin remains unchanged at the anchors. This result is consistent with the findings of Alonso-Gil et.al that SA1 cohesin can bind chromatin independently of NIPBL and impaired extrusion [38]. Similarly, a smaller loop between a pair of tandem CTCF sites is also lost (Fig. 1c). WAPL occupancy only decreases slightly at the anchors of the large loop but drop considerably at the anchors of the small one (Supplementary Fig. 1h). Since WAPL interacts directly with SA2 instead of SA1 [38], we infer from the loss of WAPL that the small loop is dependent on SA2-cohesin to recruit WAPL. To further investigate the differential reliance of cohesin subunits SA1 and SA2 on NIPBL, we identified genomic sites where cohesin occupancy is most dependent on NIPBL-mediated loading. We reasoned that sites with the highest baseline occupancy of SA2 subunit might represent NIPBL-dependent loops. Therefore, we ranked all TSS, enhancer, and CTCF loop sites by their SA2 ChIP-seq signals. At top-ranked SA2 binding sites, the putative high-loading-demand enhancers and CTCF loops, both SA1 and SA2 levels were significantly reduced upon NIPBL KD (Fig. 1b, right panels). When we applied SA2-based ranking to stable RAD21 peaks—including the most outside pair of CTCF loop anchors shown (Fig. 1c)—we observed that SA2 levels at the highest-ranked CTCF-Loop (CL) anchors were markedly diminished upon NIPBL depletion, while SA1 levels remained comparatively stable (Supplementary Fig. 8). These findings collectively define a coherent subclass of anchors where cohesin stability, particularly of the SA2 subunit, is dependent on NIPBL.

Using meta-analysis for loops N, we identified 3,985 Ctrl-specific and 2,084 si-NIPBL-specific loops (Fig. 1d). By comparing the APA scores (see Methods), we noticed that the strength of Ctrl-specific loops is much higher than that of stable loops and si-NIPBL loops (Fig. 1d). Loops identified by HICCUPS also exhibit the same trend (Supplementary Fig. 1i). These results suggest that loss of NIPBL present at promoters and enhancers can lead to global loss of strong CTCF loops.

To further distinguish the direct regulation of loops by NIPBL, we separated loops anchoring sites located at enhancers and promoters from loops anchoring other sites. Assuming that promoters and enhancers that have the highest NIPBL in Ctrl also show the greatest loss of NIPBL, we would expect that enhancer/promoter (E/P) loops should exhibit the most pronounced decrease. Interestingly, although E/P loops indeed show maximum decreases, half- and non-E/P loops are also pronouncedly reduced, suggesting that indirect regulation of non-E/P loops by NIPBL instead (Fig. 1e). These observations suggest that redistribution of cohesin by depleting NIPBL can impair both E/P and non-EP loops through comprehensive mechanisms in hESCs.

NIPBL knockdown reduces promoter and enhancer interactions without depleting local cohesin

Cell fate transitions during hESC differentiation require transcriptional regulation by distal sequences, which involves the establishment of chromatin architecture such as loops and compartments. These contacts can be highly dynamic between differentiation stages as found previously [3,56,57,62,67], which suggests that cohesin participates in the formation of dynamic loops of E/P sites controlling cell fates. To estimate whether NIPBL is required for the assembly of new E/P loops during cell differentiation, we utilized our published data obtained from different stages during pancreatic differentiation to divide all loops in hESCs into dynamic and static classes (see Methods). We observed that loops that are specific to hESC but downregulated during differentiation decrease much more in si-NIPBL(Supplementary Fig. 2a), suggesting that loops associated with cell identity are more vulnerable to loss of NIPBL.

One possible explanation for this observation is that these loops overlap tissue-specific promoters and enhancers and have more CTCF. Depending on the occupancy of CTCF at both/either/neither anchors, we divided all E/P loops into three classes and compared their changes in Ctrl and si-NIPBL (Fig. 1a). Among them, loops occupied by CTCF at both anchors are strongest in Ctrl and dramatically downregulated in si-NIPBL (Fig. 2a, Supplementary Fig. 2b). Similarly, loops not occupied by CTCF are less pronouncedly downregulated. This suggests that the amount of CTCF at E/P loop anchors correlates with their interaction frequencies, while not preventing their interaction loss after NIPBL KD. To further investigate whether loss of NIPBL impairs these loops through disrupting the loading of cohesin, we then analyzed the changes in the amount of cohesin subunits at their anchors. Interestingly, the amount of RAD21 and SA1 stays constantly high at CTCF anchors (Fig. 2b). The same tendency is observed for both subunits at non-CTCF anchors, despite both subunits being in lower amount. These results suggest that loss of E/P loops is not due to less SA1 cohesin loading, and higher occupancy of CTCF is coupled with the strength of loops only when NIPBL is functional. We also analyzed the changes of H3K27ac, normally highly enriched at enhancer and promoter sites, at the CTCF and non-CTCF anchors (Fig. 2b). Levels of H3K27ac at both anchors seem similarly high in Ctrl but drop slightly at CTCF and stays high at non-CTCF anchors, implying that CTCF coupled E/P loops rely more on NIPBL to maintain their transcriptionally active epigenetic state.

Usually, the extrusion process can be visualized as stripes in interaction matrix heatmaps, as observed in RAD21 HiChIP [30,68]. To explore the possibility that the loss of E/P loops is mainly caused by inefficient extrusion of chromatin by impaired loading and function of cohesin, we performed RAD21 HiChIP to capture signals of stripes toward E/P loops in interaction heatmaps in both Ctrl and KD cells. If the extrusion process but not the loading of cohesin is affected, then loops with stable levels of cohesin at their anchors should have reduced interaction frequencies, while having weaker the stripes in both Hi-C and RAD21 HiChIP data in si-NIPBL. By overlapping RAD21 peaks with anchors of E/P loops captured by RAD21 HiChIP, we separated E/P loops into multiple classes based on the changes of RAD21 levels, i.e. up, down and stable, and analyzed changes in loops and stripes after KD (Fig. 2c). As shown in the metaplots of interaction heatmaps, the 82 loops formed between promoters containing stable RAD21 show a mild decrease in interactions in Hi-C and RAD21 HiChIP (Fig. 2c). Besides, the stripes are also slightly weaker. The 993 enhancer-enhancer (E-E) and E-P loops that retained stable RAD21 levels also showed weakening, except that they exhibited a more substantial reduction in interaction frequencies, even though they were initially strong in Ctrl (Fig. 2d, Supplementary Fig. 2c). These correlative results suggest that loss of NIPBL can affect loops formed between enhancers and promoters by interrupting the extrusion process mediated by cohesin, while maintaining cohesin levels at loop anchors.

To further explore the effect of loss of E/P loops on the local transcription and epigenetic activities, we took these loops containing constant levels of RAD21 and analyzed the changes of H3K27ac and RNF2 at their anchors in si-NIPBL. As shown by meta-analysis of ChIP-seq data, H3K27ac signals are more enriched at anchors than inside of E-E loops, but they are randomly distributed across the whole promoter-promoter (P-P) loops (Fig. 2e). Moreover, this modification goes down specifically at anchors but not inside of enhancer loops (Fig. 2e). The opposite pattern is observed for RNF2, one of the core subunits of the Polycomb Repressive complex 1 (PRC1), in that it is more enriched at anchors than inside of promoter loops and is resistant to NIPBL KD (Supplementary Fig. 2d). These results suggest a loss of epigenetic activities of enhancers neighboring or located at loop anchors which are stably occupied by cohesin when knocking down NIPBL. Since previous studies suggest that the transcriptional machinery, TFs and epigenetic modification enzymes, can also mediate E/P interactions [69-71], we further tested whether the proper loading and extrusion of cohesin are necessary for establishing E/P loops. By performing HiChIP for RNAPII and H3K4me1, which has been utilized to capture E/P interactions, in Ctrl and si-NIPBL hESCs, we find that changes of these two HiChIP in si-NIPBL recapitulated those observed in Hi-C and RAD21 HiChIP. Interestingly, these promoter loops have fewer interactions and weaker stripes as analyzed by RNAPII HiChIP as well (Fig. 2f). Similarly, weaker enhancer loops and their associated stripes were also observed in H3K4me1 HiChIP (Fig. 2g). These observations suggest that NIPBL is necessary for establishing E/P interactions even when these E/Ps contain high levels of the transcriptional machinery.

Lost CTCF loops maintain cohesin at anchors despite the absence of NIPBL

To further explore the specific chromatin interactions that are sensitive to the perturbation of cohesin loading and loop extrusion caused by NIPBL KD, we separated loops into five classes by their sizes and analyzed their response to NIPBL KD. As shown by Hi-C data (Fig. 3a), larger loops have higher interaction frequencies than smaller ones in Ctrl, likely due to stabilizing factors like cohesin and CTCF. Upon NIPBL KD, however, loops of all distances decrease, with the largest class (>400 kb) showing the most pronounced reduction. This observation indicates that larger loops are more sensitive to the perturbation caused by NIPBL KD.

Besides E/P loops, there are still many structural loops formed between CTCF sites. Since NIPBL levels at these loop anchors are very low as described above, we asked whether NIPBL loss leads to decrease of these loops without reducing CTCF or cohesin occupancy at their anchors as well. To this end, we compared the changes in interaction frequencies in loops with stable versus down-regulated RAD21 after KD. Although Hi-C signals are stronger at the 5,763 loops that maintain RAD21 levels at their anchors in si-NIPBL, compared to the 172 loops showing reduced RAD21, both loops decreased in Hi-C signals after NIPBL KD (Fig. 3b, Supplementary Fig. 3a). This indicates that weak loops lose cohesin loading upon NIPBL KD and require NIPBL to maintain interactions, while strong loops retain cohesin but still depend on NIPBL for interaction stability. Additionally, we can infer from metaplots that loops with reduced cohesin are usually smaller than loops with stable cohesin (Fig. 3b). The maintenance of CTCF, RAD21 and SA1 together with loss of NIPBL at anchors of these two loops is also shown by meta-analysis of ChIP-seq signals (Fig. 3c). We also examined loops with up-regulated RAD21, which likewise showed a decrease in interaction frequency, albeit to a lesser extent than the previous two classes (Supplementary Fig. 3b). The up-regulated RAD21 peaks are characterized by low H3K27ac but enriched H3K27me3 (Supplementary Figs. 1g and 3c), suggesting NIPBL depletion permits aberrant cohesin accumulation at Polycomb-repressed regions. The finding that loops with up-regulated, stable, and down-regulated RAD21 all exhibit decreased interaction frequencies demonstrates that NIPBL-mediated cohesin loading is essential for maintaining chromatin architecture, regardless of cohesin chromatin occupancy.

To further explore the effects of NIPBL on the loop extrusion process in hESC, we mapped RAD21 HiChIP signals on the loops that contain stable levels of RAD21 at their anchors and analyzed the changes between conditions. Interestingly, despite strong enrichment of RAD21 HiChIP signals at these loops, stripe-to-corner peak ratios are markedly higher at CTCF loops versus E/P loops (Fig. 3d). Besides, corner peaks and stripes attenuate similarly after NIPBL loss in RAD21 HiChIP. In contrast, CTCF HiChIP shows strong corner peaks but faint stripes, and loops become nearly undetectable upon NIPBL KD (Fig. 3d). All the results above suggest that attenuation of loop extrusion by depleting NIPBL results in loss of CTCF loops without affecting cohesin occupancy.

Loss of cohesin-mediated loops and gain of compartmental contacts within PRC domains in the absence of NIPBL

Based on our observations that promoter interactions containing constant levels of RAD21 are highly enriched with RNF2, the core subunit of PRC1, we further explore the relationship between the cohesin loop extrusion process and PRC repressive compartments in the regulation of pluripotency and cell fate conversion. By setting a threshold in RNF2 ChIP-seq signals (see Methods for detail), we find 363 PRC broad domains (PDs) in ctrl overlapping 764 RAD21 peaks. Of these peaks, 146 decreased (Supplementary Fig. 4a), and 414 remain stable after NIPBL KD, with 26, 48, 244 and 96 overlapping promoters, enhancers, CTCF sites at loop anchors and CTCF sites not at loop anchors, respectively (Fig. 4a, Supplementary Figs. 3f and 4f). Meta-analysis of the ChIP-seq signals of several related proteins at these RAD21 peaks are shown separately for peaks with stable RAD21 (Fig. 4b, Supplementary Fig. 3e) and peaks with lost RAD21 (Supplementary Figs. 4a and 4b). Interestingly, slightly more CTCF and SA2 are observed at peaks overlapping promoters, loop anchors and other CTCF sites, but lower SA1, WAPL and H3K27ac are found at these peaks. This suggests that SA1 cohesin slightly comes off their targets within PDs after NIPBL KD. Meanwhile, SA2 cohesin mildly increases with concomitant loss of WAPL.. This balance between SA1 and SA2 cohesin implies that, distinct from other regions of the genome, loading of SA1 cohesin within PDs is more dependent on NIPBL. Moreover, the defects of SA2 cohesin loop extrusion are further validated by RAD21Hi-ChIP signals showing the loss of stripes after NIPBL KD (Supplementary Fig. 4e). Therefore, despite more loading of SA2 cohesin, its accumulation within PDs cannot maintain loop extrusion without NIPBL.

One possible outcome of the reduction of both SA1 cohesin binding and loop extrusion process is the upregulation of compartmental contacts between PDs as reported previously in mouse embryonic stem cells (mESCs) [36,42]. We test this possibility in hESCs by analyzing changes in Hi-C signals after NIPBL

KD. Typically, the PD containing the *HOXD* gene cluster increases its interactions with two adjacent PDs after NIPBL KD (Fig. 4c), which are also observed in Virtual 4C signals converted from Hi-C data by taking the *HOXD* gene cluster as a viewpoint. Although the occupancy of RAD21 and CTCF is relatively constant at these PDs after NIPBL KD, cohesin mediated loops, which form between anchors within the *HOXD* cluster and neighboring regions known as the C-Dom and T-Dom reported [72], are almost gone after NIPBL KD (Fig. 4c).

To further investigate the nature and promoting factors of PD contacts in hESCs, we treated si-NIPBL cells with 1,6-Hexanediol (1,6-HD), an aliphatic alcohol commonly used to dissolve liquid-liquid phase-separated condensates[13]. 1,6-HD treatment reduces the enhancement of PD contacts following NIPBL KD (Supplementary Fig. 5a). We then questioned whether the enhancement of PD contacts following NIPBL KD was due to the loss of cohesin. To test this possibility, we knocked down WAPL, the cohesin unloading factor, in si-NIPBL hESCs to maintain the cohesin occupancy. This double KD reverted the PD contacts enhancement observed in NIPBL KD alone(Supplementary Fig. 5a).

Since previous studies suggest that CTCF binding sites can load SA1 cohesin topologically without NIPBL[38], we investigated the function of CTCF in promoting PD contacts by CTCF an NIPBL double KD. Surprisingly, this double KD also diminished enhanced PD contacts (Fig. 4f). A similar result was observed when we treated the si-NIPBL hESCs with TPA, a Zn^{2+} chelator that has been reported to inhibit CTCF binding [73] (Supplementary Fig. 5a). Loss of cohesin at CTCF sites may rescue its depletion elsewhere in PDs, while cohesin loss at non-CTCF sites could increase PD contacts.

We also confirm the global enhancement of PD contacts by measuring the changes of Hi-C signals (Fig. 4d). Furthermore, when we treated the NIPBL KD cells with UNC3866, a compound that antagonizes the methyllysine reading function of the chromodomains in CBX, a core subunit of PRC1, the enhancement of PD contacts was reversed (Supplementary Fig. 4c). Notably, our RNF2 ChIP-seq data confirm that this UNC-induced disruption of foci does not substantially reduce PRC1 chromatin occupancy (Supplementary Fig. 3d). This suggests that PD contacts is partially dependent on the reader function of canonical PRC1.

Moreover, since the balance of SA1-cohesin and SA2-cohesin is maintained within PDs, the general levels of RAD21 are relatively stable but slightly lower at these sites after NIPBL KD (Fig. 4b, Supplementary Fig. 3e). However, RAD21 Hi-ChIP signals slightly increased among PD contacts (Fig 4d). This suggests that local occupancy of cohesin does not hinder compartmental contacts between PDs. By analyzing ChIP-Seq data in si-NIPBL, we find that enhanced PD contacts in hESCs are mediated by stably binding PRC1 complex (Fig. 4a, Supplementary Figs. 3f, 4a and 4b), instead of increased PRC1 complex as reported in mESCs [38].. Additionally, RNF2 Hi-ChIP signals also increase after NIPBL KD (Fig 4d). These results suggest that loss of cohesin mediated loops in PDs facilitate long-range PD contacts without increasing PRC1 complex in hESCs.

To further explore the mechanisms underlying PD interactions, we utilized immunofluorescence to detect PRC1 distribution within the nucleus. Similar to previous findings [74], there are many RNF2 foci within nucleus of Ctrl hESCs. Intriguingly, these foci are higher in intensities and numbers in si-NIPBL (Fig. 4e and Supplementary Data 2), suggesting enhanced condensation of PRC and increased contacts among PDs upon NIPBL KD. When we treated si-NIPBL cells with UNC3866, both the sizes and numbers of these foci reverted to those in Ctrl condition (Fig. 4e). Notably, our RNF2 ChIP-seq data confirm that this UNC-induced disruption of foci does not substantially reduce PRC1 chromatin occupancy (Supplementary Fig. 3d). In summary, these results suggests that PRC1 contributes to PD interactions through condensation, which mainly depends on its reader function. Our results are consistent with findings that treatment of mESCs with 1,6-hexanodial disrupts the formation of PRC1 foci [74].

Next, we asked what sequences within PDs mainly contribute to the increase of their compartmental contacts. By dividing PD interactions into multiple classes, specifically those anchoring promoters, enhancers and CTCF sites, we find that E/P loops increase more dramatically than CTCF loops after NIPBL KD(Fig. 4f). However, additional treatment with UNC3866 has minor effects on increased interactions (Supplementary Fig. 4d). These changes suggest that genes and enhancers within poised domains behave differently than those in the rest of the genome and can form more compacted PRC compartments when cohesin mediated loop extrusion is inhibited by loss of NIPBL. In contrast, CTCF interactions within PDs are more resistant to NIPBL KD than those in the rest of the genome.

To further elucidate the mechanistic basis of enhanced inter-Polycomb domain (PD) contacts following NIPBL depletion, we performed Hi-C experiments in dual KD systems combining NIPBL with WAPL or NIPBL with CTCF (Supplementary Fig. 5b). We find that NIPBL+WAPL co-depletion disrupts the enhanced PD contacts, indicating this phenotype depends on losing functional cohesin loading and subsequent loop domain disintegration rather than representing a direct gain of function. Similarly, NIPBL+CTCF co-depletion fails to further enhance PD contacts and may even suppress them. Meta-analysis further revealed that interactions between promoters (P), enhancers (E), and CTCF sites at loop anchors (CL) within PDs are similarly affected by these dual KDs (Supplementary Fig. 5b). This suggests cohesin loss at non-CTCF sites serves as the primary driver, while residual cohesin at CTCF sites helps establish the genomic context that permits increased PD interactions. Together, these results underscore that the enhancement of PD contacts is not a simple consequence of phase separation but arises from the complex interplay between distinct cohesin loading pathways and the subsequent large-scale reorganization of the chromatin landscape.

Loss of cohesin-mediated loops within super enhancer domains in the absence of NIPBL

Previous studies in somatic cells found that depleting various cohesin subunits have little impact on 3D interactions except for an increase in super enhancer (SE) interactions [41,48-51]. We therefore explored the effects of NIPBL KD on SEs in hESCs. We find 707 down-regulated RAD21 peaks at SEs (Fig. 5a). Levels of H3K27ac remain constant at these peaks . Besides, although SA2 is more abundant than SA1 at these peaks (Fig. 5b, Supplementary Fig. 5c), it decreases after NIPBL KD (Fig. 5b, Supplementary Fig. 5c), indicating that loading of SA2-cohesin cannot be replaced by SA1-cohesin at these sites. Notably, we still found 1,709 stable RAD21 peaks within super-enhancers (Supplementary Fig. 5d).

Next, we utilized Hi-C data to analyze changes of interactions between SEs after NIPBL KD. For example, around the *NANOG* gene locus, increase in domains (SDs) interactions are induced by NIPBL KD (Fig. 5c), which is prevented by concomitant 1,6-HD treatment (Supplementary Fig. 5e). Similar results were observed when we knocked down WAPL or CTCF, either individually or alongside NIPBL, or when treating NIPBL KD cells with TPA (Supplementary Fig. 5e). . Nonetheless, distinct from contacts between PDs, only half of the contacts between SDs are increased upon NIPBL KD (Fig. 5d and Supplementary Data 3). Specifically, upregulation of SD contacts usually occur between SDs that are more distant from each other, as shown in meta-plots (Fig. 5d).. This suggests that some long-range SD contacts can be induced by NIPBL KD in hESCs.

Enhanced distant SD interactions could dilute adjacent cohesin-mediated loops. We therefore examined cohesin-dependent SIP loops anchored at SDs using Hi-C.We find 998 SIP loops which overlap SDs on both sides and 69 loops which overlap SD on either side. As expected, loops overlapping SDs on both sides decrease more after NIPBL KD (Fig. 5e and Supplementary Data 3), suggesting that original cohesin mediated SD loops are impaired by increased SD interactions induced by NIPBL KD.

We then asked whether interactions of SDs and their target gene promoters correlate with the occupancy of cohesin at SDs. To this end, we divided SD loops into four classes by the changes of RAD21 at loop anchors and compared these loops in both sizes and strengths upon NIPBL KD. Intriguingly, SD loops containing stable and down-regulated levels of RAD21 upon KD are intermediate in sizes, while those containing no RAD21 are biggest (Fig. 5f). Besides, SD loops with more cohesin at anchors upon NIPBL KD have even less interaction frequencies with their target gene promoters than those containing less or constant cohesin (Fig. 5f and Supplementary Data 3). These results suggest that small SE/P loops are more dependent on cohesin than large ones. This spatial reorganization implies mechanistic compensation via cohesin redistribution, which differentially affects local and long-range contacts within SE domains.

Dissolution of interactions between enhancers and down-regulated genes within regions of shortened CTCF loops

Depletion of cohesin subunits usually has mild effects on gene transcription [41]， but the underlying mechanisms are still unclear. To obtain further insights into the effects of the loop extrusion process on transcription or vice versa, we performed RNA-seq experiments to get differentially expressed genes (DEGs) in Ctrl and NIPBL KD. Additionally, since treatment with UNC3866 can reverse enhancement of PD contacts induced by NIPBL KD, it may affect transcription as well. Therefore, we also include this condition for comparison. Notably, control samples clustered distinctly from NIPBL KD samples in the PCA analysis, indicating significant alterations in transcription (Supplementary Fig. 6a). Based on the high reproducibility of gene expression changes (Supplementary Fig. 6a), we find 1067 up-regulated and 708 -down-regulated genes (Fig. 6a, Supplementary Fig. 6b), which is close to the number of DEGs published previously. Furthermore, 502 -down-regulated and 393 up-regulated genes are common between si-NIPBL and si-NIBPL treated with UNC3866 (Fig. 6a), suggesting that these genes are regulated by cohesin loops but not poised by PRC1 repression. To get a more comprehensive view of transcriptional changes, we performed Gene Ontology (GO) analysis on the DEGss. GO analysis of DEGs revealed enrichment in embryonic development and differentiation terms (Supplementary Fig. 6d, 6f). DEGs were also enriched at anchors of down-regulated loops (Supplementary Fig. 6e), linking transcriptional changes to lost E-P interactions. Comparison of lineage markers showed that NIPBL KD mildly reduced pluripotency markers (*NANOG*, *POU5F1*; Fig. 6b) and mesoderm markers, while slightly elevating ectoderm/endoderm markers. *SA1* was up-regulated among cohesin subunits (Supplementary Fig. 6c). These shifts suggest that disrupting cohesin loading and loop extrusion erodes pluripotency and biases cells toward differentiation.

To explore the relationship between the loss of CTCF loops and inactivation of enhancers and promoters responsible for changes in transcription, we examined Hi-C data from Ctrl and si-NIPBL cells . Notably, loss of CTCF loops in si-NIPBL with respect to Ctrl cells results in decreased E-P interactions within the loops (Fig. 6c), suggesting that CTCF loops allow more frequent internal E/P interactions.

Since Ctrl-specific loops could maintain CTCF at their anchors (Fig. 3c,3d), in principle, new loops in si-NIPBL may form between proximal CTCF sites via premature extrusion arrest, while lost Ctrl-specific loops likely require efficient extrusion to link distal sites. Size analysis confirms that lost Ctrl loops are longer than stable loops, whereas gained si-NIPBL loops are shorter (Fig. 6d and Supplementary Data 4). These results suggest that new loops after NIPBL KD may form by shrinking, as if cohesin stops at nearer CTCF sites before reaching original Ctrl anchors (Fig. 6e).

To test whether si-NIPBL-specific CTCF loops form by shrinking existing Ctrl loops, we analyzed differential loops relative to CTCF site occupancy. Most lost loops contain gained loops internally. Approximately 60% maintain one CTCF anchor but now pair with a nearer CTCF site, while 20-30% dissolve into shorter internal loops (Fig. 6e). Meta-analysis of Hi-C data confirms that loops encompassing gained loops lose interaction frequency after NIPBL KD (Fig. 6e). This shortening may reflect reduced cohesin loading and impaired extrusion efficiency due to NIPBL KD.

We then examined cohesin loading by mapping NIPBL ChIP-seq across lost Ctrl-specific loops. In the *SOX2* region, CTCF loops disappear upon NIPBL KD, and NIPBL-binding sites between anchors—potential cohesin loading sites marked by H3K27ac—are also lost (Fig. 6f). Genome-wide analysis confirmed that NIPBL-binding sites between anchors of lost Ctrl loops disappear in si-NIPBL (Fig. 6g). These sites are enriched for RAD21 and H3K27ac, supporting their identity as enhancer-/promoter-like cohesin loading regions that are disrupted after NIPBL KD (Fig. 6g).Together, these results demonstrate that loss of cohesin loader NIPBL impairs both E-P interactions and CTCF-loop architecture (Fig. 6e-g), highlighting its critical role in maintaining chromatin structure during differentiation.

NIPBL knockdown during pancreatic differentiation reduces chromatin accessibility and disrupts stage-specific loops To investigate how cohesin loading and loop extrusion contribute to cell-fate establishment during development, we differentiated H9 hESCs into pancreatic islet organoids following established protocols [75,76] and performed NIPBL KD at key intermediate stages, including definitive endoderm (DE), primitive gut tube-like (PGT), pancreatic progenitors (PP), and stem cell-derived β-cell organoids (SC-β organoids) [75-77]. To ensure that observed phenotypes reflected direct chromatin regulatory roles rather than secondary effects on cell viability, we first confirmed efficient *NIPBL* KD at the PP2 stage by RT-qPCR (Supplementary Fig. 7a) and verified that NIPBL depletion did not significantly alter cell survival or apoptosis rates via flow cytometry (Supplementary Figs. 7b and 10).

We then asked whether stage-specific gene expression is affected by NIPBL loss. RNA-seq performed at PP stages revealed 3827 DEGs absent in si-NIPBL induced PP (Fig. 7a). Principal component analysis (PCA) showed clear separation between control and si-NIPBL PP samples (Supplementary Fig. 7c), indicating NIPBL KD substantially alters the transcriptional program associated with pancreatic lineage progression.

Because stage-specific enhancers frequently neighbor CTCF-loop anchors [67], we next asked whether the activation of such enhancers requires NIPBL. ATAC-seq at DE and PGT stages identified 8957 and 4371 ATAC-TF peaks, respectively, that were absent in si-NIPBL induced PGT (Fig. 7b, supplementary Figs. 7d and 7e). These accessibility changes were reproducible across replicates (Supplementary Fig. 7f) and resulted distinct clustering in PCA (Supplementary Figs. 7g and 7h). Transcription factor motif enrichment analysis at stage-specific anchor summits showed that motifs significantly enriched in Ctrl DE and PGT cells were less prominent in si-NIPBL samples (Fig. 7c), suggesting impaired activation of stage-specific enhancers upon NIPBL KD.

To assess Polycomb-specific contributions to these transcriptional changes, we treated NIPBL-KD PP2 cells with the Polycomb Repressive Complex 2 (PRC2) inhibitor UNC3866. The z-score heatmap of si-NIPBL + UNC3866 samples (Supplementary Fig. 7i)revealed that UNC3866 treatment partially restored expression of differentiation-associated genes disrupted by NIPBL depletion, indicating that PRC activity contributes specifically to the transcriptional deficits observed during pancreatic lineage commitment.

Given that new cell-type generation requires establishment of E/P interactions often associated with CTCF-loop formation [67], we performed Hi-C in si-NIPBL-induced PP cells. Meta-analysis showed that 1369 loops gained at PP stage relative to hESCs exhibited markedly lower increase in interaction frequencies of Hi-C data in si-NIPBL cells(Fig. 7d). These differences elicit us to investigate whether the extension of CTCF loops found in our previous work is generally affected at PP stage by NIPBL KD. By measuring the sizes of loops gained in ctrl and si-NIPBL induced PP cells and those of stable loops, we find that PP gained loops, which are longer than stable loops, are more abundant in ctrl than in si-NIPBL induced cells (Fig. 7e and Supplementary Data 4). Extension of CTCF loops, where either one or both anchors shift distally, was significantly reduced upon NIPBL KD. Particularly, while ~64.2% of new PP loops in control cells maintained one anchor but paired with a more distal CTCF site, and 27.2% formed between pre-existing and new anchors to create two longer loops, both proportions decreased in si-NIPBL cells (Fig. 7e, Supplementary Fig. 7j). We then asked whether E-P interactions within these CTCF loops are likewise downregulated. Mapping Hi-C signals to PP-specific CTCF loops that encompass upregulated genes and their enhancers revealed higher median signals at both CTCF loops and internal E-P loops in control relative to si-NIPBL cells (Fig. 7f). Subtraction of hESCs signals from PP-stage data also showed a greater increase in interaction frequencies at these loops in control cells (Fig. 7f), indicating that NIPBL KD reduces both CTCF-loop extension and the E-P interactions they enclose. Median and subtraction heatmaps further confirmed that interaction frequencies at extended CTCF loops were lower in si-NIPBL than in control PP cells (Fig. 7g).

Finally, we tested whether PRC inhibition could rescue these architectural defects. Meta-analysis of the same category of upregulated E-P loops within CTCF loops showed that UNC3866 treatment altered interaction patterns in both control and si-NIPBL PP2 cells but did not restore loop-level architecture in si-NIPBL cells (Supplementary Fig. 7k). Thus, while PRC inhibition can partially reverse compartment-level changes in inter-PD contacts (Supplementary Figs. 4c and 4d), it cannot compensate for the loss of specific cohesin-dependent loops resulting from impaired loading

In summary, NIPBL is required for stage-specific enhancer activation, extension of CTCF loops, and stabilization of enhancer-promoter interactions during pancreatic differentiation. These findings underscore that cohesin loading and the loop extrusion it mediates are critical for orchestrating the 3D chromatin reorganization that underlies cell-fate transitions.

Discussion

Cell differentiation is a highly orchestrated process that requires precise regulation of gene expression, primarily mediated by E-P interactions [3]. Our study elucidates the pivotal role of the cohesin loading factor NIPBL in maintaining the integrity of these interactions through its influence on chromatin architecture. Specifically, we demonstrate how NIPBL KD disrupts cohesin dynamics, leading to alterations in CTCF loops and E-P connectivity, which are crucial for the differentiation of hESCs into pancreatic lineages.

Our findings corroborate existing literature suggesting that loss of NIPBL hampers the loading of cohesin at crucial regulatory regions [30]. We observed a global depletion of NIPBL and RAD21, implicating an immediate reduction in cohesin occupancy. While RAD21 levels remained relatively stable at CTCF loop anchors, the overall cohesin dynamics shifted, suggesting a redistribution that is crucial for maintaining chromatin loop structures. Notably, our results indicate that the reduction of NIPBL leads to a greater loss of E-P interactions compared to CTCF loops, underscoring the importance of NIPBL and cohesin in mediating dynamic E-P looping.

CTCF serves as both a loop anchor and a stabilizing factor for E/P interactions. The observed decrease in both loop types following NIPBL KD indicates that cohesin-mediated loop extrusion is vital for maintaining these interactions. Interestingly, some loops retained RAD21 occupancy but showed diminished interaction frequencies alongside a loss of original CTCF loops. This highlights a potential mechanism whereby pre-existing loops fail to sustain interaction strength in the absence of effective NIPBL-mediated cohesin loading.

Parallel to the effects, our study reveals insights into PRC domain behavior. Previous findings have shown that cohesin depletion can upregulate long-range interactions between PRC target genes [42]. Here, NIPBL KD affected both SA1 and SA2 cohesin, indicating that PRC domains are influenced by cohesin-modulated chromatin architecture. Following NIPBL KD, we observed increased in E-P interactions within PRC domains. This upregulation of compartmental contacts supports the notion that when loop extrusion is disrupted, alternative mechanisms may stabilize interactions among Polycomb-associated genes. Since PRC domains typically maintain gene silencing, loosening these constraints could inappropriately activate developmental genes, potentially disrupting cell identity or facilitating differentiation, emphasizing the delicate balance required between architectural proteins like cohesin and regulatory complexes like Polycomb.

Super enhancers (SEs), known for driving high transcriptional output, exhibit complex interactions with cohesin and NIPBL. We observed a reduction in RAD21 peaks at SEs following NIPBL KD, indicating that cohesin dynamics are crucial for maintaining SE structural integrity. While H3K27ac levels remained constant, the loss of RAD21 suggests that cohesin-mediated loop structures essential for the stabilization SE-promoter interactions are compromised, implying that cohesive loading at SEs must be maintained during differentiation to ensure the proper regulation of cell fate-determining genes.

Our analysis extends to pancreatic progenitor differentiation, where NIPBL depletion leads to impaired activation of stage-specific enhancers. The transcriptional changes observed, including the downregulation of pluripotency markers and altered expression of germ layer markers, suggest a shift toward differentiation that may be exacerbated by the loss of E-P interactions. Given that these activation patterns correlate with the formation of CTCF loops, it follows that the spatial and functional reorganization of chromatin is integral to facilitating cell fate determination.

In conclusion, this study provides valuable insights into the intricate mechanisms by which cohesin dynamics, mediated by the loading factor NIPBL, influence chromatin organization and gene regulation during cellular differentiation, highlighting the multifaceted relationships between chromatin architecture, transcriptional regulation, and cellular identity.

MethodsCell lines and culture

H9 cells (WA09), obtained from WiCell, served as the starting material for pancreatic lineage specification [75]. Routine propagation was performed on Matrigel Matrix-coated (Corning, 354277) surfaces in mTeSR™ Plus medium (STEMCELL Technologies, 100-0276), with cultures maintained at 37°C under 5% CO2 and fed daily. To trigger differentiation, single-cell suspensions were prepared and seeded at 1×10^6 cells/ml in 5.5-6 ml of mTeSR™ Plus containing 10 μM Y27632 (Selleckchem, S1049) per well of 6-well spheroid plates (JetBiofil). Overnight incubation on an Innova 2000 rotator (97 rpm, 37°C, 5% CO2) induced spheroid formation. Fresh mTeSR1 with Y27632 was supplied at 24 and 48 hours post-plating. After 72 hours, spheroids were collected and the following stepwise differentiation protocol was initiated through sequential media changes with specified supplements:

Day 1: S1 basal medium supplemented with 100 ng/ml Activin A (R&D Systems, 338-AC), 3 μM CHIR99021 (Selleckchem, S2924), and 10 μM Y27632.
Day 2: S1 medium containing 100 ng/ml Activin A only.
Days 3, 5, 10, 12, 15, 17, and 19: No medium change.
Days 4 and 6: S2 medium with 50 ng/ml KGF (Peprotech, AF-100-19).
Days 7 and 8: S3 medium supplemented with 50 ng/ml KGF, 0.25 μM Sant1 (Sigma, S4572), 2 μM RA (Sigma, R2625), 200 nM LDN193189 (Sigma, SML0559; added only on Day 7), 500 nM PdBU (EMD Millipore, 524390), and 10 μM Y27632.
Days 9, 11, and 13: S3 medium containing 50 ng/ml KGF, 0.25 μM Sant1, 100 nM RA, and 10 μM Y27632.
Days 14 and 16: S5 medium with 0.25 μM Sant1, 100 nM RA, 1 μM XXI (Sigma, 565790), 20 ng/ml Betacellulin (Peprotech, 100-50), 10 μM Alk5i II (Enzo, ALX-270-445-M001), and 1 μM T3 (Sigma, 64245-250MG-M).
Days 18 and 20: S5 medium containing 20 ng/ml Betacellulin, 1 μM XXI, 10 μM Alk5i II, 1 μM T3, and 25 nM RA.
Days 21 to 35: S6 medium with 10 μM Alk5i II and 1 μM T3; medium was refreshed every other day.
Samples representing definitive endoderm (DE), pancreatic gut tube (PGT), pancreatic progenitor (PP), and SC-β organoid stages were harvested on days 4, 7, 14, and 35, respectively.

Treatment of H9 cells with siRNA and chemical agents

To knock down NIPBL, cultured hESCs were transfected with either a negative control siRNA or siRNA specifically targeting NIPBL (Genepharma) at a final concentration of 50 nM, using RNAiMAX reagent (Thermo) for a duration of 24 hours. The cells were harvested 72 hours post-transfection, and the efficiency of the KD was assessed via real-time polymerase chain reaction (PCR) analysis and Western blotting. For NIPBL KD at the pancreatic progenitor (PP) stage, cultured spheroids were treated with Accutase (Thermo) at 37°C for 5 minutes to dissociate them into single cells, 72 hours before the completion of the PP stage. These cells were then transfected with siRNA at a final concentration of 50 nM, again utilizing RNAiMAX reagent (Thermo) for 24 hours, and incubated overnight on an Innova 2000 rotator at 97 rpm, 37°C under 5% CO2 to allow them to reaggregate into spheroids. The reaggregates were collected 72 hours after transfection, and the KD efficiency was confirmed through real-time PCR analysis.
SiRNA oligo sequences are as follows,

siNIPBL: 5’-CAACAGAUCACAUAGAGUUTT-3’
siWAPL: 5’-GGAGUAUAGUGCUCGGAAUTT-3’
siCTCF: 5’-GGAGAAACGAAGAAGAGUATT-3’

The treatment with UNC3866 was carried out using the following strategy. For normally cultured hESCs, 10 mM UNC3866 (MedChemExpress, HY-100832) is added to the culture medium to reach a working concentration of 40 µM. The medium is changed daily, and cell samples are collected on Day 2 and Day 4 for subsequent immunofluorescence analysis and library construction experiments. For hESCs undergoing NIPBL transfection, 10 mM UNC3866 (MedChemExpress, HY-100832) is added to the culture medium the day before transfection to achieve a final working concentration of 40 µM. After transfection, the cells are maintained in the UNC3866-containing medium, with the medium being replaced daily. Cell samples are collected on Day 3 post-transfection for immunofluorescence analysis and library construction experiments.

The treatment with Tris (2-pyridylmethyl) amine (TPA, Thermo Fisher, Y-24201) was performed according to the following protocol. Two hours prior to sample collection, add TPA to the culture medium to achieve a final concentration of 10 mM. Incubate for an additional 2 hours before harvesting the cells for immunofluorescence and next-generation sequencing (NGS) library preparation.

The 1,6-Hexanediol (1,6-HD, Rhawn, 629-11-8) treatment was implemented according to the following strategy. For both normally cultured and transfected hESCs, add 2% 1,6-HD to the culture medium 5 minutes before sample collection. After this step, proceed to harvest the cells for immunofluorescence and library preparation experiments.

Western analysis

Cells were lysed in Laemmli SDS buffer and heated for 5 minutes. The samples were separated using 6% SDS-PAGE and transferred to a PVDF membrane (Millipore) in Tris-glycine transfer buffer containing 20% methanol for 2 hours at 100 volts. For Western blotting, the membranes were blocked in TBST (20 mM Tris, pH 7.4, 150 mM NaCl, 0.05% Tween 20) with 5% nonfat milk powder and incubated overnight with the following antibodies: RAD21 (1:5000) (Abcam, ab992), NIPBL (1:5000) (Bethyl, A301-779A), TBP (1:6000) (proteintech, 22006-1-AP), RNF2 (1:5000) (BETHYL, A302-869A-A) and CTCF (1:6000) (Abcam, ab70303). Following three washes with TBST, membranes were incubated with HRP-conjugated secondary antibodies (1:5000, Jackson ImmunoResearch Laboratories) for 1 hour. After three additional washes, the detection of various proteins was performed using SuperSignal West Pico/Dura Chemiluminescent substrate (Thermo).

Immunofluorescence microscopy

Immunofluorescence microscopy was conducted as previously described [78], with some modifications. Control and treated cells were first allowed to adhere to slides in growth media for 1 hour. Cells were then fixed in 1% formaldehyde in PBS for 10 minutes and permeabilized with 0.3% Triton X-100 in PBS for 30 minutes at room temperature. Afterward, the cells were blocked with a solution containing 1× PBS, 0.2% Tween-20, and 5% normal donkey serum for 1 hour and incubated overnight at 4°C with a primary antibody diluted in blocking buffer (1:200 rabbit anti-RING1B (BETHYL A302-869A-A)). Slides were washed three times in PBST (PBS + 0.1% Triton X-100) and then incubated for 2 hours at room temperature with goat anti-rabbit Alexa Fluor 488 (Thermo) diluted in blocking buffer (1:1000). After washing three times, the slides were mounted with Vectashield mounting medium containing DAPI (Vector Laboratories). Image acquisition was performed using Zeiss LSM880+Airyscan. Z-stack images spaced by approximately 0.5 µm intervals were captured for a total of around 100-200 cells per condition using ZEN 3.11 software. The area stained with RNF2 was quantified using Fiji Software. Statistical significance for differences in the number and intensities of PRC1 bodies was assessed using the Wilcoxon rank-sum test in R.

RNA-seq

TRIzol^{TM} reagent (Thermo, 15596018CN) was used to isolate total RNA from control and siRNA-treated H9 cells, together with samples collected at various differentiation time points. Ribosomal RNA depletion was carried out using the Ribo Minus Transcriptome Isolation Kit (Invitrogen, K1550). RNA concentrations were determined with the Qubit RNA HS Assay Kit (Thermo Fisher), after which fragmentation buffer was added to randomly break the RNA into smaller pieces. Using these fragments as templates, cDNA was synthesized with random hexamer primers. The resulting double-stranded cDNA underwent end repair, A-tailing, and adapter ligation, followed by size selection and PCR amplification to generate final libraries. Each sample was prepared in two independent biological replicates and sequenced on an Illumina NovaSeq 6000, producing paired-end 150 bp reads.

In-situ Hi-C

Hi-C libraries were constructed in situ with DpnII utilizing established methods[26]. From each stage, 2.5 million cells were crosslinked in 1% formaldehyde, quenched with glycine, washed, and permeabilized to release nuclei. Chromatin was digested overnight with DpnII, and biotin-dCTP with unlabeled dATP/dTTP/dGTP filled the 5' overhangs. Proximity ligation joined fragments, then crosslinks were reversed and DNA ethanol-precipitated. After sonication to 200-500 bp, streptavidin beads captured biotin-labeled DNA. Bead-bound material underwent end repair, A-tailing, and adapter ligation per Illumina TruSeq protocols, with two Tween Washing Buffer rinses between steps. On-bead PCR with barcoded primers and KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) ran for 5-12 cycles. Libraries were paired-end sequenced on an Illumina NovaSeq 6000. At least two biological replicates per condition were pooled after confirming high correlation.

HiChIP

HiChIP experiments were conducted following a modified version of a published protocol [68]. Control and siRNA-treated H9 cells, together with their differentiated derivatives, were collected and crosslinked in 1% formaldehyde for 10 minutes at room temperature. The reaction was quenched with glycine, followed by PBS washes and storage at -80°C. To obtain intact nuclei, thawed cell pellets were resuspended in Hi-C lysis buffer and gently disrupted using pestle A. Isolated nuclei underwent DpnII digestion, biotin-dCTP fill-in, and proximity ligation to join spatially adjacent fragments. Chromatin was then sonicated to 200-500 bp fragments and pre-cleared with Protein A/G Dynabeads for 2 hours at 4°C. Immunoprecipitation was performed overnight at 4°C using beads pre-coated with specific antibodies. On-bead tagmentation with Tn5 transposase (Vazyme, S601) was carried out, followed by elution, crosslink reversal, and ethanol precipitation. Biotinylated fragments were captured with streptavidin beads, and on-bead PCR amplification generated final libraries for Illumina sequencing. Two biological replicates were prepared for each sample.

ChIP-Seq

A published method [68] was followed for all ChIP-seq experiments. Cells first underwent fixation in 1% formaldehyde diluted in PBS for 10 minutes at ambient temperature, with glycine subsequently added to halt crosslinking. After PBS rinses, samples were either processed immediately or snap-frozen in liquid nitrogen and kept at -80°C. Nuclear lysis was performed, after which sonication fragmented chromatin to 200-500 bp. To minimize nonspecific binding, samples were incubated with Protein A/G Dynabeads for 2 hours at 4°C. Antibody-coated beads prepared overnight were then used for immunoprecipitation at 4°C. Following capture, bead-bound complexes were tagmented using Tn5 transposase. Elution, crosslink reversal, and purification were carried out using standard procedures. Final amplification employed Illumina Nextera barcoded primers with KAPA SYBR FAST qPCR Master Mix for 5-12 cycles, yielding sufficient DNA for sequencing.

ATAC-seq

ATAC-seq analysis [67] was performed on control and siRNA-treated H9 cells together with their differentiated progeny. For SC-β organoids, an electric homogenizer disrupted the samples within 10 seconds, producing small cell clusters. After PBS washes, nuclei were released by resuspending cells in cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin) on ice. A single wash with lysis buffer lacking NP40 and digitonin followed. Tagmentation was carried out at 37°C for 20 minutes with occasional shaking in a Tn5 reaction mix containing 25 µl 2x TD buffer, 2.5 µl Tn5 (100 nM final), 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20, and 5 µl H2O. DNA purification was performed using the Universal DNA Purification Kit (Tiangen, DP214). Purified fragments then served as templates for Illumina Nextera library construction, with final sequencing conducted on a NovaSeq 6000 platform.

Flow cytometric analysis of cell cycle and apoptosis

Cultured Ctrl and siRNA treated PP2 spheroids were dissociated to single cells using Accutase (Thermol, A1110501) and stained with propidium iodide (PI) using the Cell Cycle and Apoptosis Analysis Kit/Annexin V-FITC Apoptosis Detection Kit (Beyotime, C1052/C1062S) according to the manufacturer's protocols. Stained samples were subsequently analyzed on a CytExpert flow cytometer. Data was collected using software for Beckman CytoFlex machine and analyzed using FlowJo v10.8.1.

Data Processing

ChIP-seq data processing

Sequencing reads were aligned exclusively to the hg38 human reference genome using Bowtie2. Duplicate entries arising from PCR amplification were identified and discarded. Coverage tracks in bedgraph format were generated with the Bedtools genomeCoverage utility. To enable signal comparisons across different ChIP-seq datasets, read counts were equalized through random subsampling. These normalized alignments then produced bedgraph files visualized in IGV. Peak calling employed MACS2 with default settings, using IgG controls as background. The edgeR [79] R package identified differentially enriched regions, applying thresholds of $p<0.05$ and at least 20% fold change. Transcription factor motif analysis was performed on both ChIP-seq peaks and differential loop anchors using MEME. To determine motif occurrence frequencies, FIMO scanned these genomic intervals against the JASPAR_CORE_2016_vertebrates database.

RNA-seq data processing

HISAT2 (v2.2.0) was run with default settings to map raw RNA-seq reads onto the human hg38 reference genome. Gene and transcript expression levels, quantified as transcripts per million (TPM), were calculated using StringTie (v2.1.6) based on existing annotations. To identify differentially expressed genes, the edgeR package was applied, with significance defined as $p \leq 0.05$ and absolute fold change exceeding 20%.

Hi-C and HiChIP data processing

HiC-Pro v2.10.0[80] aligned paired-end reads from both Hi-C and HiChIP experiments to the hg38 human genome. After removing PCR duplicates and low-quality alignments, valid interaction pairs were assigned to DpnII restriction fragments and converted into binned contact matrices (hic format). For visualization and analysis, Knight-Ruiz (KR) normalization was applied to Hi-C matrices, while Vanilla coverage square root (VCsqrt) normalization was used for HiChIP matrices, both via dump command of juicebox [81]. CTCF-mediated loops in Hi-C data were identified with SIP v1.3.3 (https://github.com/PouletAxel/SIP/releases)[82], and significant interactions at 5, 10, and 25 kb resolutions were detected using Fit-Hi-C (https://github.com/ay-lab/fithic) with $q<0.001$. For HiChIP data, FitHiChIP (https://ay-lab.github.io/FitHiChIP/)[83] generated singleton reads to identify protein-bound regions and detected significant long-range contacts (FDR<0.05) associated with specific factors.

Chromatin interaction analysis

Juicer [81] generated contact matrices from raw Hi-C and HiChIP data. All matrices underwent distance normalization using (observed – expected) / (expected + 1) to remove distance-dependent bias. For cross-sample comparisons, matrices were randomly downsampled to equalize total contacts. A/B compartments were derived from eigen decomposition of 25 kb Pearson correlation matrices[84]. CTCF loops were called with SIP at 5 kb resolution using: -norm KR -min 2.0 -max 2.0 -mat 2000 -d 6 -res 5000 -sat 0.01 -t 2500 -nbZero 6 -factor 1 -fdr 0.05 -del true -cpu 48 -factor 4. All loop calls were merged into a master list.

Differential loop analysis

Hi-C library comparisons followed three normalization steps: 1) random subsampling to match the smallest library; 2) KR balance normalization; 3) distance normalization via (observed – expected) / (expected + 1). Differential loops were identified by: 1) combining SIP-detected loops across stages; 2) extracting KR- and distance-normalized contact frequencies for each loop; 3) running edgeR on paired frequencies (FDR < 0.1, p < 0.05, fold change ≥ 4). Stage-specific loops were defined by: 1) aggregating all differential loops; 2) calculating per-stage contact frequencies; 3) ranking loops by stage of peak (gained) or nadir (lost) values; 4) assigning stage specificity where maximum change occurred. Loops not meeting stage-specific criteria after FDR ≥ 0.1, p ≥ 0.05, or fold change < 4 filtering were deemed constitutive. MEME-ChIP identified motif enrichment at upregulated loop anchors using overlapping ATAC-seq peak summits.

Virtual 4C signal generation

Virtual 4C signals were generated from Hi-C data using Coolbox (available at https://gangcaolab.github.io/CoolBox) alongside in-house Python scripts. In summary, virtual 4C was conducted by utilizing the Hi-C KR normalized observed over expected (O/E) contact frequency for each bin that interacts with the anchor. To account for differences in intensities at various resolutions, individual high-resolution bins were combined using a sliding window corresponding to the resolution step, and the average signal across the combined bins was calculated for each examined resolution step until the lowest specified resolution was reached. The final normalized enrichment scores presented in Figures 4 and 5 were produced by averaging the intensity of each bin across all resolutions from 10 kb to 500 kb.

APA metaplot analysis of chromatin interactions

To evaluate loop enrichment relative to local interaction backgrounds while controlling for distance-dependent decay and protein occupancy artifacts, we adopted the aggregate peak analysis (APA) framework[26]. Interaction matrices binned at 10 kb resolution served as the foundation for this approach. For each loop set under investigation, we isolated KR-normalized observed-over-expected (O/E) contact values corresponding to the loop pixel itself together with flanking regions extending ten bins along both axes. These extractions produced localized 210 kb × 210 kb matrices capturing interaction landscapes surrounding each loop. Position-specific median O/E values were computed across all matrices belonging to a given loop category and rendered as heatmaps through the heatmap.2 R package. Enrichment metrics were derived by comparing the central pixel intensity against averaged signals from nine pixels clustered in the bottom-right quadrant of each aggregate plot.

Metaplots examining configurations with multiple anchor points were constructed at 10 kb resolution, with intervals between anchors evenly partitioned into ten segments[68]. For experimental setups involving three or four CTCF anchor sites, anchors exhibiting stable interaction patterns were consistently assigned to the leftmost positions. To ensure valid comparisons across libraries derived from different experimental conditions, raw interaction data underwent sequential processing: valid paired-end tags were first randomly subsampled to achieve uniform sequencing depth across all samples, followed by KR balance normalization applied uniformly to the resulting contact matrices. Differential interaction signals between treatment and control conditions at specified anchor groupings were obtained by subtracting control values from treatment values on a per-loop basis, with median measurements across all anchors subsequently aggregated for visualization.

Integration of ChIP-seq signals with chromatin loop anchors

Loop anchor coordinates from SIP and Mango[85] were first extended by ±5 kb and then intersected with ChIP-seq peak regions using bedtools intersect. To establish a background expectation, ChIP-seq peaks were randomly permuted across the genome 1,000 times, with the same intersection procedure applied to each shuffled set. Empirical p-values were derived by calculating the proportion of randomized iterations in which the observed overlap count exceeded that obtained from the actual peak distributions.

Definition of genomic features

Promoter regions were defined as RefSeq transcription start sites (TSS) extended by 1 kb in both directions. Enhancer identification relied on H3K4me1 peaks that lacked H3K4me3 signal and intersected with ATAC-seq nucleosome-free regions (NFR), while excluding positions within 1 bp of any TSS. From this enhancer pool, those also containing H3K27ac marks were classified as active. Differential activity between conditions was assessed using H3K27ac ChIP-seq signals processed through the edgeR package, with significance thresholds set at $p < 0.05$ and absolute fold change exceeding 3. Super-enhancer annotation was performed with ROSE[86] using default parameters on H3K27ac data. Polycomb-associated domains were mapped by scanning the genome with 1 kb sliding windows across 10 kb bins, retaining windows where RNF2 ChIP-seq signals exceeded 10 RPMPB, followed by merging of adjacent regions.

ChIP-seq signal visualization and clustering

Genomic intervals (anchors with flanking regions) were partitioned into uniform bins to construct an empty matrix (anchors × bin positions). For cross-sample comparisons, reads from matching antibody experiments were randomly subsampled to ensure equivalent sequencing depth. Normalized reads were assigned to individual genomic bins via bedtools intersect, generating read counts for each matrix position. Raw counts were divided by library size in millions to obtain reads per million per bin (RPMPB), visualized as heatmaps using Java TreeView. Average profiles were generated by calculating mean RPMPB values across all anchors at each distance interval from the anchor center. For clustering, signal intensities from the central six bins (center ± 3 bins) were extracted and subjected to K-means partitioning using Cluster3.

ATAC-seq data processing

A customized computational approach was implemented for ATAC-seq analysis. Using Bowtie2 with parameters -X 2000 -m 1, paired-end reads were mapped to the human hg38 reference genome while retaining default settings elsewhere. Picard Tools (http://picard.sourceforge.net; https://broadinstitute.github.io/picard/) eliminated PCR duplicate reads. Fragment size corrections involved shifting forward-strand reads by +4 bp and reverse-strand reads by -5 bp. Fragment length stratification separated reads into two categories: ATAC-TF (50-115 bp fragments) marking sub-nucleosomal regions accessible to transcription factors, and ATAC-Nuc (180-247 bp fragments) indicating mono-nucleosomal protected DNA. Bedtools genomeCoverage generated bedgraph files. Transcription factor-associated open chromatin regions were identified through MACS2 peak calling restricted to ATAC-TF reads. Enrichment patterns for both ATAC-TF and ATAC-Nuc were visualized using heatmaps and average profiles, following the same analytical framework established for ChIP-seq datasets. Cluster3 performed K-means clustering on ATAC-seq heatmaps for unsupervised classification.

Statistics and Reproducibility

For comparative analyses between two groups, statistical significance was assessed using two-tailed Student's t-test or Wilcoxon rank-sum test. P values < 0.05 were considered statistically significant, where * represents P < 0.05, ** represents P < 0.01, *** represents P < 0.001, **** represents P < 0.0001 and ns represents P>=0.05. Multiple testing correction was performed using the Benjamini-Hochberg false discovery rate (FDR) method where applicable. Specific statistical tests for each figure panel are detailed in the corresponding figure legends.

All sequencing experiments were performed with at least n = 2 biological replicates, unless otherwise stated in the figure legends. A biological replicate is defined as cells or samples derived from independent cell cultures, transfections and differentiations. Sample sizes were chosen based on standard practices in the field and similar previous studies. Experiments lacking formal statistical analysis (e.g., representative microscopy images, immunoblots) were repeated independently at least three times with consistent results. Data are presented as mean ± standard error of the mean (SEM), as indicated in the figure legends.

Data Availability

RNA-seq, ATAC-seq, ChIP-seq and HiChIP data are accessible from NCBI's Gene Expression Omnibus (GEO) under the accession number GSE288172. Hi-C data can be found under accession number GSE288728. Numerical source data underlying Fig 4e are provided in Supplementary Data 2. Numerical source data underlying Fig 5d-f are provided in Supplementary Data 3. Numerical source data underlying Fig 6d and 7e are provided in Supplementary Data 4.

REFERENCES

1 Carter, B. & Zhao, K. J. The epigenetic basis of cellular heterogeneity. *Nature Reviews Genetics* **22**, 235-250 (2021). https://doi.org:10.1038/s41576-020-00300-0

2 Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nature Reviews Genetics* **19**, 789-800 (2018). https://doi.org:10.1038/s41576-018-0060-8

3 Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nature reviews. Molecular cell biology* (2019). https://doi.org:10.1038/s41580-019-0132-4

4 Jerkovic, I. & Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nature Reviews Molecular Cell Biology* **22**, 511-528 (2021). https://doi.org:10.1038/s41580-021-00362-w

5 Dekker, J. & Mirny, L. A. The chromosome folding problem and how cells solve it. *Cell* **187**, 6424-6450 (2024). https://doi.org:10.1016/j.cell.2024.10.026

6 Bouwman, B. A. & de Laat, W. Getting the genome in shape: the formation of loops, domains and compartments. *Genome Biol* **16**, 154 (2015). https://doi.org:10.1186/s13059-015-0730-1

7 Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-+ (2018). https://doi.org:10.1016/j.cell.2018.10.042

8 Strom, A. R. *et al.* Phase separation drives heterochromatin domain formation. *Nature* **547**, 241-+ (2017). https://doi.org:10.1038/nature22989

9 Akilli, N., Cheutin, T. & Cavalli, G. Phase separation and inheritance of repressive chromatin domains. *Current Opinion in Genetics & Development* **86** (2024). https://doi.org:ARTN 102201

10.1016/j.gde.2024.102201

10 Nichols, M. H. & Corces, V. G. Principles of 3D compartmentalization of the human genome. *Cell Reports* **35** (2021). https://doi.org:ARTN 109330

10.1016/j.celrep.2021.109330

11 Harris, H. L. *et al.* Chromatin alternates between A and B compartments at kilobase scale for subgenic organization. *Nature communications* **14** (2023). https://doi.org:ARTN 3303

10.1038/s41467-023-38429-1

12 Wei, M. T. *et al.* Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* **22**, 1187-+ (2020). https://doi.org:10.1038/s41556-020-00578-6

13 Larson, A. G. *et al.* Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236-+ (2017). https://doi.org:10.1038/nature22822

14 Spracklin, G. *et al.* Diverse silent chromatin states modulate genome compartmentalization and loop extrusion barriers. *Nature Structural & Molecular Biology* **30**, 38-+ (2023). https://doi.org:10.1038/s41594-022-00892-7

15 Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nature Reviews Molecular Cell Biology* **22**, 653-670 (2021). https://doi.org:10.1038/s41580-021-00387-1

16 Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944 e922 (2017). https://doi.org:10.1016/j.cell.2017.05.004

17 Merkenschlager, M. & Nora, E. P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annual Review of Genomics and Human Genetics* (2016). https://doi.org:10.1146/annurev-genom-083115-022339

18 Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A* **115**, E6697-E6706 (2018). https://doi.org:10.1073/pnas.1717730115

19 Magnitov, M. & de Wit, E. Attraction and disruption: how loop extrusion and compartmentalisation shape the nuclear genome. *Current Opinion in Genetics & Development* **86** (2024). https://doi.org:ARTN 102194

10.1016/j.gde.2024.102194

20 Fan, X., Abbott, T. E., Larson, D. & Chen, K. BreakDancer: Identification of Genomic Structural Variation from Paired-End Read Mapping. *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]* **45**, 15.16.11-11 (2014). https://doi.org:10.1002/0471250953.bi1506s45

21 Bauer, B. W. *et al.* Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. *Cell* **184**, 5448-+ (2021). https://doi.org:10.1016/j.cell.2021.09.016

22 Davidson, I. F. & Peters, J. M. Genome folding through loop extrusion by SMC complexes. *Nature Reviews Molecular Cell Biology* **22**, 445-464 (2021). https://doi.org:10.1038/s41580-021-00349-7

23 Kim, Y., Shi, Z. B., Zhang, H. S., Finkelstein, I. J. & Yu, H. T. Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345-+ (2019). https://doi.org:10.1126/science.aaz4475

24 Davidson, I. F. *et al.* DNA loop extrusion by human cohesin. *Science* **366**, 1338-+ (2019). https://doi.org:10.1126/science.aaz3418

25 Li, Y. *et al.* The structural basis for cohesin-CTCF-anchored loops. *Nature* **578**, 472-+ (2020). https://doi.org:10.1038/s41586-019-1910-z

26 Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014). https://doi.org:10.1016/j.cell.2014.11.021

27 Busslinger, G. A. *et al.* Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503-+ (2017). https://doi.org:10.1038/nature22063

28 Mach, P. *et al.* Cohesin and CTCF control the dynamics of chromosome folding. *Nature Genetics* **54** (2022). https://doi.org:10.1038/s41588-022-01232-7

29 Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6** (2017). https://doi.org:ARTN e25776

10.7554/eLife.25776

30 Vian, L. *et al.* The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* **173**, 1165-1178 e1120 (2018). https://doi.org:10.1016/j.cell.2018.03.072

31 Lengronne, A. *et al.* Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573-578 (2004). https://doi.org:10.1038/nature02742

32 Davidson, I. F. *et al.* Rapid movement and transcriptional re-localization of human cohesin on DNA. *The EMBO journal* (2016). https://doi.org:10.15252/embj.201695402

33 Ciosk, R. *et al.* Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Molecular Cell* **5**, 243-254 (2000). https://doi.org:Doi 10.1016/S1097-2765(00)80420-7

34 Petela, N. J. *et al.* Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. *Molecular Cell* **70**, 1134-+ (2018). https://doi.org:10.1016/j.molcel.2018.05.022

35 Alonso-Gil, D. & Losada, A. NIPBL and cohesin: new take on a classic tale. *Trends in Cell Biology* **33**, 860-871 (2023). https://doi.org:10.1016/j.tcb.2023.03.006

36 Cuadrado, A. *et al.* Specific Contributions of Cohesin-SA1 and Cohesin-SA2 to TADs and Polycomb Domains in Embryonic Stem Cells. *Cell Reports* **27**, 3500-+ (2019). https://doi.org:10.1016/j.celrep.2019.05.078

37 Rhodes, J., Mazza, D., Nasmyth, K. & Uphoff, S. Scc2/Nipbl hops between chromosomal cohesin rings after loading. *Elife* **6** (2017). https://doi.org:ARTN e30000

10.7554/eLife.30000

38 Alonso-Gil, D., Cuadrado, A., Giménez-Llorente, D., Rodríguez-Corsino, M. & Losada, A. Different NIPBL requirements of cohesin-STAG1 and cohesin-STAG2. *Nature communications* **14** (2023). https://doi.org:ARTN 1326

10.1038/s41467-023-36900-7

39 Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports* **15**, 2038-2049 (2016). https://doi.org:10.1016/j.celrep.2016.04.085

40 Gómez-Marín, C. *et al.* Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proceedings of the National Academy of Sciences* **112**, 7542-7547 (2015). https://doi.org:10.1073/pnas.1505463112

41 Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320 e324 (2017). https://doi.org:10.1016/j.cell.2017.09.026

42 Rhodes, J. D. P. *et al.* Cohesin Disrupts Polycomb-Dependent Chromosome Interactions in Embryonic Stem Cells. *Cell Reports* **30**, 820-+ (2020). https://doi.org:10.1016/j.celrep.2019.12.057

43 Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900-910 (2015). https://doi.org:10.1016/j.cell.2015.07.038

44 Gomes, M. V. *et al.* The cohesin ATPase cycle is mediated by specific conformational dynamics and interface plasticity of SMC1A and SMC3 ATPase domains. *Cell Reports* **43** (2024). https://doi.org:ARTN 114656

10.1016/j.celrep.2024.114656

45 Yuan, X. Y. *et al.* Molecular mechanism and functional significance of Wapl interaction with the Cohesin complex. *Proceedings of the National Academy of Sciences of the United States of America* **121** (2024). https://doi.org:ARTN e2405177121

10.1073/pnas.2405177121

46 Hsieh, T. H. S. *et al.* Enhancer-promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nature Genetics* **54** (2022). https://doi.org:10.1038/s41588-022-01223-8

47 van Ruiten, M. S. *et al.* The cohesin acetylation cycle controls chromatin loop length through a PDS5A brake mechanism. *Nature Structural & Molecular Biology* **29**, 586-+ (2022). https://doi.org:10.1038/s41594-022-00773-z

48 Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-+ (2017). https://doi.org:10.1016/j.cell.2017.05.004

49 Haarhuis, J. H. I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693-707 (2017). https://doi.org:10.1016/j.cell.2017.04.013

50 Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51-+ (2017). https://doi.org:10.1038/nature24281

51 Liu, N. Q. *et al.* WAPL maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. *Nature Genetics* **53** (2021). https://doi.org:10.1038/s41588-020-00744-4

52 Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51-56 (2017). https://doi.org:10.1038/nature24281

53 Cremer, M. *et al.* Cohesin depleted cells rebuild functional nuclear compartments after endomitosis. *Nat Commun* **11**, 6146 (2020). https://doi.org:10.1038/s41467-020-19876-6

54 Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-+ (2017). https://doi.org:10.1016/j.cell.2017.09.026

55 Cremer, M. *et al.* Cohesin depleted cells rebuild functional nuclear compartments after endomitosis. *Nature communications* **11** (2020). https://doi.org:ARTN 6146

10.1038/s41467-020-19876-6

56 Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572 e524 (2017). https://doi.org:10.1016/j.cell.2017.09.043

57 Fraser, J. *et al.* Hierarchical folding and reorganization of chromosomes are linked to transcriptional changesin cellular differentiation. *Mol Syst Biol* **11** (2015). https://doi.org:ARTN 852

DOI 10.15252/msb.20156492

58 Pekowska, A. *et al.* Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. *Cell Systems* **7**, 482-+ (2018). https://doi.org:10.1016/j.cels.2018.09.003

59 Arzate-Mejía, R. G., Recillas-Targa, F. & Corces, V. G. Developing in 3D: the role of CTCF in cell differentiation. *Development* **145** (2018). https://doi.org:ARTN dev137729

10.1242/dev.137729

60 Parenti, I. & Kaiser, F. J. Cornelia de Lange Syndrome as Paradigm of Chromatinopathies. *Front Neurosci-Switz* **15** (2021). https://doi.org:ARTN 774950

10.3389/fnins.2021.774950

61 Heller, S. *et al.* Transcriptional changes and the role of ONECUT1 in hPSC pancreatic differentiation. *Commun Biol* **4** (2021). https://doi.org:ARTN 1298

10.1038/s42003-021-02818-3

62 Xie, R. *et al.* Dynamic chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. *Cell Stem Cell* **12**, 224-237 (2013). https://doi.org:10.1016/j.stem.2012.11.023

63 Li, J. F. *et al.* TET1 dioxygenase is required for FOXA2-associated chromatin remodeling in pancreatic beta-cell differentiation. *Nature communications* **13** (2022). https://doi.org:ARTN 3907

10.1038/s41467-022-31611-x

64 Bramswig, N. C. *et al.* Epigenomic plasticity enables human pancreatic α to β cell reprogramming. *J Clin Invest* **123**, 1275-1284 (2013). https://doi.org:10.1172/Jci66514

65 Xu, C. R. *et al.* Dynamics of genomic H3K27me3 domains and role of EZH2 during pancreatic endocrine specification. *Embo J* **33**, 2157-2170 (2014). https://doi.org:10.15252/embj.201488671

66 Chen, A. C. H. *et al.* Hyperglycemia impedes definitive endoderm differentiation of human embryonic stem cells by modulating histone methylation patterns. *Cell Tissue Res* **368**, 563-578 (2017). https://doi.org:10.1007/s00441-017-2583-2

67 Lyu, X., Rowley, M. J., Kulik, M. J., Dalton, S. & Corces, V. G. Regulation of CTCF loop formation during pancreatic cell differentiation. *Nature communications* **14** (2023). https://doi.org:ARTN 6314

10.1038/s41467-023-41964-6

68 Lyu, X. W., Rowley, M. J. & Corces, V. G. Architectural Proteins and Pluripotency Factors Cooperate to Orchestrate the Transcriptional Response of hESCs to Temperature Stress. *Molecular Cell* **71**, 940-+ (2018). https://doi.org:10.1016/j.molcel.2018.07.012

69 Weintraub, A. S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573-1588 e1528 (2017). https://doi.org:10.1016/j.cell.2017.11.008

70 Jiang, Y. P. *et al.* Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III. *Genome Biology* **21** (2020). https://doi.org:ARTN 158

10.1186/s13059-020-02067-3

71 Yan, J. *et al.* Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. *Cell Research* **28**, 204-220 (2018). https://doi.org:10.1038/cr.2018.1

72 Rodríguez-Carballo, E. *et al.* The HOXD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. *Genes & Development* **31**, 2264-2281 (2017). https://doi.org:10.1101/gad.307769.117

73 Damon, L. J., Aaron, J. & Palmer, A. E. Single molecule microscopy to profile the effect of zinc status on transcription factor dynamics. *Scientific reports* **12** (2022). https://doi.org:ARTN 17789

10.1038/s41598-022-22634-x

74 Williamson, I., Boyle, S., Grimes, G. R., Friman, E. T. & Bickmore, W. A. Dispersal of PRC1 condensates disrupts polycomb chromatin domains and loops. *Life Sci Alliance* **6** (2023). https://doi.org:ARTN e202302101

10.26508/lsa.202302101

75 Pagliuca, F. W. *et al.* Generation of Functional Human Pancreatic beta Cells In Vitro. *Cell* **159**, 428-439 (2014). https://doi.org:10.1016/j.cell.2014.09.040

76 Veres, A. *et al.* Charting cellular identity during human in vitro beta-cell differentiation. *Nature* **569**, 368-+ (2019). https://doi.org:10.1038/s41586-019-1168-5

77 Alvarez-Dominguez, J. R. *et al.* Circadian Entrainment Triggers Maturation of Human In Vitro Islets. *Cell Stem Cell* **26**, 108-122 e110 (2020). https://doi.org:10.1016/j.stem.2019.11.011

78 Li, L. *et al.* Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Molecular Cell* **58**, 216-231 (2015). https://doi.org:10.1016/j.molcel.2015.02.023

79 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010). https://doi.org:10.1093/bioinformatics/btp616

80 Servant, N. *et al.* HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology* **16** (2015). https://doi.org:ARTN 259

10.1186/s13059-015-0831-x

81 Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems* **3**, 95-98 (2016). https://doi.org:10.1016/j.cels.2016.07.002

82 Rowley, M. J. *et al.* Analysis of Hi-C data using SIP effectively identifies loops in organisms from C. elegans to mammals. *Genome Research* **30**, 447-458 (2020). https://doi.org:10.1101/gr.257832.119

83 Bhattacharyya, S., Chandra, V., Vijayanand, P. & Ay, F. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nature communications* **10** (2019). https://doi.org:ARTN 4221

10.1038/s41467-019-11950-y

84 Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289-293 (2009). https://doi.org:10.1126/science.1181369

85 Phanstiel, D. H., Boyle, A. P., Heidari, N. & Snyder, M. P. Mango: a bias-correcting ChIA-PET analysis pipeline. *Bioinformatics* **31**, 3092-3098 (2015). https://doi.org:10.1093/bioinformatics/btv336

86 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319 (2013). https://doi.org:10.1016/j.cell.2013.03.035

AUTHOR CONTRIBUTIONS

X.L. and L.Y. designed the project. X.L. wrote the manuscript. L.Y., Y.L., J.Z., C.H., H.F., S.F., X.Z., W.G. performed all experiments. X.L. and L.Y. performed data analyses.

Ethics declarations

Competing interests

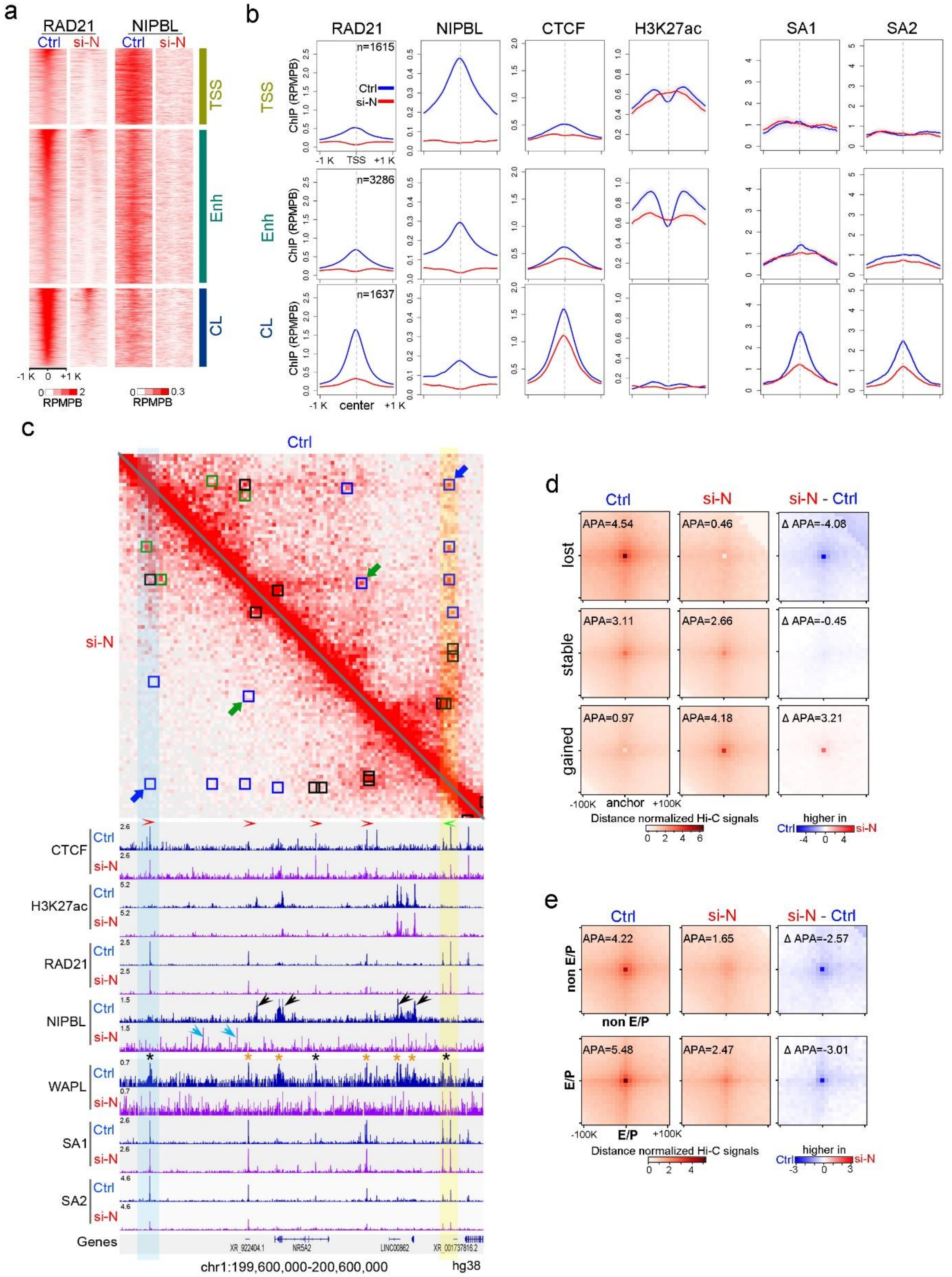
The authors declare no competing interests.

ACKNOWLEDGEMENTS

We thank Dr. Victor Corces for critical reading of the manuscript and valuable suggestions. This work was supported by funds from National Key Research and Development Program of China (2022YFC2704802 to X.L.), National Natural Science Foundation of China (32170546 and 32470573 to X.L.), and National Key Research and Development Program of China (2021YFC2700303 to X.L.).

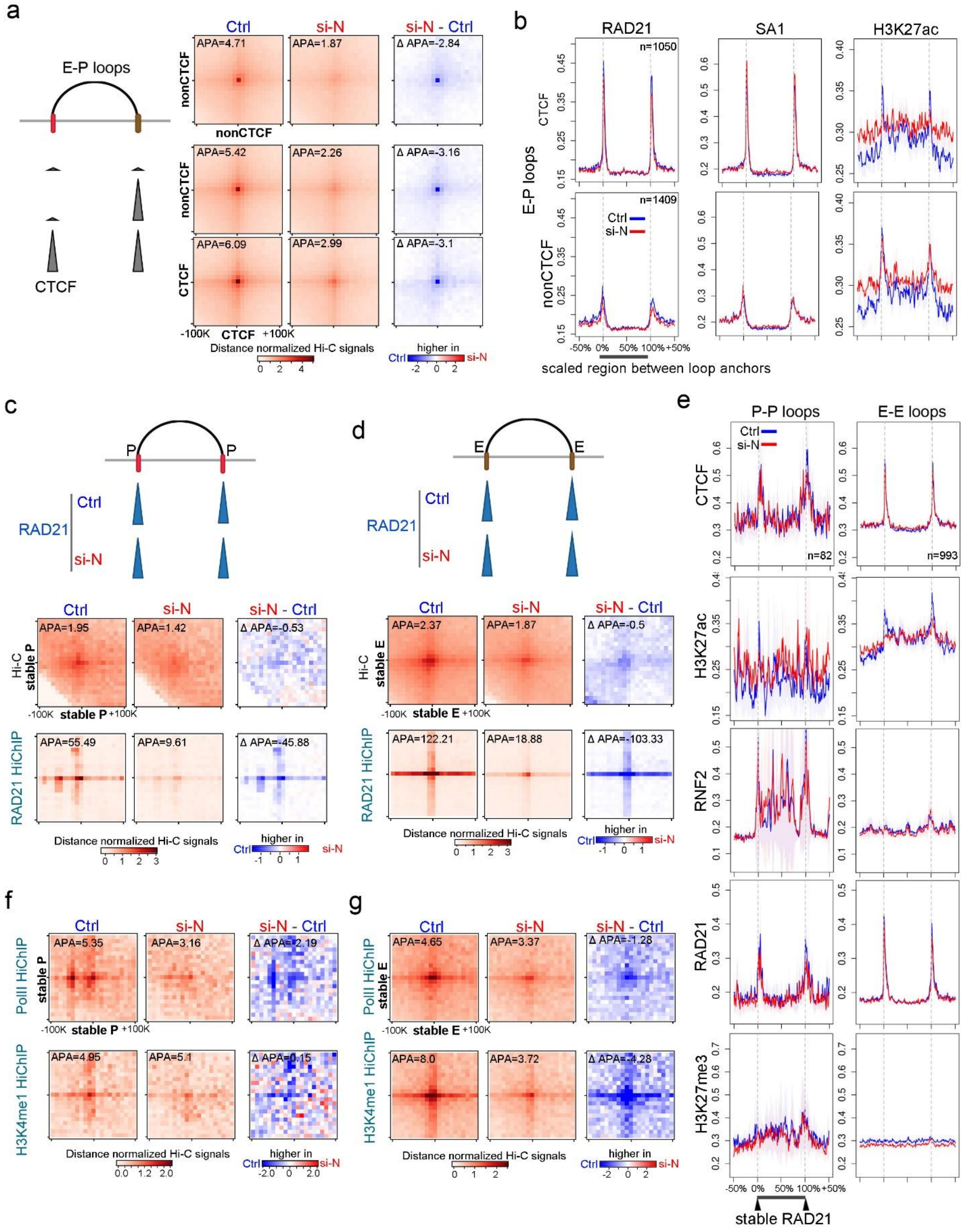
Figures and Legends

Fig. 1 | Redistribution of CTCF loops and enhancer-promoter interactions in hESCs following the knockdown of cohesin loading factor NIPBL.



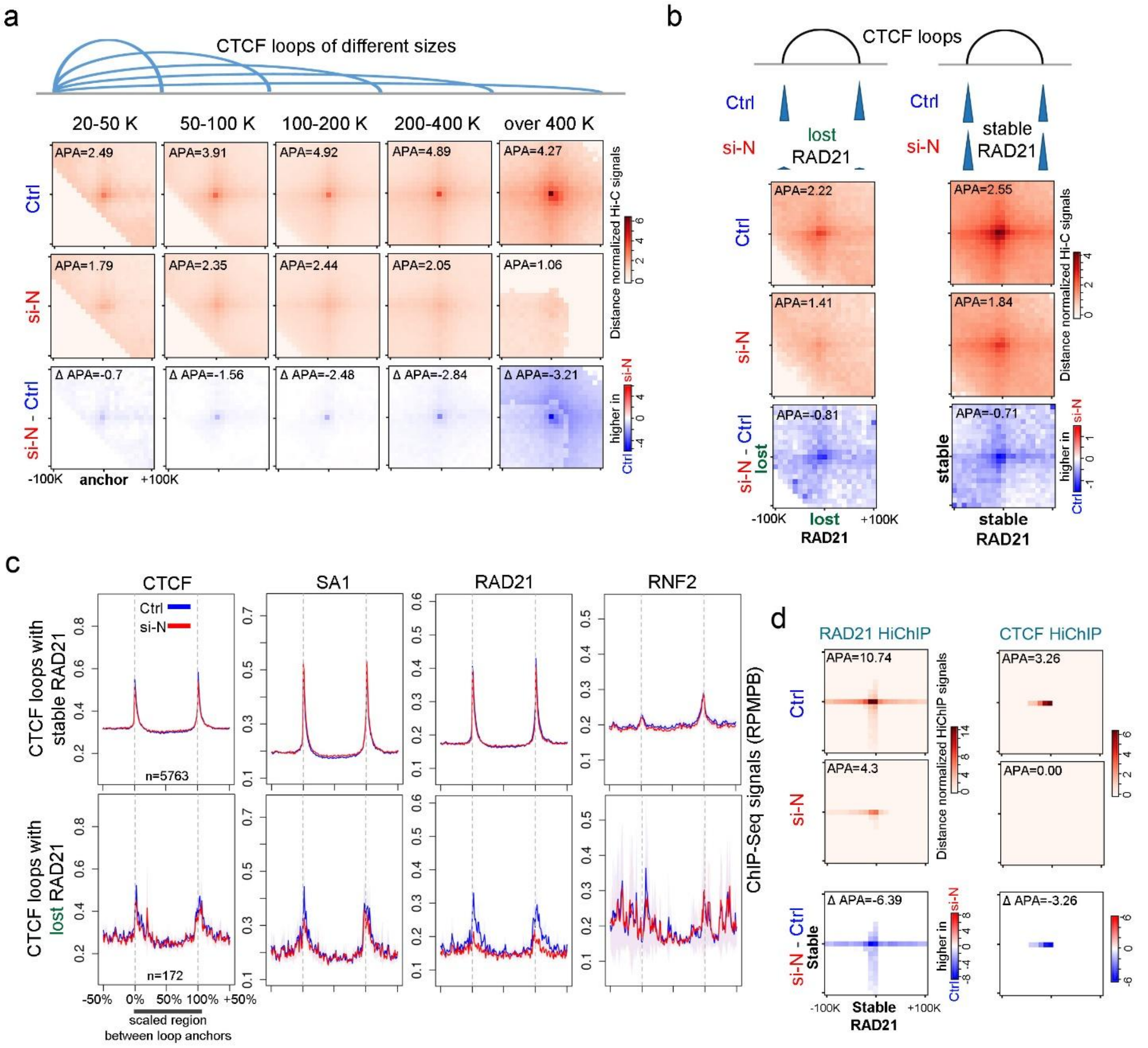
a Heatmaps showing changes induced after NIPBL knockdown (KD) for RAD21 and NIPBL at down-regulated RAD21 peaks. These peaks are divided into four distinct classes, which are promoters (TSS), enhancers (Enh), CTCF sites at loop anchors (CL) and CTCF sites outside loop anchors (CnL, supplementary figure1b) from top to bottom. ChIP-Seq signals within ±1 kb region surrounding transcription start sites (TSSs) or peak summits are plotted. RPMPB: reads per million per 50bp bin. Ctrl: control hESCs transfected with negative control siRNA; si-N: hESCs transfected with siRNA targeting NIPBL. **b** Average profiles of ChIP-seq signals for RAD21, NIPBL, CTCF, H3K27ac, centered on the same set of genomic anchors as in panel (a). Anchors are categorized into three classes: transcription start sites (TSS), enhancers (Enh), and CTCF sites at chromatin loop anchors (CL). Meta-analysis quantifying changes in cohesin subunits SA1 and SA2 after NIPBL KD. For SA1 and A2 , analysis was focused on TSS, Enh, and CL sites, which were first ranked by their baseline SA2 ChIP-seq signal intensity in control cells. The plots display the change in SA1 and SA2 levels specifically for the top quintile (20%) of sites with the highest initial SA2 binding. Lines and shaded regions represent the mean signal and the 95% confidence interval, respectively. **c** Example of SIP loops lost, stable and gained in si-NIPBL with respect to Ctrl hESCs. Blue squares: lost loops; green: gained; black: stable. Green arrows point to the smaller loop located between a pair of tandem CTCF sites, which is lost when RAD21 and SA1 levels are maintained at anchors. Blue arrows indicate the larger loop located between a pair of convergent CTCF sites in the same region, which is also lost upon NIPBL KD. Blue and yellow shading is used to highlight loop anchors formed by forward and reverse pairs of CTCF motifs, respectively. The distribution of input normalized ChIP-Seq signals of several proteins in the same region in both cells is shown. Above the CTCF ChIP-seq track, red and green arrowheads indicate the orientation of individual CTCF motifs. On the NIPBL ChIP-seq track, black and blue arrows mark stable and gained NIPBL binding events, respectively. On the WAPL ChIP-seq track, black and gold asterisks denote stable and decreased WAPL occupancy, respectively. Reference genome: hg38. **d** Aggregate peak analysis (APA) of Hi-C data obtained in ctrl and si-NIPBL hESCs at lost, stable and gained SIP loops. Each square represents the aggregate of the signals present in all corner dots corresponding to loops detected by SIP in either condition (see Methods) and the surrounding 100 kb regions up- and down-stream. Top row: lost loops, middle row: stable loops, bottom row: gained loops. Subtraction of APA metaplots of si-NIPBL minus Ctrl is shown in the last column of each row. **e** APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at SIP loops containing promoters or enhancers at neither, one or both anchors. E: enhancers; P: promoters. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), NIPBL ChIP-seq (n = 3), ChIP-seq for other proteins (n = 2).

Fig. 2 | NIPBL knockdown reduces promoter and enhancer interactions without depleting local cohesin.



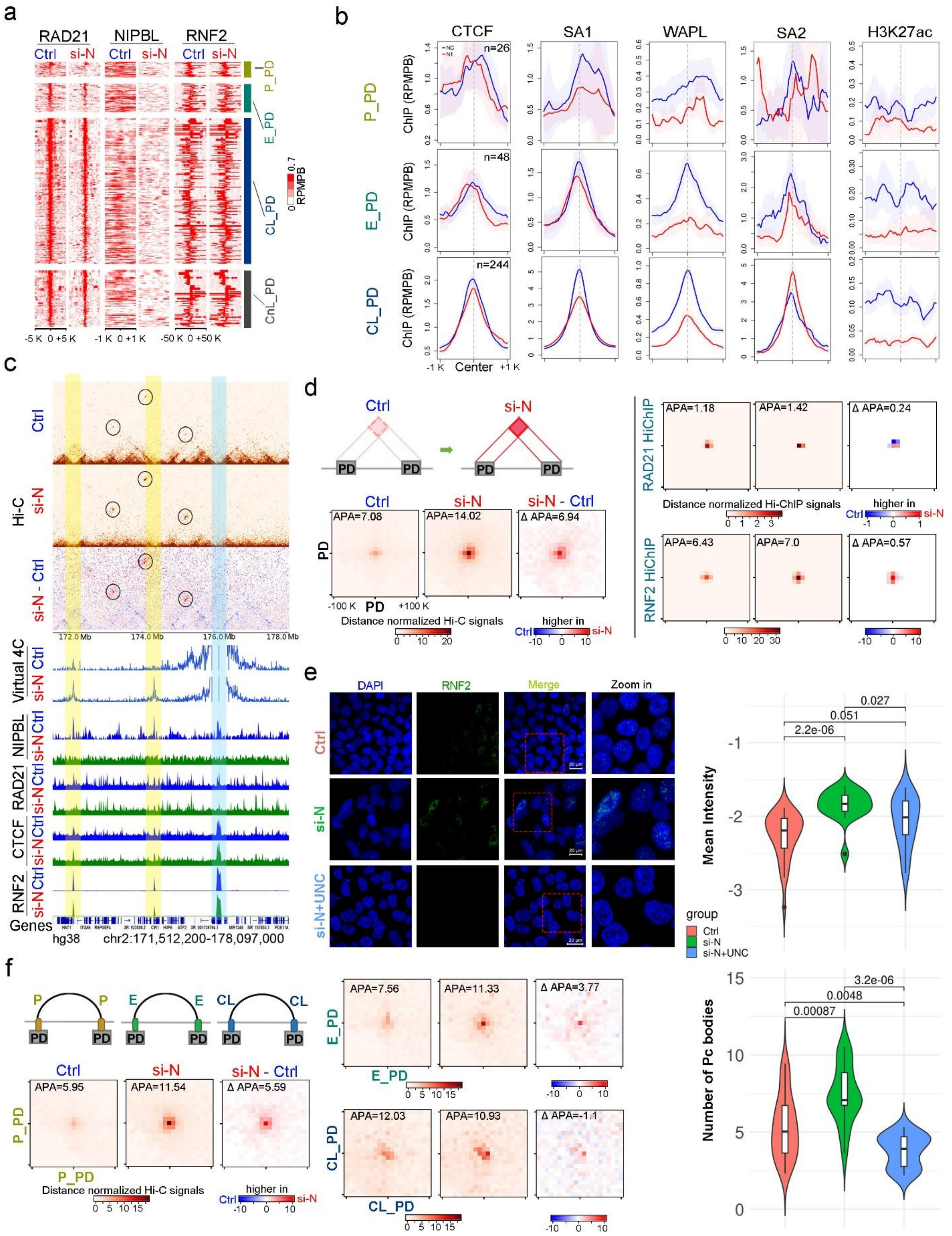
a Diagram showing distinct enhancer-promoter (E-P) loops containing CTCF at neither, one or both anchors. APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at these loops. Subtraction of APA metaplots is listed in the last column of each row. **b** Average profiles of ChIP-seq signals of RAD21, SA1 and H3K27ac at the scaled regions ± 50% of their sizes between loop anchors in (**a**). E: enhancers; P: promoters. Left column: E/P loops containing CTCF at both anchors; right column: E/P loops lacking CTCF at both anchors. **c-d** Diagrams showing promoter (P-P) loops (**c**) and enhancer (E-E) loops (**d**) containing stable levels of RAD21 at both anchors after NIPBL KD. APA metaplots of Hi-C and RAD21 HiChIP data obtained in ctrl and si-NIPBL hESCs at these P-P loops (**c**) and E-E loops (**d**). Subtraction of APA metaplots is listed in the last column of each row. **e** Average profiles of ChIP-seq signals of CTCF, RAD21, H3K27ac, RNF2 and H3K27me3 at the scaled regions ± 50% of their sizes between loop anchors in (**c**) and (**d**). Lines and shaded regions represent the mean signal and the 95% confidence interval, respectively. **f-g** APA metaplots of Pol II and H3K4me1 HiChIP data obtained in ctrl and si-NIPBL hESCs at these P-P loops (**f**) and E-E loops (**g**). Subtraction of APA metaplots is listed in the last column of each row. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), HiChIP (n = 2), ChIP-seq (n = 2).

Fig. 3 | Lost CTCF loops maintain cohesin at anchors despite NIPBL absence.



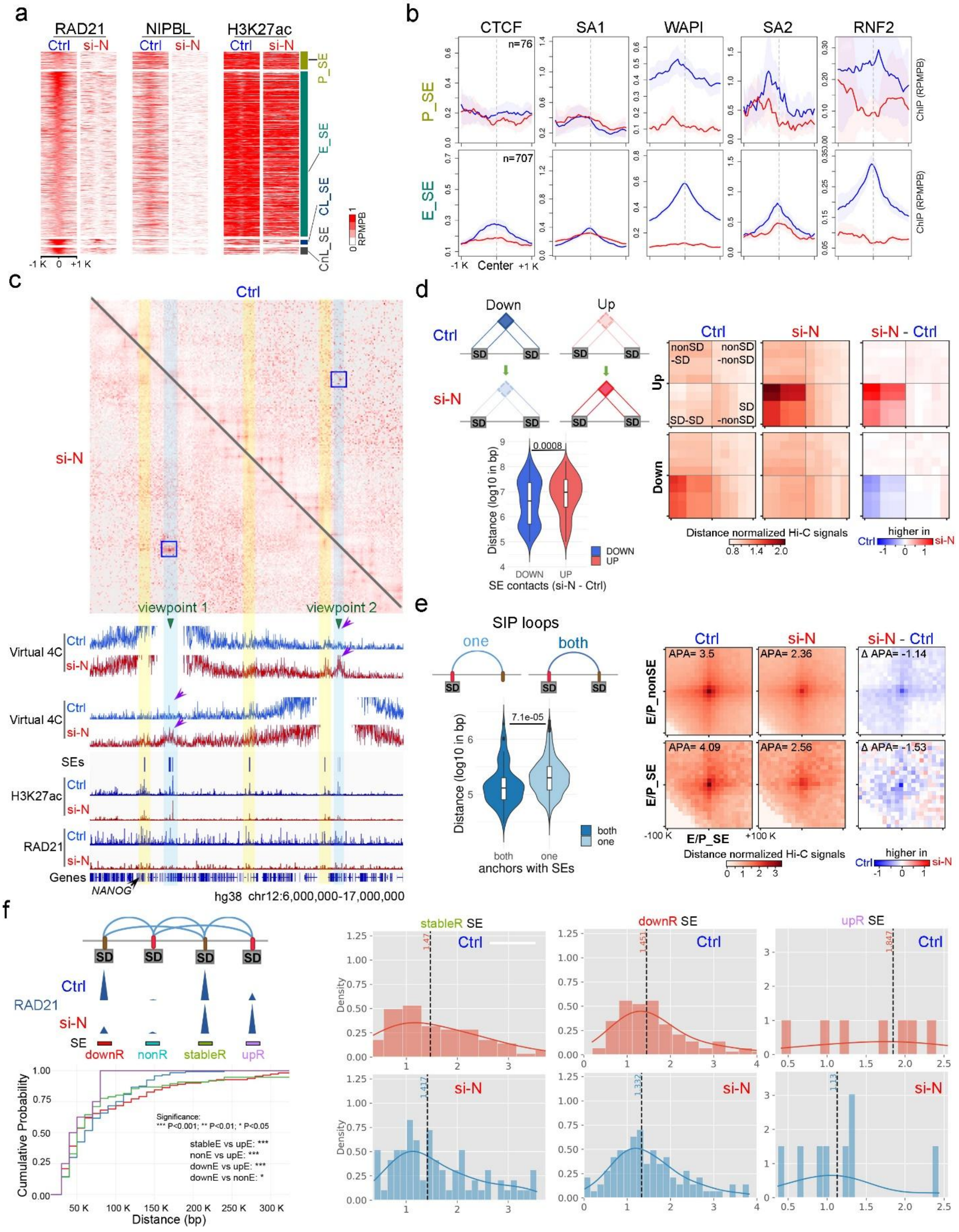
a Diagram describing CTCF loops at different sizes. APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at SIP loops of distinct sizes. Subtraction of APA metaplots is listed in the last row of each column. **b** Diagram depicting CTCF loops anchoring at lost and stable RAD21 peaks upon NIPBL KD. APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at these CTCF loops. Subtraction of APA metaplots is listed in the last row of each column. **c** Average profiles of ChIP-seq signals of CTCF, SA1, RAD21 and RNF2 at the scaled regions ± 50% of their sizes between loop anchors containing stable RAD21 peaks in (**b**) and lost RAD21 peaks in (**b**). Lines and shaded regions represent the mean signal and the 95% confidence interval, respectively. RPMPB: reads per million per 50bp bin. **d** APA metaplots of RAD21 and CTCF HiChIP data obtained in ctrl and si-NIPBL hESCs at CTCF loops containing stable RAD21 peaks at both anchors upon NIPBL KD. Subtraction of APA metaplots is listed in the last row of each column. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), HiChIP (n = 2), ChIP-seq (n = 2).

Fig. 4 | Loss of cohesin-mediated loops and gain of compartmental contacts between PRC domains in the absence of NIPBL.



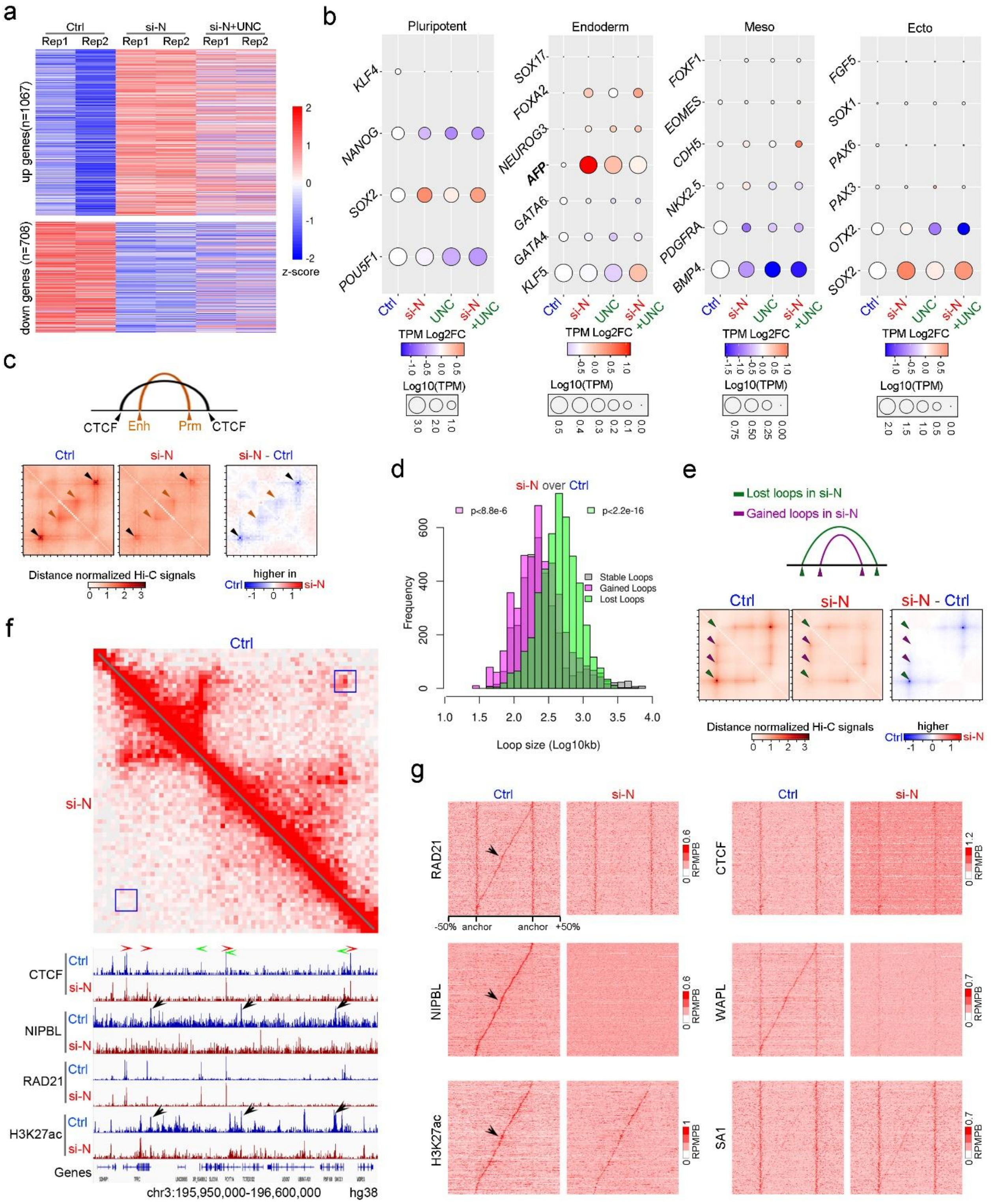
a Heatmaps are presented with protein-specific genomic windows: ±5 kb for RAD21, ±1 kb for NIPBL and ±50 kb for RNF2, optimized to best visualize the distribution of each factor (see methods). These peaks within PRC domains (PDs) are divided into four distinct classes, which are promoters (P_PD), enhancers (E_PD), CTCF sites at loop anchors (CL_PD) and CTCF sites outside loop anchors (CnL_PD) from top to bottom. RPMPB: reads per million per 50bp bin. **b** Average profiles of ChIP-seq signals of CTCF, SA1, WAPL, SA2 and H3K27ac on the same anchors as (**a**). Lines and shaded regions indicate the mean value and 95% confidence interval, respectively. TSSs: transcription start sites; Enh: enhancers; CL: CTCF sites at loop anchors; PD: PRC domains. **c** Example of interactions between PRC domains in si-*NIPBL* with respect to Ctrl hESCs. Black circles in Hi-C heatmaps highlight the interactions between three adjacent PRC domains including *HOXD* gene cluster on chr2. The virtual 4C signals originating from the *HOXD* cluster derived by converting Hi-C data and the distribution of several proteins in the same region in both cells are also shown. Arcs represent the SIP loops. Blue shading: *HOXD* gene cluster as the viewpoint of virtual 4C; yellow shading: two adjacent PRC domains. Reference genome: hg38. **d** Diagram describing the enhanced contacts between PDs upon NIPBL KD. APA metaplots of Hi-C, RAD21 and RNF2 Hi-ChIP data obtained in ctrl and si-NIPBL hESCs among PDs. Sample sizes: Hi-C (n = 3 biologically independent experiments), HiChIP (n = 2). PD: PRC domains. Subtraction of APA metaplots is listed in the last column of each row. **e** Immunofluorescence microscopy of hESCs of Ctrl, si-NIPBL and si-NIPBL treated with UNC3866 stained with antibodies to RNF2 (green). 4’,6-diamidino-2-phenylindole (DAPI) is indicated in blue. Red dashed circles highlighted cells are enlarged for better view on the right side of each row. White scale bars: 20µm. Violin plots and boxplots show the mean intensities and number of PcG bodies within each nucleus for the three groups of hESCs. Boxplots within violin: median (center line), 25th-75th percentiles (box), and 1.5×interquartile range (whiskers). P values are calculated by Wilcoxon rank-sum test. Sample sizes: Immunofluorescence (n=3 biologically independent experiments). **f** Diagram describing the promoter (P_PD), enhancer (E_PD) and CTCF (CL_PD) loops located within PDs. APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at these loops. Sample sizes: Hi-C (n = 3 biologically independent experiments). TSSs: transcription start sites; Enh: enhancers; CL: CTCF sites at loop anchors; PD: PRC domains. Subtraction of APA metaplots is listed in the last column of each row. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), HiChIP (n = 2), NIPBL ChIP-seq (n = 3), ChIP-seq for other proteins (n = 2), Immunofluorescence (n=3).

Fig. 5 | Loss of cohesin-mediated loops within super enhancer domains in the absence of NIPBL.



a Heatmaps showing changes induced after NIPBL knockdown for ChIP-Seq signals of RAD21, NIPBL and H3K27ac at lost RAD21 peaks within super enhancers (SEs) (see methods). These peaks within SEs are divided into four distinct classes, which are promoters (P_SE), enhancers (E_SE), CTCF sites at loop anchors (CL_SE) and CTCF sites outside loop anchors (CnL_SE). RPMPB: reads per million per 50bp bin. **b** Average profiles of ChIP-seq signals of CTCF, SA1, WAPL, SA2 and H3K27ac on the same anchors as (**a**). Lines and shaded regions indicate the mean values and 95% confidence interval, respectively. **c** Example of interactions between SEs in si-NIPBL with respect to Ctrl hESCs as shown in the Hi-C heatmaps. The virtual 4C signals originating from two viewpoints of SEs highlighted in blue rectangles derived by converting Hi-C data and the distribution of H3K27ac and RAD21 in the same region in both cells are also shown. The yellow rectangles highlight the SEs whose interactions are not increased upon NIPBL KD. Reference genome: hg38. **d** Diagrams showing contacts between SEs that are down- and up-regulated upon NIPBL KD. Violin and boxplots showing the distances between SEs of these two classes. Boxplots within violin: median (center line), 25th-75th percentiles (box), and 1.5×interquartile range (whiskers). P values are calculated by Wilcoxon rank-sum test. Sample sizes: Hi-C ($n = 3$ biologically independent experiments). APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at enhanced and weakened interactions between SEs upon NIPBL KD. SD: SE domain. Subtraction of APA metaplots is listed in the last row of each column. **e** Diagrams depicting SIP loops containing SEs at one or both anchors. Violin and boxplots showing sizes of loops of these two classes. Boxplots within violin: median (center line), 25th-75th percentiles (box), and 1.5×interquartile range (whiskers). P values are calculated by Wilcoxon rank-sum test. Sample sizes: Hi-C ($n = 3$ biologically independent experiments). APA metaplots of Hi-C data obtained in ctrl and si-NIPBL hESCs at these two loops. Subtraction of APA metaplots is listed in the last column of each row. **f** Diagrams describing the changes of RAD21 at loop anchors within SDs upon NIPBL KD. Cumulative curves comparing sizes of loops containing no, stable, up- and down-regulated levels of RAD21 at SEs of loop anchors. P values are calculated by Welch's t-test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. APA histograms of Hi-C data obtained in ctrl and si-NIPBL hESCs at these SE loops containing different levels of RAD21. Median values of APA scores are labeled as dashed black lines on histograms. non/stable/up/downR SE: SEs containing non, stable, up- and down-regulated levels of RAD21. nonE/stableE/upE/downE: SEs containing non, stable, up- and down-regulated levels of RAD21. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C ($n = 3$ biologically independent experiments), NIPBL ChIP-seq ($n = 3$), ChIP-seq for other proteins ($n = 2$).

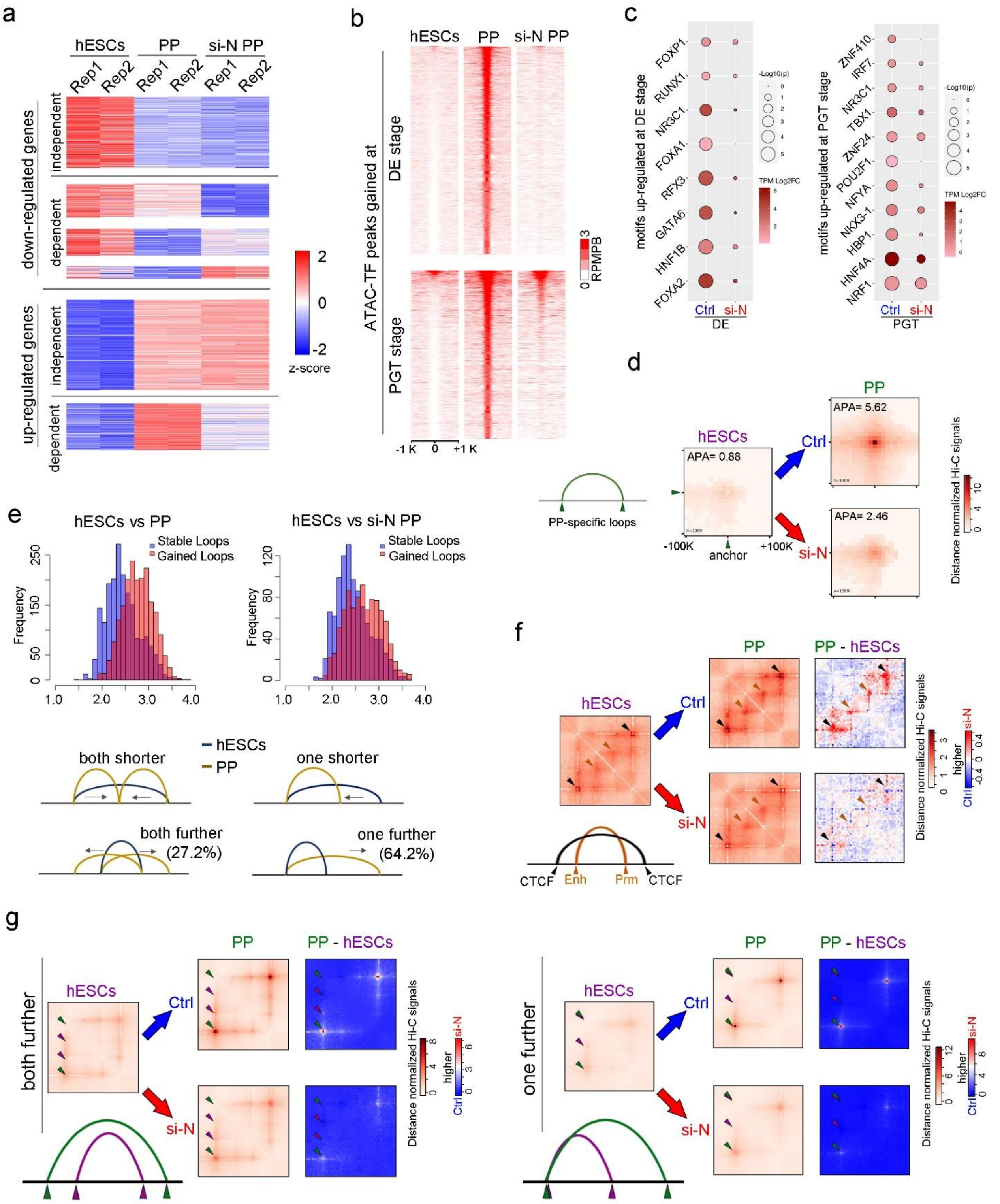
Fig. 6 | Dissolution of interactions between enhancers and down-regulated genes within regions of shortened CTCF loops.



a Heatmaps showing changes in transcription levels of up- and down-regulated genes in different hESCs based on RNA-seq data. **b** Bubble plots showing expression of marker genes of pluripotency and three germ layers in different hESCs based on RNA-seq data. Log2FC: log2 fold change over Ctrl. Ctrl: control;

si-N: si-NIPBL; UNC: UNC3866. **c** Metaplot analysis of interactions between CTCF loop anchors and between enhancers and promoters located within the loop anchors. The top diagram indicates the specific scenario in which the CTCF loops encompass E-P interactions analyzed here. Changes in interactions between E-P pairs are analyzed in the context of changes of the CTCF loops inside which they are contained. The bottom panel shows median heatmaps of Hi-C interactions in Ctrl and si-NIPBL hESCs for CTCF loops that contain pairs of down-regulated genes and their enhancers. The subtraction of heatmaps from both cells is also shown. Decreases of these loops in si-NIPBL hESCs (black arrowheads indicating decreased interactions at the CTCF loop anchors) correlates with decreased E-P interactions inside the loops (orange arrowheads). **d** Comparison of the number and sizes of lost versus gained loops upon NIPBL KD. p values of lost and gained versus stable are labeled on top-right and top-left corner respectively. **e** Metaplot analysis of interactions between CTCF loop anchors that are shortened following the KD of NIPBL. The top diagram illustrates the primary scenario in which Ctrl-specific CTCF loops (green arcs) are shortened by the loss of one of their old anchors (green arrowheads) after NIPBL KD, and smaller loops (purple arcs) are maintained between newly formed anchors (purple arrowheads) and the remaining old anchors. The bottom panel shows the median heatmaps of Hi-C data in both cells on the left side and a subtraction heatmap of Hi-C data comparing both cell types on the right side. **f** One example showing the loss of NIPBL loading site within lost CTCF loops upon NIPBL KD at *SOX2* gene locus. The top panel shows the Hi-C heatmaps. The bottom panel shows the distribution of NIPBL and other proteins at this site. Green and red arrowheads indicate orientation of the CTCF motifs at this site. Black arrowheads show the location of NIPBL sites present in Ctrl but not in si-NIPBL hESCs which contain high levels of H3K27ac. Reference genome: hg38. **g** Distribution of NIPBL and several other proteins at loop anchors and loading sites of CTCF loops identified from Hi-C data present in Ctrl cells but not si-NIPBL hESCs. Black arrowheads show the enrichment of NIPBL and RAD21 at cohesin loading sites, which contain high levels of H3K27ac, between CTCF loops present in Ctrl but not in si-NIPBL hESCs. RPMPB: reads per million per 50bp bin. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), RNA-seq (n = 2), NIPBL ChIP-seq (n = 3), ChIP-seq for other proteins (n = 2).

Fig. 7 | Reduced chromatin accessibility and stage-specific cohesin loops following NIPBL knockdown during pancreatic differentiation.



a Heatmaps showing changes in transcription levels of up- and down-regulated genes during pancreatic differentiation in Ctrl and si-NIPBL derived PP cells. Up- and down-regulated genes at PP stage in Ctrl but not in si-NIPBL cells are clustered as dependent DEGs. **b** Heatmaps showing the ATAC-TF signals of hESCs and differentiated derivatives of Ctrl and si-NIBPL at ATAC-TF peaks gained at DE stage and PGT stage. RPMPB: reads per million per 50bp bin. **c** Changes in the enrichment of transcription factor motifs found at the summits of ATAC-TF peaks specific to DE and PGT stages in Ctrl and si-NIPBL derived differentiated cells. **d** APA metaplots of Hi-C data obtained from both control and si-NIPBL differentiation at the PP stage focusing on the loops that are gained at this specific stage. **e** Comparison of the number and size of stable versus gained loops formed at PP stage during pancreatic cell differentiation in both control and NIPBL KD conditions. Diagram depicting different processes by which CTCF loops change during differentiation by moving the location of one or both loop anchors. Subtraction values of ratios representing different processes occurring in both control and si-NIPBL differentiation at the PP stage are labeled. **f** Metaplot analysis of interactions between CTCF loop anchors and between enhancers and promoters located within the loop anchors. Differentiated PP cells in both Ctrl and NIPBL KD conditions are compared with undifferentiated hESCs. The bottom left diagram indicates the specific scenario in which the CTCF loops encompass E-P interactions analyzed here. Changes in interactions between E-P pairs in two different stages are analyzed in the context of changes of the CTCF loops inside which they are contained. The top panel shows median heatmaps of Hi-C interactions in Ctrl and si-NIPBL PP cells for CTCF loops that contain pairs of up-regulated genes and their enhancers. The subtraction heatmaps from both cells are also shown. Increases of these loops at PP stage (black arrowheads indicating increased interactions at the CTCF loop anchors) correlates with increased E-P interactions inside the loops (orange arrowheads) in Ctrl condition. Less increases of both interactions are found at the same PP stage in NIPBL KD condition. **g** Metaplot analysis of interactions between CTCF loop anchors that are extended at PP stage. The bottom-left diagram illustrates the primary scenario in which hESCs-specific CTCF loops (purple arcs) are extended by replacing both of their old anchors (purple arrowheads) with new anchors (green arrowheads) to form longer loops (green arcs) during differentiation in Ctrl condition. The left panel shows the median heatmaps of Hi-C data and the subtraction heatmaps during pancreatic differentiation in both Ctrl and NIPBL KD conditions. The bottom-right diagram illustrates the primary scenario in which hESCs-specific CTCF loops are extended by replacing one of their old anchors. The right panel shows the meta-analysis of Hi-C data in this scenario. Higher in: the signed difference (treatment − Ctrl) of distance normalized Hi-C signals. Sample sizes: Hi-C (n = 3 biologically independent experiments), RNA-seq (n = 2), ATAC-seq (n = 2).