HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2023 August 14.

Published in final edited form as:

Nat Genet. 2023 June ; 55(6): 1048–1056. doi:10.1038/s41588-023-01391-1.

Region Capture Micro-C reveals coalescence of enhancers and promoters into nested microcompartments

Viraat Y. Goel[1,2,3], **Miles K. Huseyin**[1,2,3], **Anders S. Hansen**[1,2,3,*]

[1]Department of Biological Engineering, Massachusetts Institute of Technology; Cambridge, MA 02139, USA

[2]The Broad Institute of MIT and Harvard; Cambridge, MA 02139, USA

[3]Koch Institute for Integrative Cancer Research; Cambridge, MA, 02139, USA

Abstract

Although enhancers are central regulators of mammalian gene expression, the mechanisms underlying Enhancer-Promoter (E-P) interactions remain unclear. Chromosome conformation capture (3C) methods effectively capture large-scale 3D genome structure but struggle to achieve the depth necessary to resolve fine-scale E-P interactions. Here, we develop Region Capture Micro-C (RCMC) by combining MNase-based 3C with a tiling region-capture approach and generate the deepest 3D genome maps reported with only modest sequencing. By applying RCMC in mouse embryonic stem cells and reaching the genome-wide equivalent of ~317 billion unique contacts, RCMC reveals previously unresolvable patterns of highly nested and focal 3D interactions, which we term microcompartments. Microcompartments frequently connect enhancers and promoters and though loss of loop extrusion and inhibition of transcription disrupts some microcompartments, most are largely unaffected. We therefore propose that many E-P interactions form through a compartmentalization mechanism, which may explain why acute cohesin depletion only modestly affects global gene expression.

INTRODUCTION:

3D genome structure regulates vital cellular processes including gene expression, DNA repair, genome integrity, DNA replication, and somatic recombination[1,2]. Many insights into 3D genome structure have come from Chromosome Conformation Capture (3C)

*Corresponding author. ashansen@mit.edu.

Author Contributions

V.Y.G. and A.S.H. designed the project. V.Y.G. developed and optimized the Region Capture Micro-C protocol with input from M.K.H. and A.S.H. V.Y.G. performed cell culture and RCMC, analyzed sequencing data, compared RCMC to previous studies, performed loop and compartment calling, identification of individual microcompartments, and validation of ICE balancing. M.K.H. performed and analyzed western blotting and Pol II ChIP-seq experiments, and analyzed reproducibility of RCMC datasets and features of microcompartment interactions. V.Y.G. and M.K.H. performed pileup and microcompartment interaction strength analysis. A.S.H. supervised the project. All authors contributed to drafting and editing the manuscript and figures.

Competing Interests

Region Capture Micro-C (RCMC) is being considered for a patent application by the authors.

Code Availability

All custom code and scripts used for data analyses in this manuscript are available on Github at https://github.com/ahansenlab/RCMC_analysis_code and on Zenodo at https://zenodo.org/record/7641852.

assays, which have revealed structural hallmarks across at least three scales. First, active and inactive chromatin segregates into A- and B-compartments through a poorly understood compartmentalization mechanism[3,4]. Second, the genome is folded into loops and local domains called Topologically Associating Domains (TADs) or loop domains[5–8] by loop-extruding cohesin complexes halted at CTCF boundaries[9,10]. Third, while A/B-compartments and TADs generally span hundreds to thousands of kilobases, recent work has hinted at finer scale 3D chromatin interactions, including those linking enhancers and promoters[11–17]. Because enhancers are the primary units of gene expression control in mammals, there has been intense interest in resolving fine-scale enhancer-promoter (E-P) interactions; however, it has remained challenging to resolve fine-scale E-P interactions with current methods[8,18]. This motivated us to develop a 3C method that effectively captures E-P interactions.

Advances in our understanding of 3D genome structure have been primarily driven by: (1) deeper sequencing; (2) improved 3C protocols; and (3) perturbation studies. First, A/B-compartments, TADs, and loops were uncovered as deeper sequencing increased the number of captured unique contacts in 3C experiments from ~8 million[3] to ~450 million[5] to ~5 billion[7], respectively. Second, Micro-C overcomes the resolution limits imposed by Hi-C's dependence on restriction enzymes by digesting chromatin with micrococcal nuclease (MNase); this grants Micro-C nucleosome-resolution and allows it to better resolve finer-scale regulatory interactions, including those between enhancers and promoters[8,11–13,15,19,20]. Third, perturbation studies have yielded profound mechanistic insights into 3D genome structure; for example, protein-depletion studies were pivotal in elucidating the roles of CTCF, cohesin, and associated factors in the formation of TADs and loops[12,21–27].

Nevertheless, despite decreasing sequencing costs, sequencing remains the key bottleneck for 3C assays. For a genome with n linear bins, sequencing costs to populate an n^2 pairwise contact matrix grow quadratically with n. For example, we estimate approximately $1.6 billion in sequencing costs alone to average one read per nucleosome-sized bin across the human genome (a total of $(3.3 \times 10^9 \text{ bp}/150 \text{ bp})^2/2 = 2.4 \times 10^{14}$ reads). To overcome the prohibitive cost of sequencing inherent to current methods and facilitate the study of fine-scale 3D genome structure and enhancer-promoter interactions at loci of interest, we therefore sought to develop a 3C method that (1) strongly increases effective sequencing depth, (2) incorporates the latest advances in 3C-derived protocols, and (3) is cost-effective for perturbation experiments.

Here, we address these three points by combining Micro-C with a tiling region capture approach[28,29] to enrich for entire regions of interest in a method we call Region Capture Micro-C (RCMC). We use RCMC to generate the deepest maps of 3D genome organization reported so far, achieving nucleosome resolution with a fraction of the sequencing of other methods. By reaching the local equivalent of ~317 billion unique contacts genome-wide, we discover patterns of previously unseen, fine-scale, focal, and highly nested 3D interactions in gene-dense loci that we call microcompartments. Microcompartments frequently connect enhancers and promoters, and are largely robust to the loss of loop extrusion and inhibition of transcription, though some microcompartmental loops do change. Taken together, our

results suggest that interactions between enhancers and promoters, now highly resolved by RCMC, may be driven by compartmentalization mechanisms rather than loop extrusion.

Region Capture Micro-C (RCMC): Development and benchmarking

To develop Region Capture Micro-C (RCMC), we optimized the regular Micro-C protocol[11,13,15] to maximize library complexity and combined it with a tiling region capture approach[28,29] (Fig. 1a). Briefly, mouse embryonic stem cells (mESCs) were crosslinked with disuccinimidyl glutarate (DSG) and formaldehyde (FA) and digested to nucleosomes with MNase (Extended Data Fig. 1a–b), after which fragment ends were repaired with biotin-labeled nucleotides and then proximity ligated. After protein removal and reversal of crosslinks, we size-selected and pulled down ligated dinucleosomal fragments, and prepared a Micro-C sequencing library. Avoiding repetitive regions, we designed 80-mer biotinylated oligos tiling 5 regions of interest, each spanning between 425 kb and 1,900 kb (Extended Data Fig. 1c), and pulled-down the tiled regions of interest with 35-49% efficiency in a single step (Extended Data Fig. 1d). After paired-end sequencing and normalization[30] (Extended Data Fig. 1e), we obtained contact maps (Fig. 1a). To validate our RCMC contact maps, we compared them to high-resolution Hi-C[31] and Micro-C[13] for the same regions. Our RCMC data matched both Hi-C[31] and Micro-C[13] data at 2-kb resolution (Extended Data Fig. 1f), was reproducible (Extended Data Fig. 1g), and gave the expected contact frequency scaling (Extended Data Fig. 2a). Thus, RCMC captures all information in target regions obtained in prior multi-billion contact studies[13,31].

Having validated RCMC, we next benchmarked it against other 3C datasets. Despite capturing ~2.6-3.3 billion unique contacts, the deepest Hi-C[31] and Micro-C[13] datasets in mESCs give sparse contact maps at fine (sub-kb) resolutions (Fig. 1b). In contrast, since RCMC focuses its sequencing reads in only regions of interest, almost all 100 bp-sized interaction bins showed at least one interaction for our most deeply sequenced region (*Klf1* Fig. 1b; Extended Data Fig. 2), and our RCMC maps matched genome-wide Micro-C[13] even after downsampling by ~100-fold (Extended Data Fig. 2d–f). Indeed, with relatively modest sequencing (Extended Data Fig. 2c) we captured the genome-wide equivalent of ~317 billion unique contacts at the *Klf1* region.

To visualize the improvements afforded by RCMC, we plotted contact maps comparing RCMC to Hi-C[31] and Micro-C[13] at our 5 captured regions (Extended Data Fig. 3–4). While A/B-compartments, TADs, and CTCF and cohesin-mediated structural loops are well-resolved in prior high-resolution Hi-C[31] and Micro-C[13] studies, resolving enhancer-promoter (E-P) interactions has proven more challenging[8,18]. To test the ability of RCMC to resolve E-P interactions, we captured a region around the *Sox2* gene and its regulatory elements (Fig. 2a). *Sox2* encodes a key pluripotency transcription factor, whose expression in mESCs is controlled by a well-characterized ~100 kb distal enhancer (*Sox2* Control Region (SCR))[32–34]. While long-range *Sox2*-SCR interactions are visible in Hi-C and Micro-C, RCMC resolved the fine-scale substructure of the *Sox2*-SCR interactions: rather than one broad loop, *Sox2* forms multiple individual focal interactions with subelements of the SCR marked by Mediator binding and ATAC peaks (Fig. 2a). Furthermore, RCMC also revealed previously unobservable long-range interactions between a ~600-700 kb distal

region near the *Fxr1* gene and *Sox2* and the SCR as well as strong compartmental exclusion of a ~550 kb intervening region (Extended Data Fig. 4a). Next, we focused on a ~300 kb segment of our most deeply sequenced region, the region around *Klf1* (Fig. 2b). Notably, RCMC revealed patterns of highly focal and nested interactions in the *Klf1* region which are not visible in genome-wide Hi-C or Micro-C data (Fig. 2b). We name these interactions microcompartments (see Discussion for rationale and definition). We conclude that for mapping genomic interactions within specific regions, RCMC outperforms genome-wide Hi-C and Micro-C at a fraction of the cost.

Finally, while our studies were ongoing, the related methods Micro-Capture-C (MCC)[16] and Tiled-Micro-Capture-C (TMCC)[17] were reported. Unlike RCMC, (T)MCC uses only formaldehyde for fixation[35], skips the pull-down of ligation products and the gel purification of dinucleosomes (Fig. 1a), and instead uses sonication to generate small fragments containing both ligated and unligated DNA. This allows (T)MCC to precisely sequence the ligation junction, which for RCMC requires longer-read sequencing. Thus, this affords (T)MCC base-pair resolution when capturing the interactions between regulatory elements[16,17]. However, by not enriching for the informative ligation products, (T)MCC mainly captures unligated DNA fragments, resulting in most sequencing reads being uninformative (Fig. 1b). Indeed, with only slightly deeper sequencing, RCMC captured ~200 million unique >1 kb *cis* contacts in the target regions compared to just ~9-13 million for TMCC, underscoring the more than one order of magnitude higher efficiency of RCMC (Extended Data Fig. 2c). To directly compare RCMC to TMCC, we designed probes against the same *Nanog* region used in TMCC[17]. Due to the less efficient nature of TMCC, even with almost 4-fold higher sequencing at the *Nanog* region, TMCC maps were noisier than RCMC, which became even more evident when we subsampled TMCC's sequencing depth to match RCMC (Fig. 2c; Extended Data Fig. 4b). In summary, we conclude that RCMC is more efficient for general 3D genome structure mapping of a region, while (T)MCC may be applied when it is necessary to resolve ligation junctions with base-pair resolution.

RCMC reveals nested focal interactions in gene-rich regions

RCMC data revealed highly nested and focal interactions in both the *Klf1* and *Ppm1g* regions which were not visible in multi-billion contact genome-wide Hi-C[31] and Micro-C[13] datasets (Fig. 2b, 3a–b; Extended Data Fig. 5a–b). We applied existing loop[36,37] and compartment calling algorithms[36] to identify these interactions, but they did not reliably detect them (Extended Data Fig. 5c). We therefore manually identified 132 anchors forming a total of 1091 focal interactions in the gene-rich *Klf1* and *Ppm1g* regions (Fig. 3a–b; Extended Data Fig. 5d). Furthermore, we validated that these interactions were not due to incomplete contact map normalization[30] (Extended Data Fig. 1e, 6a) nor an artifact of increased accessibility at the anchors (only about half of all ATAC peaks result in "dots" and not all "dots" are anchored by ATAC peaks; Extended Data Fig. 6b–d).

Next, we observed that these interactions resemble both loops and compartments. Like loops, they give rise to focal enrichments ("dots" in Fig. 3a–b) between two anchors and occasionally form contact domains as small as a few kilobases ("squares" in Fig. 3a–b). Like A/B-compartments, they result in nested, tessellated interactions in a checkerboard-like

fashion, with a mean of ~17 interactions per anchor (mean interaction length: ~240 kb) and the most nested anchor forming 52 focal interactions (Fig. 3c–d). Because these highly nested and focal interactions ("dots") resemble fine-scale compartmental interactions (see Discussion), we refer to them as *microcompartments*.

To understand which genomic elements form microcompartments, we investigated the chromatin states of microcompartment anchors (Fig. 3c; Extended Data Fig. 7). About two-thirds of the identified microcompartment anchors overlapped either promoter (~46%) or enhancer (~21%) features (Fig. 3e; Extended Data Fig. 7), with the remaining anchors either corresponding to CTCF and cohesin-bound anchors or unknowns ("Other"). Notably, however, promoters and enhancers formed many more focal interactions (Fig. 3f). Specifically, promoters and enhancers formed a mean of 24 and 18 interactions, respectively, compared to just 5.5 and 7.4 for CTCF and cohesin, and "other" anchors, respectively (Fig. 3f). Indeed, 74% of all annotated microcompartmental dots represented either P-P or E-P interactions, while only 4% of interactions were between anchors which exclusively overlapped CTCF and cohesin (Fig. 3g). Taken together, these observations suggest that microcompartments largely represent nested interactions between promoter and enhancer regions as well as some currently poorly understood "other" regions.

Most microcompartments are robust to loss of loop extrusion

Having identified microcompartments as nested interactions frequently linking enhancers and promoters (Fig. 3a–b), we next took advantage of the cost-effective nature of RCMC to test the roles of loop extrusion and transcription (below) in forming these interactions.

First, we explored the role of cohesin and cohesin-mediated loop extrusion. Acute loss of cohesin strengthens large-scale A/B compartments while simultaneously causing the global loss of TADs, loop domains, and CTCF and cohesin-mediated structural loops[12,21,24,25,27,38]. Therefore, to understand whether cohesin regulates microcompartments, we used our previously validated mESC cell line to acutely deplete the cohesin subunit RAD21 (mESC clone F1M RAD21-mAID-BFP-V5)[12,38] and performed RCMC across all 5 regions with and without 3 hours of cohesin depletion (Fig. 4a; Extended Data Fig. 8a). The cohesin depletion was ~97% efficient (Fig. 4b), diminished the well-characterized CTCF and cohesin-mediated *Fbn2* loop[38] (Extended Data Fig. 8a), led to the expected change in contact frequency[21,23,24] (Extended Data Fig. 8b), and was reproducible between replicates (Supplementary Fig. 1), thus validating the cohesin depletion. As expected, the small fraction of interactions between CTCF and cohesin-bound sites showed large reductions in strength upon cohesin depletion (Fig. 4a,c; Extended Data Fig. 8a). However, the strengths of microcompartmental interactions, including E-P, E-E, and P-P[12,17], were largely unaffected by cohesin depletion (Fig. 4c). Specifically, though we do see clear individual examples of especially P-P interactions that either slightly strengthen (Fig. 4e, i) or strongly weaken (Fig. 4e, ii–iii) after cohesin depletion (Fig. 4d–e), most microcompartmental interactions were largely unaffected (Fig. 4c). We therefore refine the microcompartment definition to interactions largely robust to cohesin depletion (see Discussion for full definition).

Most microcompartments are robust to loss of transcription

Second, we explored the role of transcription. We observed that microcompartments are largely formed between active promoter and enhancer regions (Fig. 3e,g; Extended Data Fig. 7), suggesting a relationship between active transcription and microcompartments. To understand if microcompartments are a downstream consequence of transcription, we abolished transcription by inhibiting transcription initiation by RNA Pol II using triptolide. We chose two timepoints: 45 min, which was previously reported to modestly affect global E-P and P-P stripes[13], and 4 hours, which was recently reported to greatly reduce punctate H3K4me3 (found at active promoters) and H3K27ac (found at active enhancers) marks in mESCs in addition to inhibiting transcription[39]. We performed RCMC across all 5 captured regions and ChIP-Seq gave the expected reduction of RNA Pol II signal, with the 4-hour triptolide treatment more thoroughly eliminating RNA Pol II at promoters and throughout gene bodies (Fig. 5a–b; Extended Data Fig. 9). We observed both weakened and strengthened E-P and P-P interactions (Fig. 5c–e), as well interesting dynamically changing interactions (e.g., Fig. 5c, i increases in strength with 45 min of triptolide treatment, but then weakens after 4 hours). Nevertheless, the strong majority of microcompartmental interactions were largely unaffected by the inhibition of transcription (Fig. 5a, c–e). Our findings differ somewhat from recent preprints reporting global weakening of E-P interactions after RNA Pol II depletion[20] or inhibition[19]. In addition to differences in cell type, treatment, and treatment length, this may be due to the much lower depth (~1-1.7 billion Micro-C contacts)[19,20] used in these studies, which cannot resolve microcompartmental interactions and fine-scale E-P and P-P interactions (Fig. 3a; Extended Data Fig. 2e–f, 5b). Alternatively, since we only observe microcompartments in the very gene-dense *Klf1* and *Ppm1g* regions, prior findings[13,19,20] may apply more to individual, isolated E-P/P-P interactions instead of dense and nested microcompartments.

In summary, we conclude that microcompartments generally do not require transcription at short timescales and are more likely either independent from or formed upstream of transcription rather than forming as a downstream consequence of transcription.

DISCUSSION

Here we introduce RCMC as a new, accessible, and affordable method for mapping 3D genome structure at unprecedented depth. Compared with Micro-Capture-C[16] methods such as TMCC[17], RCMC is much more efficient (Fig. 1b; Extended Data Fig. 2c), thus affording much higher depth with less sequencing. Another approach is to brute-force genome-wide Hi-C or Micro-C; by performing 150 separate Hi-C experiments and sequencing deeper than ever before, a recent preprint by Gu *et al.* reached 33 billion contacts[14]. However, such efforts[14] are expensive, not accessible to most labs, and poorly compatible with perturbation experiments vital to uncovering mechanisms of organization. Instead, with RCMC we reach the local equivalent of 317 billion contacts with relatively modest sequencing (Extended Data Fig. 2c). Thus, while genome-wide Micro-C may still be preferred for unbiased genome-wide 3D genome structure mapping, we propose RCMC as an ideal method for generating ultra-deep 3D contact maps and for perturbation experiments, albeit only for individual regions.

What molecular processes might drive microcompartment formation? Although cohesin-mediated loop extrusion is well-established to generate focal interactions (loops)[9,10], microcompartmental loops are largely robust to acute cohesin removal, and therefore likely not dependent on loop extrusion (Fig. 4a,c). Furthermore, although most microcompartmental loops connect enhancers and promoters, microcompartments are also generally robust to the acute loss of RNA Pol II transcription initiation (Fig. 5a,e). Instead, we propose that nested, multiway, and focal microcompartments correspond to small, punctate A-compartments[14,40,41] that form through a compartmentalization mechanism, perhaps mediated by factors upstream of RNA Pol II initiation such as transcription factors and co-factors or active chromatin states[42]. Indeed, in the field of polymer physics, it is well-established that block copolymers undergo microphase separation[4,43–45] when composed of distinct monomers that preferentially self-interact (Fig. 5f). Intuitively, if active chromatin regions at microcompartment anchors are selectively “sticky” with each other, they will tend to co-segregate, resulting in the formation of nested, focal interactions (Fig. 5f). Microphase separation due to preferential interactions among active loci within a block copolymer might thus explain the formation of the striking pattern of interactions we observe (Fig. 3a–b; 5f). In summary, we tentatively define microcompartments as follows: 1) highly nested, focal interactions that frequently connect promoters and enhancer regions often in gene-rich loci; 2) are formed through a compartmentalization mechanism; and 3) for the most part are independent of loop extrusion and transcription, at least on short timescales.

How do microcompartments compare to previously described 3D genome features? First, previous genome-wide Micro-C studies uncovered widespread short-range Promoter-Promoter and Enhancer-Promoter links (P-P and E-P links)[12,13]. Similarly, many microcompartmental interactions connect promoters and enhancers. RCMC now better resolves these interactions, revealing them to be highly nested, frequently forming dozens of microcompartmental loops. Second, while differences in cell type preclude a direct comparison, the microcompartments described here also share features with the fine-scale A-compartment interactions recently described by Gu *et al.* that were proposed to segregate active enhancers and promoters into small A-compartments[14]. Indeed, examining the Hi-C data of Gu *et al.* at 1 kb resolution reveals structures with similarities to microcompartments, suggesting that microcompartments may be conserved to human cells (Extended Data Fig. 10). Further, along the lines of Gu *et al.*, the microcompartments we observe form small contact domains, and their loops are more punctate as compared to CTCF and cohesin-mediated loops, which are more diffuse[14] (Fig. 4c, 5e).

Finally, our study provides insights into E-P interactions. While some studies propose that cohesin is largely required for E-P interactions[27,46], others have suggested that cohesin is most important for very long-range[47–49] or inducible[48,50] E-P interactions, or that cohesin is largely not required for the maintenance of E-P interactions[12,17]. Except for some CTCF and cohesin-bound enhancers and promoters, our data suggest that most P-P and E-P interactions are mediated by a compartmentalization mechanism distinct from loop extrusion. This may offer a mechanistic explanation for the observation that cohesin is not required for the short-term maintenance of most E-P interactions and that the effects of cohesin depletion on global gene expression are modest[12,17,25].

We end by noting some limitations and future directions. Although we can detect microcompartments with RCMC, further work will be necessary to fully understand their function and effect on gene expression. Furthermore, although we show that microcompartments are largely robust to loss of cohesin and transcription, further work is necessary to identify perturbation(s) that disrupt microcompartments. Subsequent studies will also be necessary to test our proposed compartmentalization mechanism of microcompartment formation, as well as to understand how microcompartments change during differentiation and across the cell cycle. Future imaging studies will also be required to understand the frequency and lifetime of microcompartmental interactions in live cells[38], as well as their multi-way nature. Additionally, new computational tools will be required for automated microcompartment calling and analysis. Nevertheless, many of these questions can now be addressed with RCMC, and RCMC provides an accessible method to deeply resolve 3D genome structure in general and enhancer-promoter interactions in particular across loci, cell types, and disease states.

Methods

Experimental Procedure

Overview of the Region Capture Micro-C experiment—Region Capture Micro-C (RCMC) was developed by merging Micro-C[13] with tiling Region-Capture of a locus[28,29]. An overview of the RCMC protocol is provided below, and a detailed protocol as Supplementary Information. The data generated in this manuscript come from merging of two RCMC biological replicates for each of the five tested conditions (wild-type, transcriptional inhibition for 45 min or for 4 hr, cohesin depletion, and a cohesin depletion control). For four of the tested conditions (all except transcriptional inhibition for 4 hr), the first biological replicate is a compilation of three technical replicates generated from the same batch of harvested cells. Biological replicates were generated by harvesting (culturing, crosslinking, aliquoting, and snap-freezing) 125-200M cells for each tested condition, after which downstream RCMC steps (Micro-C and Capture) were applied to five snap-frozen 5M cell aliquots to generate 1-3 technical replicates for each biological replicate. Reaction volumes for performing RCMC on both 1M and 5M cell samples are provided in the Supplementary Protocol.

Cell culture—Mouse embryonic stem cells (JM8.N4 mESCs[56]; Research Resource Identifier: RRID:CVCL_J962; obtained from the KOMP Repository at UC Davis) were cultured at 37°C with 5% CO_2 on plates coated with 0.1% gelatin solution (Sigma-Aldrich #G1890) under feeder-free conditions in medium consisting of KnockOut DMEM (ThermoFisher #10829-018) with 15% FBS (HyClone, SH30396.03, Lot. No. AE28209315) and 1000U/mL LIF (home-made[57]), 1 mM MEM Non-Essential Amino Acid Solution (ThermoFisher #11140-050), 2 mM GlutaMAX (ThermoFisher #35050061), 100 μg/ml Penicillin-Streptomycin (ThermoFisher #15140-122), and 0.1 mM 2-mercapoethanol (ThermoFisher #31350010) supplemented with 2i, 10 μM MEK inhibitor (Tocris #PD0325901) and 3 μM GSK inhibitor (Sigma-Aldrich #SML1046). F1M RAD21-mAID-BFP-V5 JM8.N4 mESCs were previously generated and validated in the lab[38]. mESCs were fed daily by replacing half of the medium and passaged every two days with TrypLE

Express Enzyme (ThermoFisher #12605036). One day prior to treatment and harvesting, cells were swapped to medium as described above without 2i.

HEK293T cells were obtained from ATCC (CRL-3216) and were cultured at 37°C with 5% CO_2 in DMEM supplemented with 10% FBS, 2 mM L-glutamine, 1x penicillin-streptomycin, and 0.5 mM β-mercaptoethanol.

Depletion of cohesin—Depletion of cohesin was achieved using indole-3-acetic acid (IAA) treatment of the cell line clone F1M RAD21-mAID-BFP-V5 as previously described[12,38]. A 250 mM IAA (BioAcademia #30-003-10) stock was prepared by dissolving the drug in DMSO. F1M RAD21-mAID-BFP-V5 JM8.N4 mESCs[38] were grown to ~80% confluency in medium as described above, with a swap to 2i-free medium 24 hours prior to treatment. Cells were washed once with PBS and fed fresh 2i-free medium containing either only DMSO (untreated control) or 500 μM IAA (cohesin depleted), incubated for 3 hours, and then harvested.

Inhibition of transcription—Inhibition of RNA Pol II activity was achieved using triptolide (Trp) treatment as previously described[13]. A 1 mM triptolide (Sigma-Aldrich #T3652) stock was prepared by dissolving the drug in DMSO. Wild-type (WT) JM8.N4 mESCs[56] were grown to ~80% confluency in medium as described above, with a swap to 2i-free medium 24 hours prior to treatment. Cells were washed once with PBS and fed fresh 2i-free medium containing 1 μM triptolide, incubated for 45 min or for 4 hours, and then harvested.

Crosslinking—Cells were doubly crosslinked to fix protein-protein and protein-DNA interactions using DSG (disuccinimidyl glutarate, 7.7Å) (ThermoFisher #20593) and formaldehyde (ThermoFisher #28906), respectively. Crosslinking medium was prepared by diluting freshly made DSG stock solution (300 mM DSG in DMSO) to 3 mM in 1X PBS (ThermoFisher #10010031). Trypsinized cells were resuspended to single cells, counted, washed in PBS, and then resuspended in crosslinking medium at a concentration of 1M cells per mL. The crosslinking reaction was gently mixed at room temperature for 35 minutes, after which formaldehyde was added to a final concentration of 1%. The double crosslinking reaction was mixed at room temperature for an additional 10 minutes before quenching with Tris buffer pH = 7.5 (K-D Medical #RGE-3370) at a final concentration of 0.375 M. Treatments for non-WT samples (1 μM triptolide, 500 μM IAA, or DMSO) were added to all harvesting reagents used prior to Tris quenching (PBS, trypsin, trypsin-quenching media, and crosslinking medium) to avoid post-treatment rescue during the crosslinking process. Crosslinked cells were washed twice with 1X PBS, re-counted to quantify any sample loss during fixation, and then partitioned into 5M cell aliquots that were pelleted and snap-frozen in liquid nitrogen for storage at −80°C.

Micrococcal nuclease (MNase) titration—Digesting the crosslinked genome to the nucleosome-sized fragments (150-200 bp) necessary to capture nucleosome-resolution DNA contacts requires a titration to identify the ideal MNase digestion concentration and reaction conditions. Accordingly, MNase titrations were performed for each batch of crosslinked cells before performing the RCMC protocol. The titration involved MNase digestion of 1M

or 5M cell samples varying MNase concentrations, reversal of crosslinks, DNA purification, and gel-based separation to visualize the distribution of fragment sizes (see corresponding sections below). Ideal digestion concentrations were identified by samples digested to primarily (~80%) mononucleosomal fragments (150-200 bp), few (~15-20%) dinucleosomal fragments (250-350 bp), and a faint but visible band (<5%) of trinucleosomal fragments (400-500 bp) (Extended Data Fig. 1a).

Micrococcal nuclease digestion—Cell membranes were solubilized to extract intact nuclei by resuspending crosslinked 5M cell pellets in Micro-C Buffer #1 (MB#1; 50 mM NaCl, 10 mM Tris-HCl pH = 7.5, 5 mM MgCl2, 1M CaCl2, 0.2% NP-40 Alternative (Millipore Sigma #492018), 1x Protease Inhibitor Cocktail (Sigma-Aldrich #5056489001)) at 1M cells per 100 μL for 20 min on ice. Following an MB#1 wash, samples were resuspended in 100 μL MB#1 and the ideal amount of 20 U/μL MNase (Worthington Biochem #LS004798) determined by the MNase titration was added. This digestion reaction was mixed at 37°C for 20 min on a thermomixer before being quenched with 4 mM EGTA (bioWORLD #40520008) and heat inactivated at 65°C for 10 min. Digested nuclei were washed twice with ice-cold Micro-C Buffer #2 (50 mM NaCl, 10 mM Tris-HCl pH = 7.5, 10 mM MgCl2, 100 μg/mL BSA (Sigma-Aldrich #B8667)).

End repair and labeling—To generate blunt ends on digested DNA fragments prior to proximity ligation and add biotinylated nucleotides, a series of enzymatic processing steps were performed. First, to catalyze the addition of 5'-phosphate groups and the removal of 3'-phosphate groups, digested samples generated from 5M cell inputs were incubated in end-repair reactions (50 U T4 Polynucleotide Kinase (New England BioLabs #M0201), 50 mM NaCl, 10 mM Tris-HCl pH = 7.5, 10 mM MgCl2, 100 ug/mL BSA, 2 mM ATP (ThermoFisher #R1441), 5 mM DTT (Sigma-Aldrich #10197777001), in water) at 37°C for 15 min while mixing. To create 5' fragment overhangs for end-blunting and labelling, 50 U of DNA Polymerase I Klenow Fragment (New England BioLabs #M0210) was added to the reaction and incubated at 37°C for 15 min while mixing. Next, a mixture of dNTPs in end-labelling buffer (66 μM each of dTTP (Jena Bioscience #NU-1004), dGTP (Jena Bioscience #NU-1003), biotin-dATP (Jena Bioscience #NU-835-BIO14), and biotin-CTP (Jena Bioscience #NU-809-BIOX), 1X T4 DNA Ligase Buffer, 100 μg/mL BSA, in water) was added to the reaction. This reaction was incubated at room temperature for 45 min with interval mixing before being quenched by 30 mM EDTA (Invitrogen #15575020) and heat inactivated at 65°C for 20 min. Finally, end-blunted and biotin-labeled nuclei were washed once with Micro-C Buffer #3 (50 mM Tris-HCl pH = 7.5, 10 mM MgCl2, 100 μg/mL BSA).

Proximity ligation and removal of unligated biotin—Proximity ligation was performed by incubating labeled chromatin in a ligation reaction (10,000 U T4 DNA Ligase (New England BioLabs #M0202), 1X T4 DNA Ligase Buffer, 100 μg/mL BSA, in 500 μL water) at room temperature for at least 2.5 hours with gentle mixing. To remove biotinylated dNTPs from all unligated fragment ends, samples were digested by 1,000 U of Exonuclease III (New England BioLabs #M0206) in reaction buffer (1X NEBuffer #1 in water) at 37°C for 15 min with interval mixing.

DNA purification and size-selection—In order to prepare ligated DNA for library generation, DNA was reverse crosslinked and proteins and RNA were digested by adding 1% SDS (Sigma-Aldrich #L3771), 2 mg/mL Proteinase K (Viagen Biotech #501-PK), 250 mM NaCl, and 100 μg/mL RNaseA (ThermoFisher #EN0531) to the samples and incubating at 65°C overnight. DNA was extracted using phenol:chloroform:isoamylic alcohol (PCI) (Sigma-Aldrich #P2069) in a 1:1 volumetric ratio using 5PRIME Phase Lock Gel Light tubes (Quantabio #2302820). The aqueous phase was further purified using the Zymo DNA Clean & Concentrator kit (Zymo Research #D4034) according to the kit manual.

Dinucleosome-sized DNA fragments (250-350 bp) were isolated by extraction from a 1% agarose gel (VWR #97062) (Extended Data Fig. 1b). Gel extracts were purified using the Zymo Gel Purification kit (Zymo Research #D4008), and samples were quantified by Qubit 1X dsDNA High Sensitivity Assay (Invitrogen #Q33231). Sample ends were polished and blunted again using the End-It enzyme reaction (Lucigen #ER81050) at 25°C for 45 minutes, followed by reaction inactivation at 65°C for 10 min.

Ligated DNA contact fragments were isolated by pulling down biotin-bound fragments using Dynabeads MyOne Streptavidin T1 (Invitrogen #65601). DNA samples were bound to beads in a Binding and Wash Buffer (1 M NaCl, 5 mM Tris-HCl pH = 7.5, 500 μM EDTA, 0.1% Tween-20 (Sigma-Aldrich #P8074)) at room temperature for at least 30 minutes with mixing. After two washes with the Binding and Wash Buffer, the bead-bound samples were washed once with 10 mM Tris-HCl pH = 7.5 prior to library prep.

Library preparation—Illumina library preparation was performed using the NEBNext Ultra II kit (New England BioLabs #E7645) to end-repair, A-tail, and adaptor ligate the bead-bound samples. All steps were performed as directed by the manual, except that incubations included interval shaking (1 minute on, 3 minutes off) at 1000 rpm. Sample washes were performed using Binding and Wash Buffer and 10 mM Tris-HCl pH = 7.5 washes. To determine the minimum number of PCR cycles to meet input material guidelines for capture or sequencing, a test library amplification was performed with 5% or less of the prepped library to quantify the yield. The test PCR reaction mixture was run on an agarose gel and yield quantified using image quantification software Image Studio Lite (LI-COR Biosciences). 10 or fewer PCR cycles to meet Capture input requirements is optimal to reduce PCR duplicates, and the RCMC replicates in this manuscript used 7-8 PCR cycles for final library amplification. All library amplifications were done using sequencing indices from the NEB Multiplex Oligos for Illumina Primer Set 1 (New England BioLabs #E7335) and the KAPA HiFi HotStart ReadyMix enzyme (Roche #07958927001). Following library amplification, the T1 Dynabeads containing the original bead-bound samples were removed and the amplified libraries were purified to remove adaptor dimers, primers, and contaminants using AmPure XP beads (Beckman Coulter #A63880). Purified libraries were quantified via Fragment Analyzer and qPCR at the MIT BioMicro Center to determine library concentrations for pooling prior to Capture.

Capture probe design—Target loci of interest were identified based on genomic features or enhancer-promoter relationships of interest. *Klf1* and *Ppm1g* were selected as gene-rich loci, *Fbn2* was selected as a gene-poor control with a well-established CTCF- and

cohesin-mediated loop[38], and *Sox2* and *Nanog* were later selected as loci for comparing RCMC against TMCC (Extended Data Fig. 3). Using the UCSC Genome Browser and HiGlass visualization of existing mESC 3C datasets, locus bounds were selected to include visible local structures and genomic features in roughly 1 Mb-sized regions. Once loci had been selected, 80-mer probes were designed to tile end-to-end without overlap across the Capture loci through Twist Bioscience (Extended Data Fig. 1c). Probes with high predicted likelihoods of off-target pulldown (e.g., such as those in high-repeat regions) were masked and removed from the probe tiling, and probe coverage was double-checked to ensure the inclusion of key genomic features (e.g., all promoters and CTCF sites in the locus) before finalization. Probe panels were synthesized and purchased as Custom Target Enrichment Panels from Twist Bioscience.

Capture of target loci—Capture was performed in accordance with Twist Bioscience's Standard Hybridization Target Enrichment Protocol. Briefly, pooled sample libraries were dried and mixed with Hybridization Mix (Twist Bioscience #104178), Custom Panels (Twist Bioscience #101001), and Universal Blockers (Twist Bioscience #100578), as well as Mouse Cot-1 DNA (Invitrogen #18440016). The library pool was hybridized to the biotinylated probe panel overnight, after which streptavidin beads (Twist Bioscience #100983) were used to pull down probes with hybridized ligated fragments and then washed (Twist Bioscience #104178) to remove unbound fragments. Another round of PCR amplified the target-enriched library using the Equinox Library Amplification Mix (Twist Bioscience #104178), with including a test PCR (as described above) to identify the number of amplification cycles necessary to meet sequencing requirements. With 2-4 μg of input library for Capture, the RCMC samples generated in this manuscript needed 5-6 cycles of post-Capture PCR amplification. Following PCR amplification, the Captured library was purified (Twist Bioscience #100983) and then quantified via both Fragment Analyzer and qPCR at the MIT BioMicro Center in preparation for sequencing submission.

Three technical replicates of the pre-Capture Micro-C library were generated for each of the four initially tested conditions (wild-type, 45 min transcriptional inhibition with Trp, RAD21 depletion, and a RAD21 depletion control), after which each replicate was simultaneously Captured for the *Klf1*, *Ppm1g*, and *Fbn2* loci. After the publication of TMCC[17], additional probes for the *Sox2* and *Nanog* loci were designed and a single additional Capture experiment was conducted pooling all three pre-Capture Micro-C libraries for simultaneous *Sox2* and *Nanog* Capture. Subsequently, a biological replicate was generated for each of the initially tested conditions, with the inclusion of the additional perturbation of 4 hr transcriptional inhibition with Triptolide. Pre-Capture Micro-C libraries were constructed for each of the five conditions, after which each library was simultaneously Captured for all five target loci. Finally, a biological replicate of the 4 hr transcriptional inhibition perturbation was generated; once a pre-Capture Micro-C library for it was generated, it was pooled with the WT library from the first technical replicate of the first biological replicate, and the pooled libraries were simultaneously Captured for all five target loci.

Sequencing—Following qPCR quantification, post-Capture libraries across samples (wild-type, transcriptionally inhibited, cohesin depleted, and DMSO-treated control) were pooled in a 1:1 molar ratio. Pooled libraries were sequenced by paired-end 2×50 cycle sequencing kits with Illumina NovaSeq SP or S1 flow cells on a NovaSeq 6000 system by the Broad Institute of MIT and Harvard's Walk-Up Sequencing services. Basecalls for NovaSeq output were performed using bcl2fastq v2.20.0.422.

Data Analysis

Mapping and normalizing RCMC—RCMC paired-end reads generated by the Illumina NovaSeq sequencers were downloaded as .fastq files for each sample, pair mate, and flow cell lane. Read quality was verified using FastQC (v0.11.9). Paired end reads were aligned to the UCSC mm39 genome using bowtie2 (v2.3.5.1) with --local --reorder --very-sensitive-local. Aligned paired end reads were then parsed with pairtools (v0.3.0) parse with --add-columns mapq --walks-policy mask --min-mapq 2. Parsed reads were filtered for PCR duplicates and unmapped/mutiple mapping reads with pairtools dedup with --max-mismatch 1. Remaining reads were indexed (pairix v0.3.7) and filtered (pairtools select) to retain only those reads where both read mates lie in a locus of interest. These filtered reads were subsequently converted to .cool format using cooler (v0.8.11) cload pairs, creating binned read counts across the genome for 50 bp bins. Finally, .cool files were converted to the .mcool format with cooler zoomify including the --balance option, compiling read counts for bins from 50 bp up to 10 Mb in size.

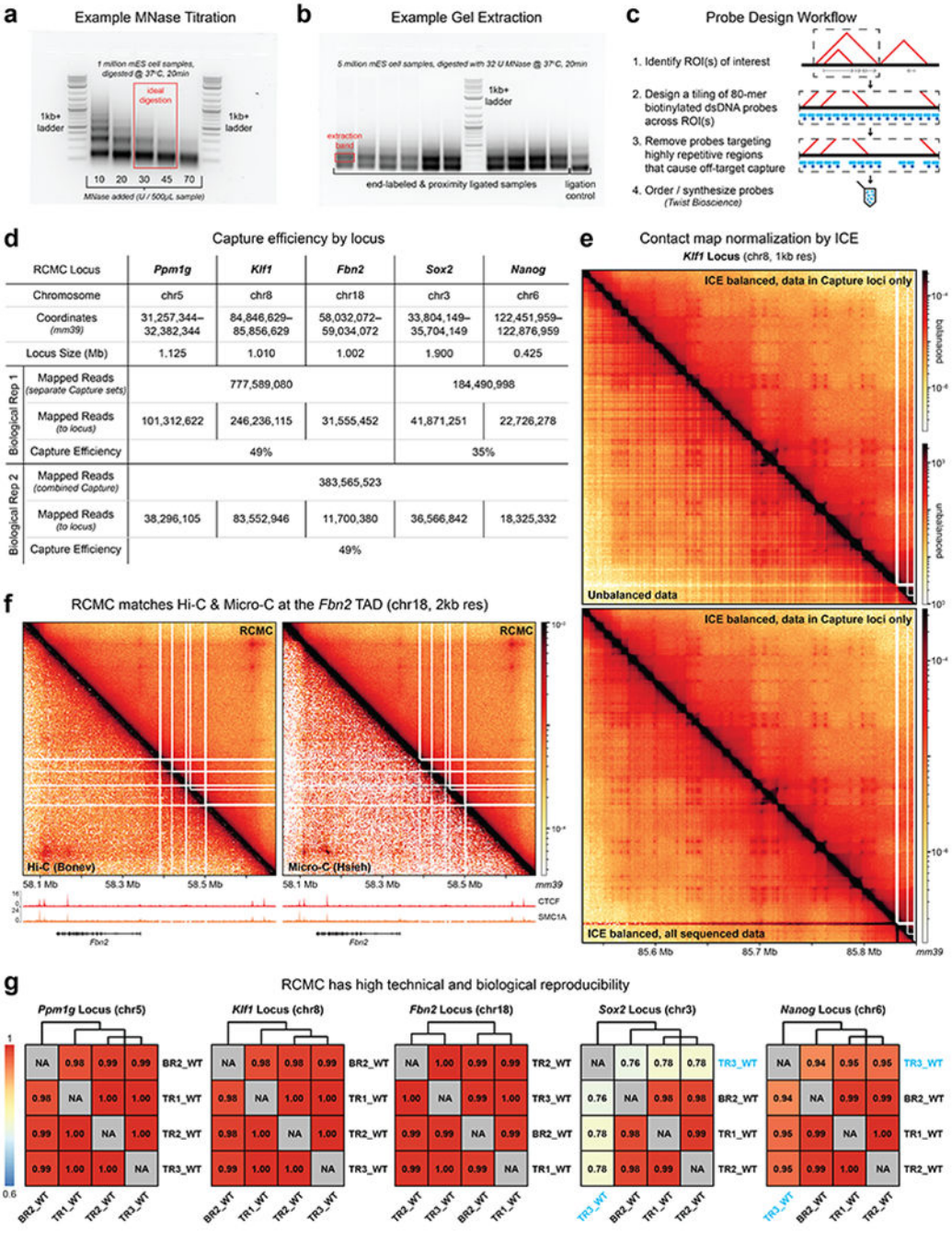
Contact matrices were balanced using iterative correction and eigendecomposition (ICE)[30], which normalizes all rows and columns of a contact matrix sum to the same value. Applying ICE balancing to all mapped reads generated subpar normalization and generated an artifact where "stripes" containing no Capture probe coverage appeared to have greater contact densities than adjacent probe-covered regions (Extended Data Fig. 1e). ICE balancing to .mcool files containing data only within Captured regions of interest (ROIs) did not result in these artifacts, and was therefore used in for all RCMC data in this study. The success of ICE balancing applied to these ROI-only .mcools was evaluated against published whole-genome Hi-C[31] and Micro-C[13] datasets in mESCs (Extended Data Fig. 6a). The sum of each row of each of the RCMC, Micro-C, and Hi-C balanced contact matrices at 250 bp resolution within Capture ROIs was calculated, plotted as a histogram distribution of row sums, and verified to match the distribution of column sums. The subset of RCMC rows containing microcompartment anchors was also plotted to confirm that they match the distribution of row sums across the whole locus, ruling out that microcompartments are an artifact of incomplete ICE-normalization[30].

Visualizing RCMC—RCMC contact maps were visualized alongside genomic annotations, published ChIP-seq, RNA-seq, and ATAC-seq datasets using the HiGlass[58] browser (http://higlass.io/) and software (v0.8.0). Contact maps shown in figures were generated using cooltools (v0.5.0) (https://cooltools.readthedocs.io/). Genomic tracks (i.e., ChIP-seq, RNA-seq, and ATAC-seq) and gene annotations for manuscript figures were generated using CoolBox[59] (v0.3.3). In generating our genomic tracks, we analyzed 27 public datasets (Supplementary Table 1) using processed bigWig files which were

CrossMapped[60] (v0.6.1) (http://crossmap.sourceforge.net/) to the mm39 reference genome. Tracks were visualized using the Integrative Genomics Viewer (IGV)[61] (v2.10.3) to scale tracks by identifying local maxima and minimizing noise.

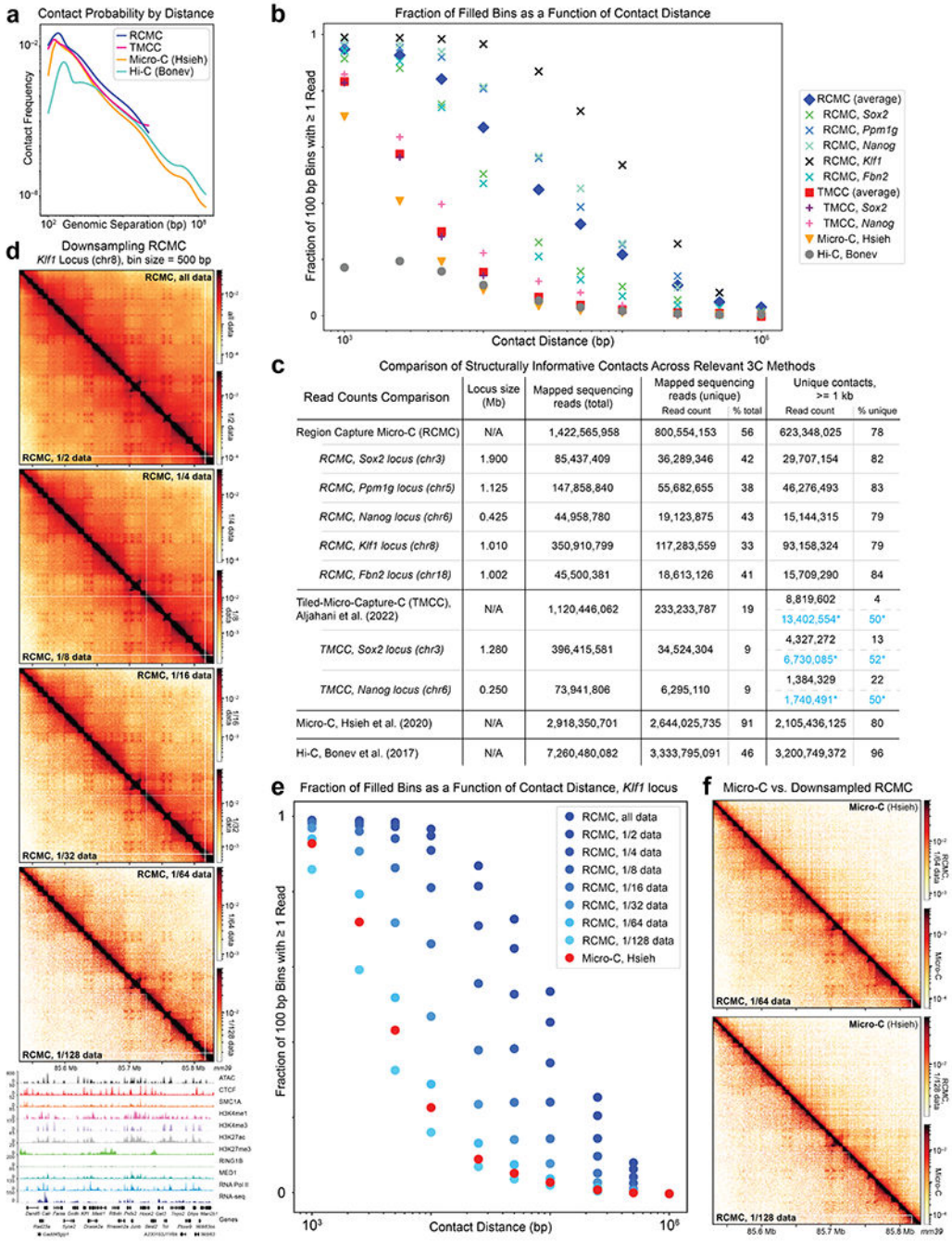
All other bioinformatic and data analyses are provided as Supplementary Methods.

Extended Data



Extended Data Figure 1. RCMC efficiently and reproducibly captures ligated dinucleosomal fragments, giving rise to deep contact maps.

(**a**) Representative MNase titration DNA gel indicating the ideal level of digestion by MNase, based on the ratio of fragment sizes, for the RCMC protocol. (**b**) Representative size selection gel for the RCMC protocol showing the dinucleosomal band that is extracted to obtain ligated fragments. (**c**) Overview of the capture probe design workflow for RCMC. 80-mer probes tiling the region of interest are designed, removing those which overlap highly repetitive regions. (**d**) Summary of the capture efficiency for each of the five regions for which probes were designed. The locations and sizes of the regions, the number of ligated fragments which mapped at single loci at both ends in total and in the region, and the capture efficiencies are given. Because different capture probe sets were used for Biological Replicates 1 (two separate sets of capture probes) and 2 (simultaneous capture for all five loci), numbers are separately provided for each Biological Replicate. (**e**) Contact maps comparing raw, unbalanced data (upper panel, lower triangle), ICE-balanced[30] to all aligned reads (lower panel, lower triangle) and ICE-balanced to reads in Capture loci only (both panels, upper triangle). Balancing only to data entirely within Capture loci was necessary to remove artifacts due to capture bias. (**f**) Contact maps comparing the entire *Fbn2* TAD in RCMC and in Hi-C[31] and Micro-C[13]. Gene annotations and ChIP-seq signal tracks are shown below the contact maps. (**g**) Measurement of reproducibility between wild-type replicates across all five Capture loci, with reproducibility scores determined using HiCRep[52] at 10 kb resolution, clustered according to similarity. Three technical RCMC replicates (denoted by “TR#”) comprise Biological Replicate 1, while “BR2” denotes Biological Replicate 2. TR3_WT is noted in blue text at the *Sox2* and *Nanog* loci because very little TR3_WT pre-Capture library remained for input to *Sox2* & *Nanog* Capture after the initial *Ppm1g*, *Klf1* and *Fbn2* Capture experiment; accordingly, relative to all other replicates, TR3_WT has much lower sequencing depth (0.5-2.4% the number of unique contacts) at the *Sox2* & *Nanog* loci.

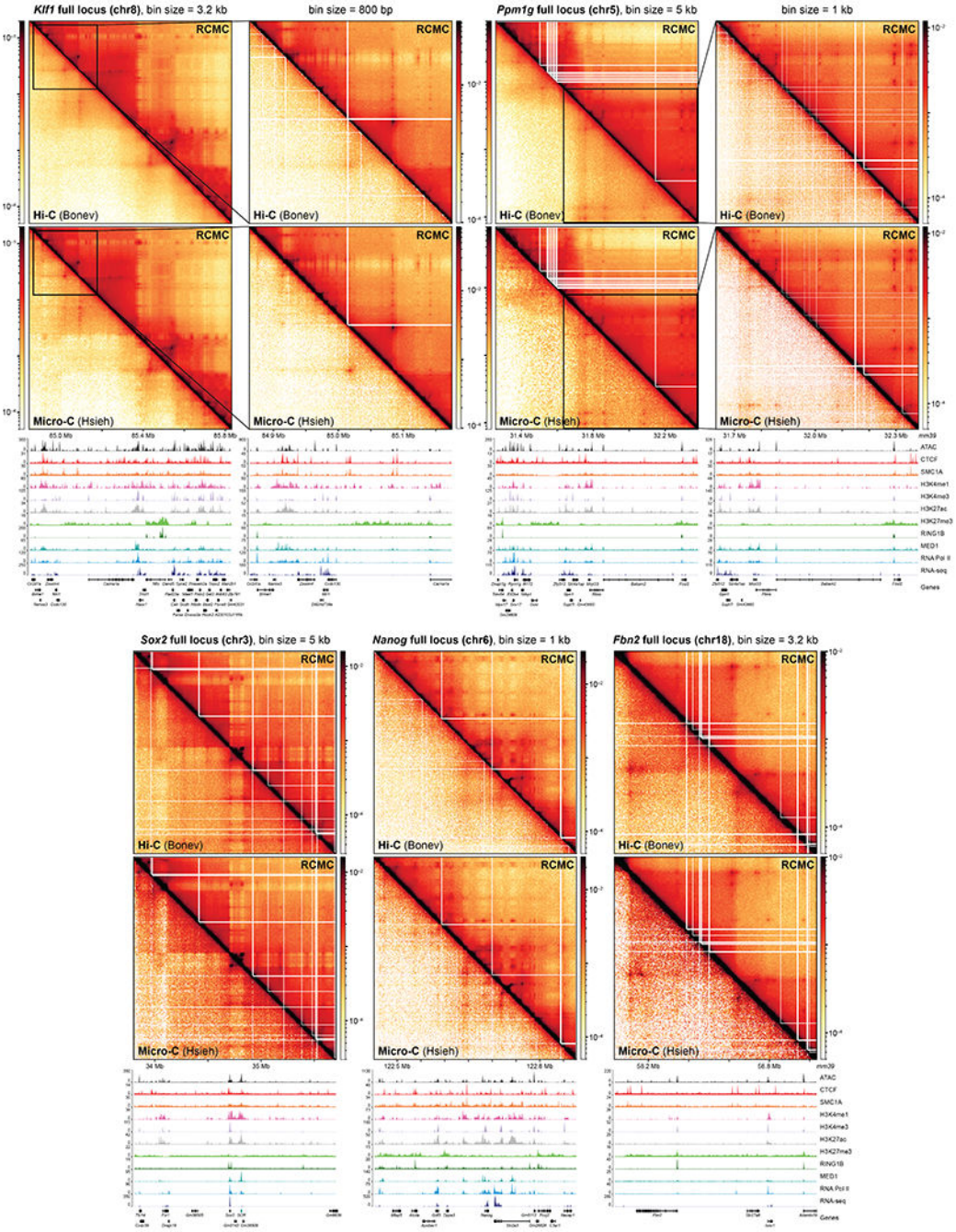


Read Counts Comparison	Locus size (Mb)	Mapped sequencing reads (total)	Mapped sequencing reads (unique)		Unique contacts, >= 1 kb	
			Read count	% total	Read count	% unique
Region Capture Micro-C (RCMC)	N/A	1,422,565,958	800,554,153	56	623,348,025	78
RCMC, Sox2 locus (chr3)	1.900	85,437,409	36,289,346	42	29,707,154	82
RCMC, Ppm1g locus (chr5)	1.125	147,858,840	55,682,655	38	46,276,493	83
RCMC, Nanog locus (chr6)	0.425	44,958,780	19,123,875	43	15,144,315	79
RCMC, Klf1 locus (chr8)	1.010	350,910,799	117,283,559	33	93,158,324	79
RCMC, Fbn2 locus (chr18)	1.002	45,500,381	18,613,126	41	15,709,290	84
Tiled-Micro-Capture-C (TMCC), Aljahani et al. (2022)	N/A	1,120,446,062	233,233,787	19	8,819,602	4
					13,402,554*	50*
TMCC, Sox2 locus (chr3)	1.280	396,415,581	34,524,304	9	4,327,272	13
					6,730,085*	52*
TMCC, Nanog locus (chr6)	0.250	73,941,806	6,295,110	9	1,384,329	22
					1,740,491*	50*
Micro-C, Hsieh et al. (2020)	N/A	2,918,350,701	2,644,025,735	91	2,105,436,125	80
Hi-C, Bonev et al. (2017)	N/A	7,260,480,082	3,333,795,091	46	3,200,749,372	96

Extended Data Figure 2. Benchmarking of RCMC against other 3C methods.

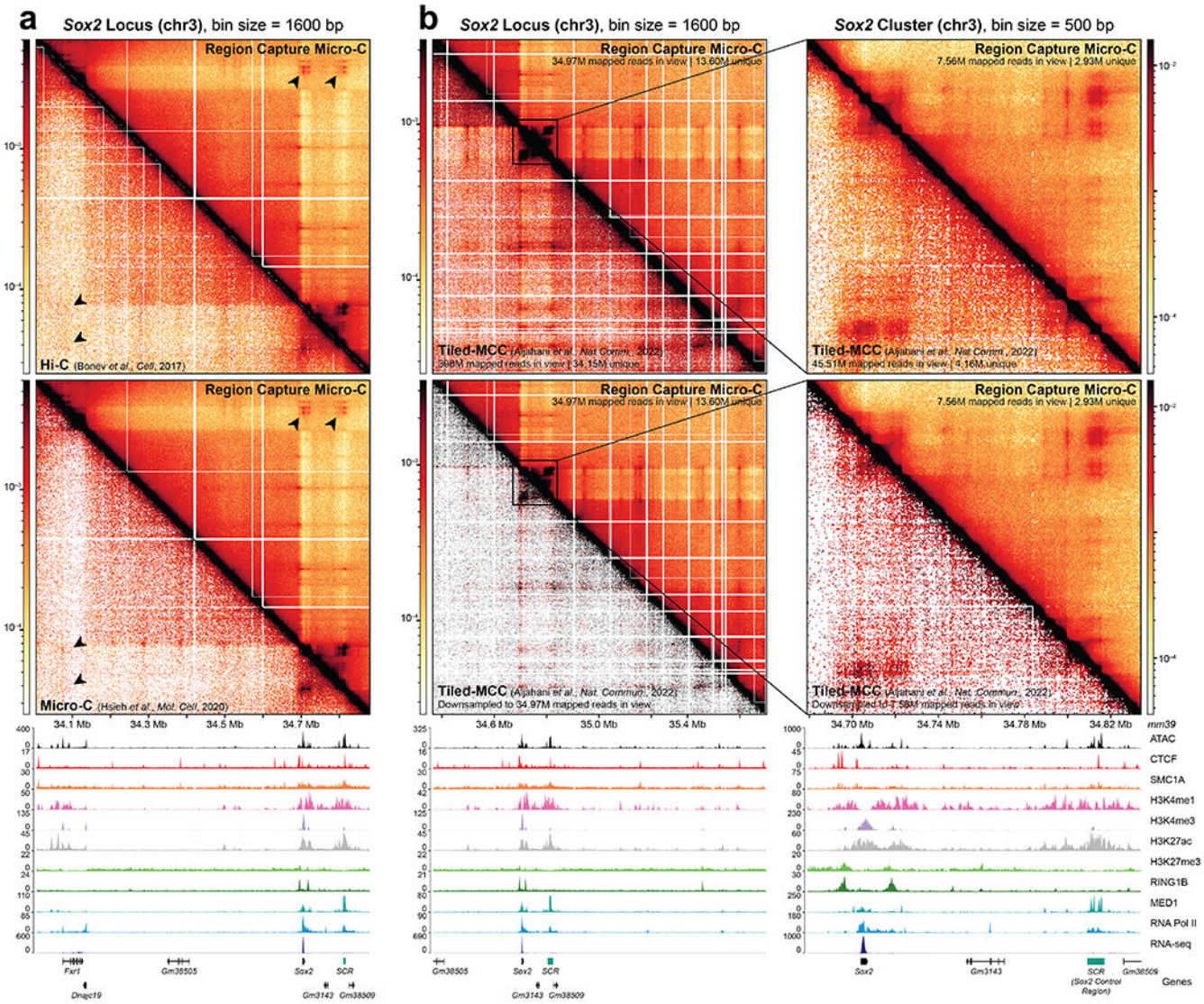
(**a**) Contact probability curves comparing RCMC against the highest resolution Tiled-Micro-Capture-C (TMCC)[17], Micro-C[13], and Hi-C[31] mESC datasets across contact distances. (**b**) Benchmarking comparison of RCMC's ability to fill out high-resolution contact matrices against TMCC[17], Micro-C[13], and Hi-C[31]. Region-averaged calculations are shown for all methods, and calculations for individual Captured regions are also shown for RCMC and TMCC. The x-axis shows the contact distance in bp, and the y-axis shows the fraction of all bins at a given contact distance within the Captured locus that contain at least

one read at 100 bp resolution. (**c**) Summary of read counts across RCMC, TMCC[17], Micro-C[13], and Hi-C[31]. The number of mapped sequencing reads, the fraction of unique reads, and the fraction of structurally informative (defined as *cis* contacts >=1 kb) unique reads are given for each method. Two versions of quantification are provided for TMCC. In black are numbers processed using the same bioinformatic pipeline as for RCMC. Capture region-specific quantifications (defined here as all reads with at least one of two read mates mapped to the locus) are also provided for all RCMC loci and the *Sox2* and *Nanog* TMCC loci; the *Oct4* and *Prdm14* TMCC loci are not considered in this manuscript. In blue are numbers kindly provided by Dr. A Marieke Oudelaar, obtained using the custom TMCC-specific bioinformatic pipeline from Aljahani *et al.*[17]. Values with asterisks denote quantifications of all unique contact pairs mapped to Captured loci (not filtered to be >= 1 kb in size). (**d**) Contact map comparisons of RCMC data generated in this manuscript, starting from the full dataset (topmost) and successively downsampled by orders of two down to 1/128th of the data (bottommost), shown for the *Klf1* locus at 500 bp resolution. (**e**) As in (b), benchmarking comparison of successively downsampled RCMC's ability to fill out high-resolution contact matrices against Micro-C[13] at the *Klf1* locus. (**f**) Contact map comparisons of 1/64th and 1/128th downsampled RCMC (left) against the highest-resolution available mESC Micro-C[13] (right; Hsieh 2020) dataset, shown for the *Klf1* locus at 500 bp resolution.



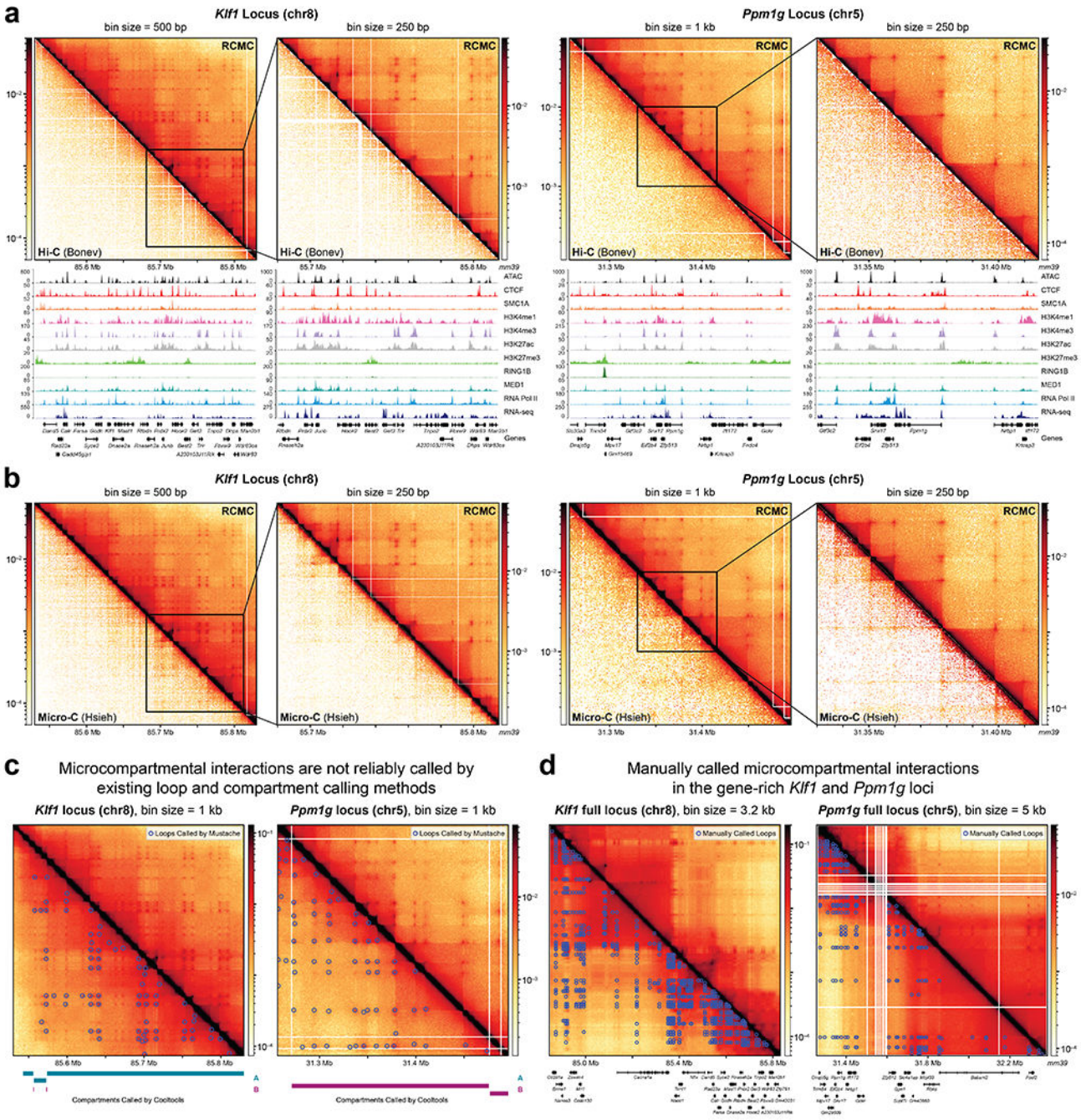
Extended Data Figure 3. RCMC generates deeper contact maps than other 3C methods across all 5 Captured loci.

Contact map comparisons of RCMC against the highest-resolution available mESC Hi-C[31] (top; Bonev 2017) and Micro-C[13] (bottom; Hsieh 2020) datasets at the *Klf1*, *Ppm1g*, *Sox2*, *Nanog*, and *Fbn2* loci. Full Capture regions are shown for each locus at resolutions ranging from 1-5 kb, as well as *Klf1* and *Ppm1g* zoom-ins at 800 and 1000 bp, respectively. Gene annotations and ATAC, ChIP-seq, and RNA-seq tracks (Supplementary Table 1) are shown below the contact maps, while the contact intensity scales are shown next to the maps.



Extended Data Figure 4. RCMC maps the *Sox2* locus more deeply and efficiently than sister methods, uncovering previously unseen interactions.

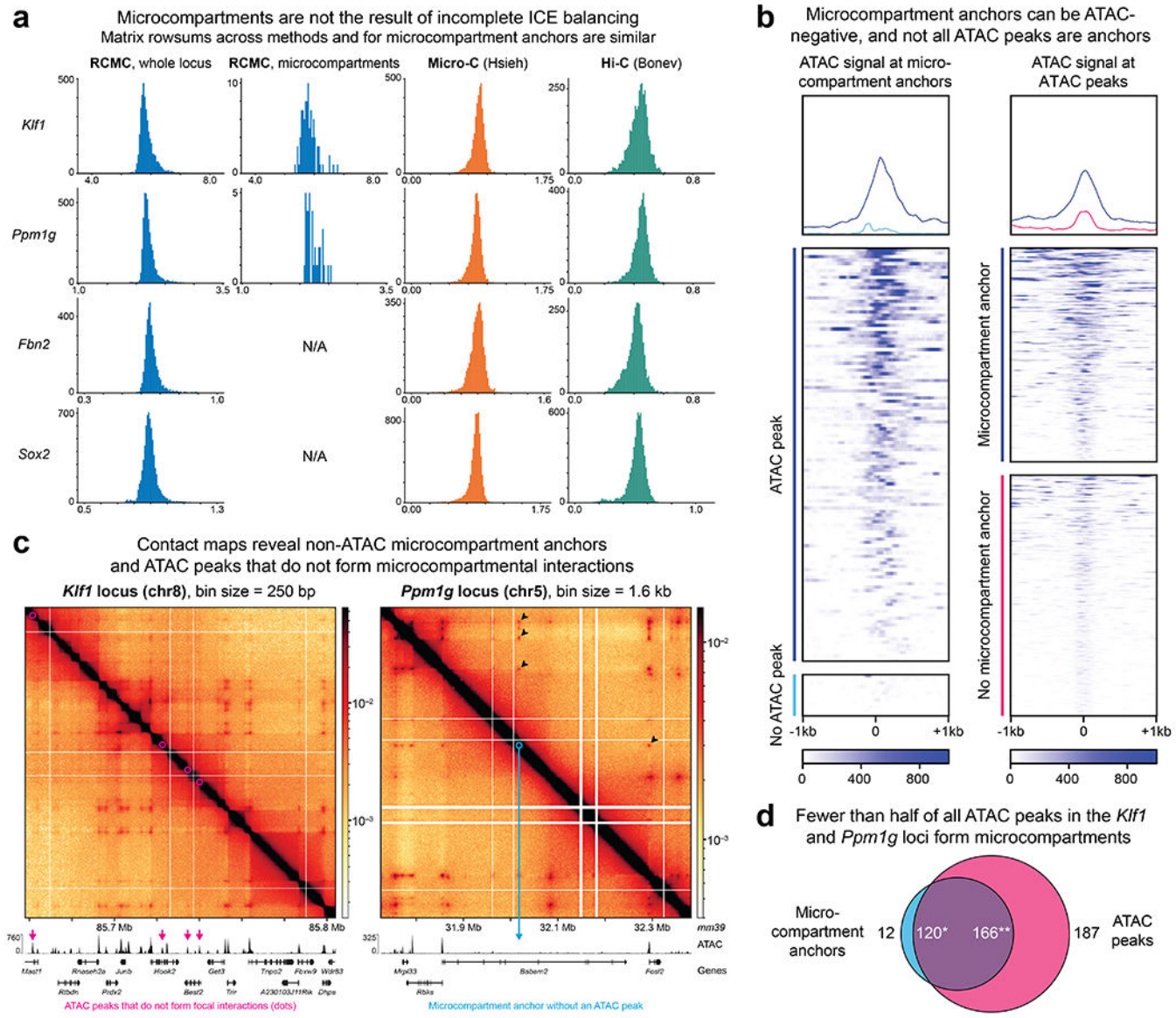
(**a**) Contact map comparisons of RCMC against Hi-C[31] (top) and Micro-C[13] (bottom) at the *Sox2* locus at 1.6 kb resolution. Arrows mark contacts between *Sox2*, the SCR, and *Fxr1* not mapped by Hi-C and Micro-C. (**b**) Contact map comparisons of RCMC against Tiled-Micro-Capture-C[17] (TMCC) across the whole TMCC-Captured locus (left, 1600 bp resolution) and in the *Sox2* and SCR regulatory cluster (right, 500 bp resolution). Full datasets are visualized in the top contact maps, and TMCC has been downsampled to match the total number of RCMC sequencing reads in view in the bottom contact maps.



Extended Data Figure 5. RCMC identifies microcompartments, which are not visible in other methods and not reliably called by existing algorithms.

(**a-b**) Contact maps comparison of RCMC (top) against Hi-C[31] (bottom, **a**) and Micro-C[13] (bottom, **b**) at the *Klf1* locus at 500 and 250 bp resolutions and at the *Ppm1g* locus at 1000 and 250 bp resolutions. (**c**) Contact maps of the *Klf1* and *Ppm1g* loci at 1 kb resolution with loop calls by Mustache[37] overlaid on the bottom half of the map and compartment calls by cooltools[53,54] shown below the map. (**d**) Contact maps of the entire *Klf1* (3.2 kb

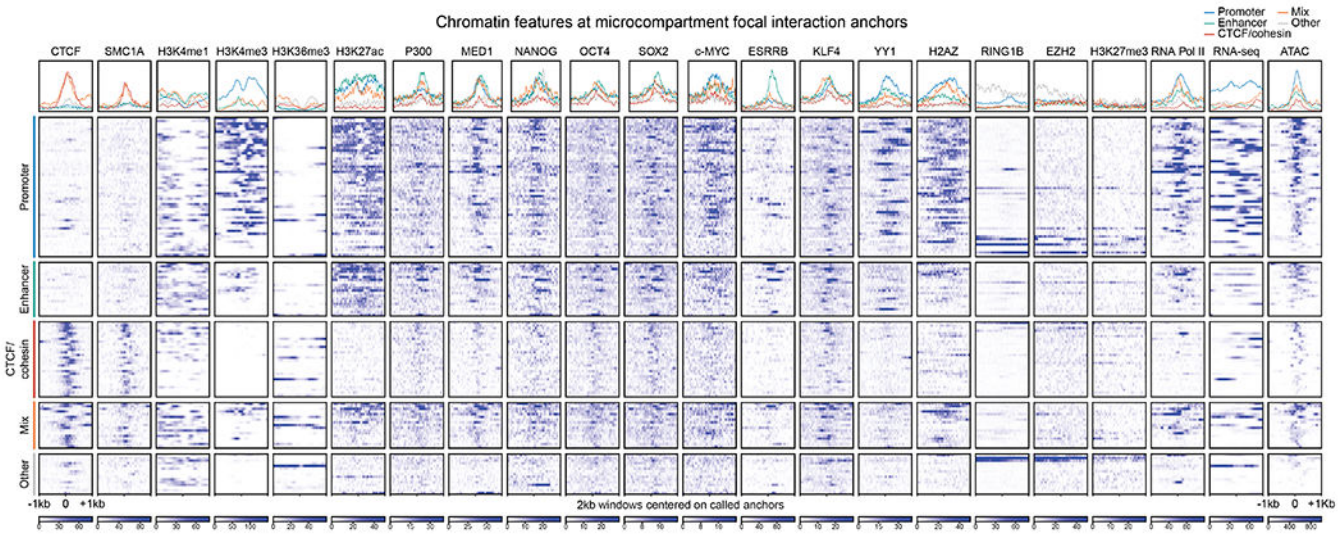
resolution) and *Ppm1g* (5 kb resolution) Captured loci with manually called loops (see Methods) overlaid on the bottom halves of the maps.



Extended Data Figure 6. Microcompartments are not artifacts resulting from incomplete ICE balancing nor chromatin accessibility.

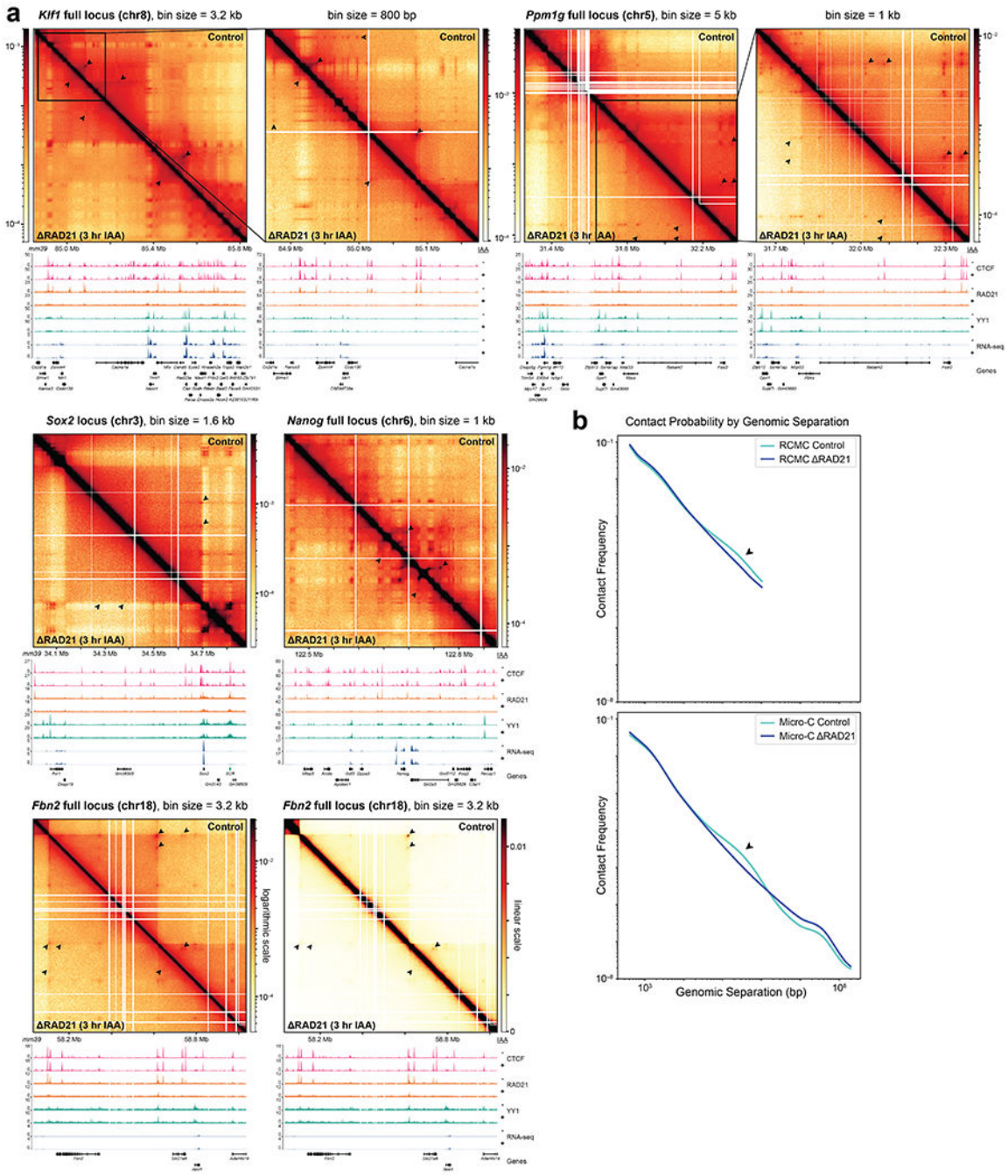
(**a**) Comparison of ICE balancing across methods and Capture loci. Distributions of the sums of ICE-balanced contact matrix rows at 250 bp resolution are shown at the *Klf1*, *Ppm1g*, *Fbn2*, and *Sox2* loci for RCMC, Micro-C[13], and Hi-C[31], as well as for the subset of RCMC rows containing microcompartment anchors. A sharp unimodal peak is consistent with ICE's baseline assumption that all contact matrix rows and columns must sum to the same value. (**b**) Metaplots (above) and heatmaps (below) depicting ATAC signal at microcompartment anchors (left, separated by whether anchors coincide with an ATAC peak) and at all ATAC peaks in the *Klf1* and *Ppm1g* Capture loci (right, separated by whether peaks coincide

with a microcompartment anchor). Signals are plotted in a 2 kb window centered on the anchor (left) or the ATAC peak (right). (**c**) RCMC contact maps at the *Klf1* (left, 250 bp resolution) and *Ppm1g* (right, 1.6 kb resolution) indicating ATAC peaks that do not form microcompartments (left, magenta arrows) and a microcompartment anchor that does not coincide with an ATAC peak (right, cyan arrow). Black arrows (right) indicate microcompartmental loops involving the ATAC-negative microcompartment anchor. (**d**) Venn diagram breakdown of the overlap between all manually annotated microcompartment anchors and all ATAC peaks across the *Klf1* and *Ppm1g* Capture loci. Of 132 annotated microcompartment anchors, 12 do not coincide with ATAC peaks (cyan) while 120 do (purple, *). Of 353 called ATAC peaks, 187 do not form microcompartment anchors (magenta) while 166 do (purple, **). The apparent discrepancy of 120 microcompartment anchors being anchored by 166 ATAC peaks is due to two close ATAC peaks occasionally anchoring a single microcompartment.



Extended Data Figure 7. Categories of microcompartment anchors can be defined by their chromatin features.

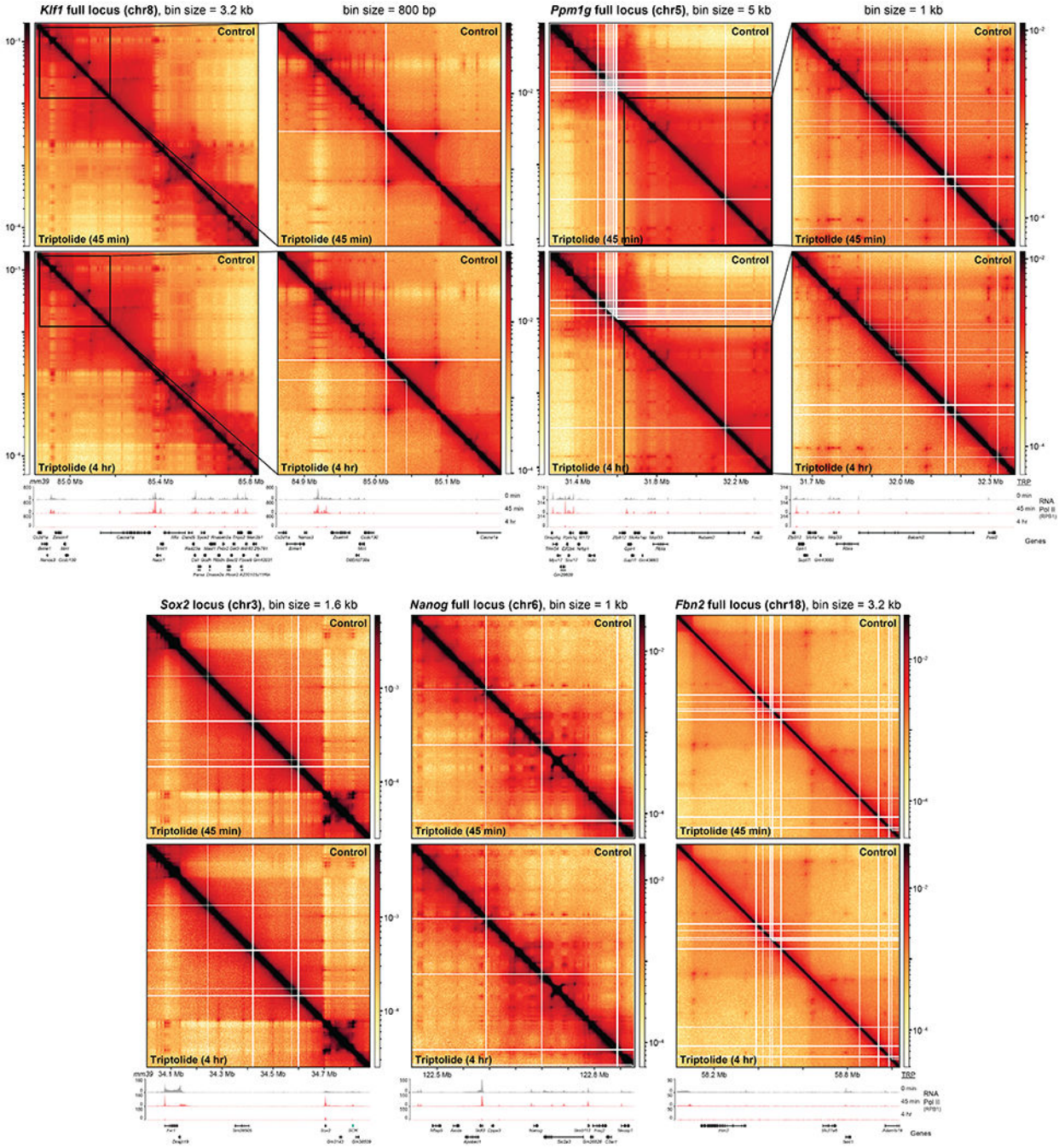
Metaplots (above) and heatmaps (below) depicting ATAC, ChIP-seq, and RNA-seq (Supplementary Table 1) signal at microcompartment loop anchors for classes of microcompartment anchors as defined in Fig. 3e. Features are plotted in a 2 kb window centered on the anchor.



Extended Data Figure 8. Cohesin depletion disrupts CTCF/Cohesin loops, but generally not most microcompartmental loops.

(**a**) Contact maps comparing a DMSO control (above) and RAD21-depleted samples (below) are shown for the *Klf1*, *Ppm1g*, *Sox2*, *Nanog*, and *Fbn2* loci at resolutions spanning 800 bp – 5 kb in F1M RAD21-mAID-BFP-V5 mESCs[12,38]. Arrows mark contacts lost upon RAD21 depletion. ChIP-seq data from Hsieh *et al.*, *bioRxiv* (2021)[12] is shown below the maps before and after the IAA treatment (500 μM, 3 hours). Two versions of the *Fbn2* locus are shown, with the left using logarithmic contact frequency scaling and the right using

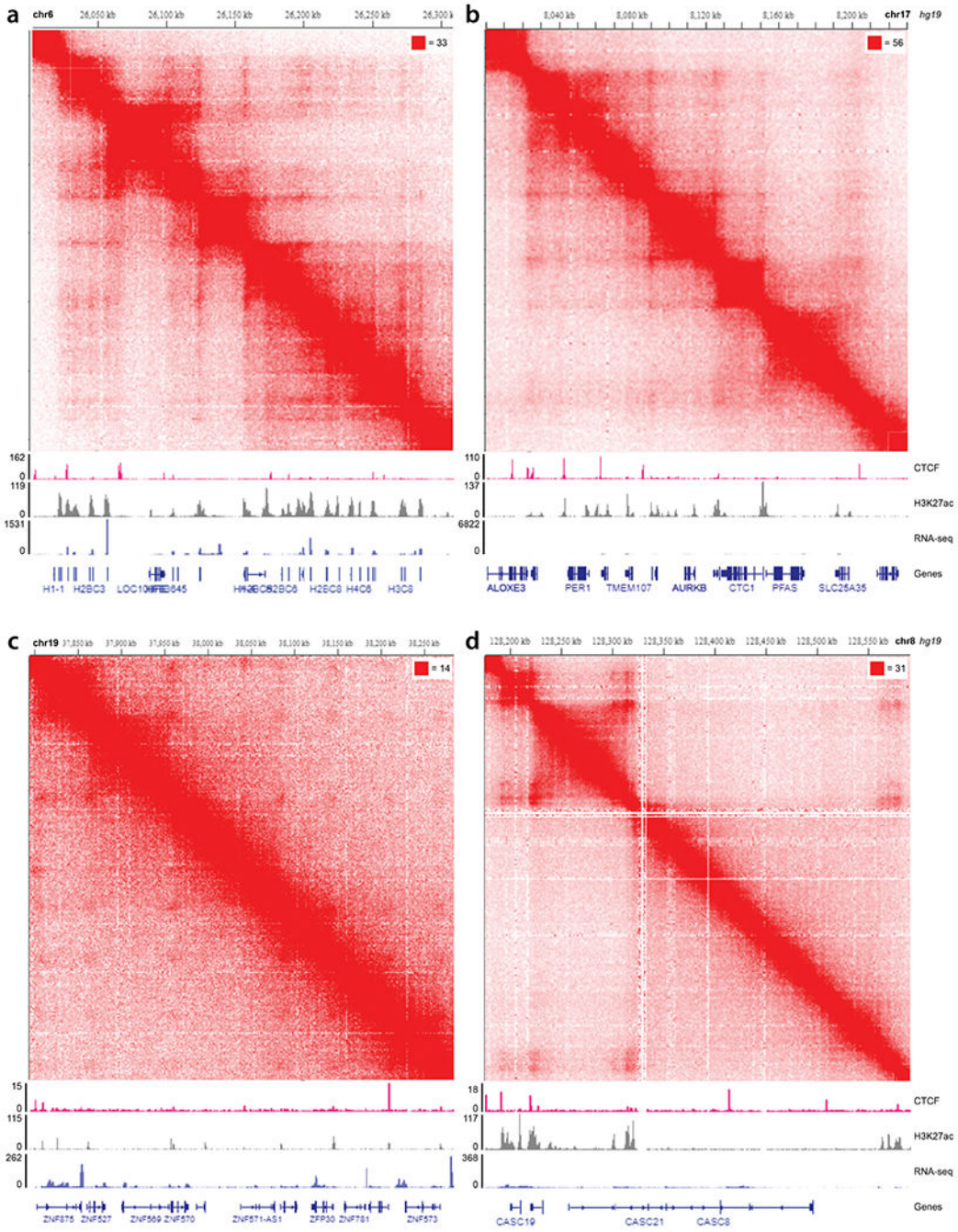
linear scaling. Loss of the *Fbn2* loop[38] is most clearly seen on linear scale. (**b**) Contact probability curves comparing RAD21-depleted RCMC samples against a DMSO control (top) RAD21-depleted Micro-C samples against a DMSO control (bottom). Arrows indicate the contact frequency “bump” lost upon RAD21 depletion.



Extended Data Figure 9. Inhibition of transcription does not significantly alter genome organization in Captured loci.

Contact maps comparing control data against 45 min (top) and 4 hr (bottom) transcriptional inhibition data (from 1 μM triptolide treatments) are shown for the *Klf1*, *Ppm1g*, *Sox2*,

Nanog, and *Fbn2* loci at resolutions spanning 800 bp – 5 kb in mESC WT cells. RNA Pol II ChIP-seq data is shown below the maps for each treatment condition.



Extended Data Figure 10. Microcompartment-like structures are also visible in ultra-deep Hi-C data.

(**a-d**) Contact maps of ultra-deep Hi-C data in human lymphoblastoid cells (Gu *et al.*, *bioRxiv* (2021)[14] showing loci with structures sharing many microcompartmental features. Maps were generated using Juicebox's web interface[55] kindly provided by Dr. Jordan Rowley. Maps are shown at 1 kb resolution, with GM12878 gene annotations,

CTCF (ENCFF364OXN) and H3K27ac (ENCFF180LKW) ChIP-seq, and RNA-seq (ENCFF604VIC) signal tracks shown below the contact maps.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank TH Stanley Hsieh, the co-inventor of Micro-C, for holding a Micro-C workshop to teach us the protocol. We also thank Stanley Hsieh, Claudia Cattoglio, and Leonid Mirny for extensive and insightful discussions throughout this project. We thank Leonid Mirny and Geoff Fudenberg for insightful discussions on ICE normalization. We thank Emilia Dimitrova and Amy Hughes for their guidance on performing RNA Pol II ChIP-seq. We thank Job Dekker, Simon Grosse-Holz, Emily Navarette, Sameer Abraham, and the Hansen lab for helpful discussions throughout this project. We thank Elphege Nora, Miriam Huntley, Alistair Boettiger, Stanley Hsieh, Domenic Narducci, Sarah Nemsick, Asmita Jha, Michele Gabriele, Jin Harvey Yang, Christine Eyler, Martin Aryee, Sarah Johnstone, Christian Cerda-Smith, and Vijay Sankaran for critical feedback on the manuscript. We thank Marieke Oudelaar for providing TMCC numbers and feedback on the TMCC vs. RCMC comparison and the manuscript. We thank Jordan Rowley for sharing the Hi-C data generated in Gu *et al.* (*bioRxiv*, 2021). We thank the MIT Koch Institute's Robert A. Swanson (1969) Biotechnology Center for technical support, specifically the Integrated Genomics and Bioinformatics Core and MIT BioMicroCenter, and this work was supported in part by the Koch Institute Support (core) Grant P30-CA14051 from the National Cancer Institute. We also thank the Walk-Up Sequencing services of the Broad Institute of MIT and Harvard. This work was supported by NIH grants DP2GM140938 (ASH), R33CA257878 (ASH), and UM1HG011536 (ASH), NSF grant 2036037 (ASH), a Solomon Buchsbaum Research Support Committee award (ASH) and the Koch Institute Frontier Research Fund (ASH). VYG is supported by a graduate research fellowship from the Ludwig Center at MIT's Koch Institute. MKH is supported by an NIH F32GM140548 fellowship and a non-stipendiary EMBO fellowship.

Data Availability

The data generated in this study can be found at NCBI GEO under accession number GSE207225 at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE207225.

References

1. Dekker J et al. The 3D Genome as Moderator of Chromosomal Communication. Cell (2016). doi:10.1016/j.cell.2016.02.007
2. Oudelaar AM et al. The relationship between genome structure and function. Nat. Rev. Genet 22, 154–168 (2021). [PubMed: 33235358]
3. Lieberman-Aiden E et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science (80-.). 326, 289–293 (2009).
4. Nuebler J et al. Chromatin organization by an interplay of loop extrusion and compartmental segregation. Proc. Natl. Acad. Sci 115, E6697–E6706 (2018). [PubMed: 29967174]
5. Dixon JR et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380 (2012). [PubMed: 22495300]
6. Nora EP et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385 (2012). [PubMed: 22495304]
7. Rao SSP et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665–1680 (2014). [PubMed: 25497547]
8. Goel VY et al. The macro and micro of chromosome conformation capture. WIREs Dev. Biol n/a, e395 (2020).
9. Sanborn AL et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. U. S. A (2015). doi:10.1073/pnas.1518552112
10. Fudenberg G et al. Formation of Chromosomal Domains by Loop Extrusion. Cell Rep. (2016). doi:10.1016/j.celrep.2016.04.085

11. Krietenstein N et al. Ultrastructural Details of Mammalian Chromosome Architecture. Mol. Cell (2020). doi:10.1016/j.molcel.2020.03.003
12. Hsieh T-HS et al. Enhancer-promoter interactions and transcription are maintained upon acute loss of CTCF, cohesin, WAPL, and YY1. bioRxiv 2021.07.14.452365 (2021). doi:10.1101/2021.07.14.452365
13. Hsieh T-HS et al. Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. Mol. Cell (2020). doi:10.1016/j.molcel.2020.03.002
14. Gu H et al. Fine-mapping of nuclear compartments using ultra-deep Hi-C shows that active promoter and enhancer elements localize in the active A compartment even when adjacent sequences do not. bioRxiv 2021.10.03.462599 (2021). doi:10.1101/2021.10.03.462599
15. Hansen AS et al. Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. Mol. Cell (2019). doi:10.1016/j.molcel.2019.07.039
16. Hua P et al. Defining genome architecture at base-pair resolution. Nature 595, 125–129 (2021). [PubMed: 34108683]
17. Aljahani A et al. Analysis of sub-kilobase chromatin topology reveals nano-scale regulatory interactions with variable dependence on cohesin and CTCF. Nat. Commun 13, 2139 (2022). [PubMed: 35440598]
18. Gasperini M et al. A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. Cell 176, 377–390.e19 (2019). [PubMed: 30612741]
19. Barshad G et al. RNA polymerase II and PARP1 shape enhancer-promoter contacts. bioRxiv 2022.07.07.499190 (2022). doi:10.1101/2022.07.07.499190
20. Zhang S et al. Enhancer-promoter contact formation requires RNAPII and antagonizes loop extrusion. bioRxiv 2022.07.04.498738 (2022). doi:10.1101/2022.07.04.498738
21. Schwarzer W et al. Two independent modes of chromatin organization revealed by cohesin removal. Nature (2017). doi:10.1038/nature24281
22. Nora EP et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell (2017). doi:10.1016/j.cell.2017.05.004
23. Gassler J et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. EMBO J. 36, 3600–3618 (2017). [PubMed: 29217590]
24. Wutz G et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J. (2017). doi:10.15252/embj.201798004
25. Rao SSP et al. Cohesin Loss Eliminates All Loop Domains. Cell (2017). doi:10.1016/j.cell.2017.09.026
26. Haarhuis JHI et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell 169, 693–707.e14 (2017). [PubMed: 28475897]
27. El Khattabi L et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. Cell 178, 1145–1158.e20 (2019). [PubMed: 31402173]
28. Oudelaar AM et al. Dynamics of the 4D genome during in vivo lineage specification and differentiation. Nat. Commun 11, 2722 (2020). [PubMed: 32483172]
29. Jäger R et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. Nat. Commun 6, 6178 (2015). [PubMed: 25695508]
30. Imakaev M et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods 9, 999–1003 (2012). [PubMed: 22941365]
31. Bonev B et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell 171, 557–572.e24 (2017). [PubMed: 29053968]
32. Zhou HY et al. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. Genes Dev. 28, 2699–2711 (2014). [PubMed: 25512558]
33. Li Y et al. CRISPR Reveals a Distal Super-Enhancer Required for Sox2 Expression in Mouse Embryonic Stem Cells. PLoS One 9, e114485 (2014). [PubMed: 25486255]
34. Chakraborty S et al. High affinity enhancer-promoter interactions can bypass CTCF/cohesin-mediated insulation and contribute to phenotypic robustness. bioRxiv 2021.12.30.474562 (2022). doi:10.1101/2021.12.30.474562

35. Akgol Oksuz B et al. Systematic evaluation of chromosome conformation capture assays. Nat. Methods 18, 1046–1055 (2021). [PubMed: 34480151]
36. Abdennur N et al. Cooltools: enabling high-resolution Hi-C analysis in Python. bioRxiv 2022.10.31.514564 (2022). doi:10.1101/2022.10.31.514564
37. Roayaei Ardakany A et al. Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. Genome Biol. 21, 256 (2020). [PubMed: 32998764]
38. Gabriele M et al. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. Science (80-.). 376, 496–501 (2022).
39. Wang Z et al. Prediction of histone post-translational modification patterns based on nascent transcription data. Nat. Genet 54, 295–305 (2022). [PubMed: 35273399]
40. Rosencrance CD et al. Chromatin Hyperacetylation Impacts Chromosome Folding by Forming a Nuclear Subcompartment. Mol. Cell 78, 112–126.e12 (2020). [PubMed: 32243828]
41. You Q et al. Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution. Nat. Biotechnol 39, 225–235 (2021). [PubMed: 32839564]
42. Rippe K et al. Functional organization of RNA polymerase II in nuclear subcompartments. Curr. Opin. Cell Biol 74, 88–96 (2022). [PubMed: 35217398]
43. Leibler L Theory of Microphase Separation in Block Copolymers. Macromolecules 13, 1602–1617 (1980).
44. Meier DJ Theory of block copolymers. I. Domain formation in A-B block copolymers. J. Polym. Sci. Part C Polym. Symp 26, 81–98 (1969).
45. Fujishiro S et al. Generation of dynamic three-dimensional genome structure through phase separation of chromatin. Proc. Natl. Acad. Sci 119, e2109838119 (2022). [PubMed: 35617433]
46. Thiecke MJ et al. Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. Cell Rep. 32, (2020).
47. Kane L et al. Cohesin is required for long-range enhancer action at the Shh locus. Nat. Struct. Mol. Biol 29, 891–897 (2022). [PubMed: 36097291]
48. Calderon L et al. Cohesin-dependence of neuronal gene expression relates to chromatin loop length. Elife 11, e76539 (2022). [PubMed: 35471149]
49. Rinzema NJ et al. Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. Nat. Struct. Mol. Biol 29, 563–574 (2022). [PubMed: 35710842]
50. Cuartero S et al. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. Nat. Immunol 19, 932–941 (2018). [PubMed: 30127433]

Methods References

51. Navarro Gonzalez J et al. The UCSC Genome Browser database: 2021 update. Nucleic Acids Res. 49, D1046–D1057 (2021). [PubMed: 33221922]
52. Yang T et al. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. Genome Res. 27, 1939–1949 (2017). [PubMed: 28855260]
53. Venev Sergey, Abdennur Nezar, Goloborodko Anton, Flyamer Ilya, Fudenberg Geoffrey, Nuebler Johannes, Galitsyna Aleksandra, Akgol Betul, Abraham Sameer, Kerpedjiev Peter, & M. I. open2c/cooltools: v0.4.1 (v0.4.1). Zenodo 10.5281/zenodo.5214125 (2021).
54. Abdennur N et al. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics 36, 311–316 (2020). [PubMed: 31290943]
55. Robinson JT et al. Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. Cell Syst. 6, 256–258.e1 (2018). [PubMed: 29428417]
56. Pettitt SJ et al. Agouti C57BL/6N embryonic stem cells for mouse genetic resources. Nat. Methods 6, 493–495 (2009). [PubMed: 19525957]
57. Hansen AS et al. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. Elife 6, e25776 (2017). [PubMed: 28467304]

58. Kerpedjiev P et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol. 19, 125 (2018). [PubMed: 30143029]

59. Xu W et al. CoolBox: a flexible toolkit for visual analysis of genomics data. BMC Bioinformatics 22, 489 (2021). [PubMed: 34629071]

60. Zhao H et al. CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics 30, 1006–1007 (2014). [PubMed: 24351709]

61. Robinson JT et al. Integrative genomics viewer. Nat. Biotechnol 29, 24–26 (2011). [PubMed: 21221095]

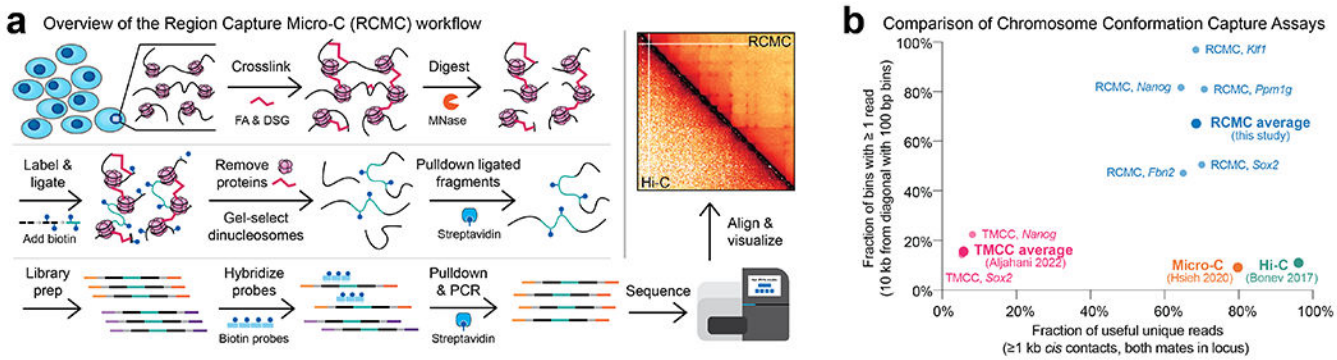


Figure 1. Region Capture Micro-C captures chromosome conformation at nucleosome resolution.

(**a**) Overview of the Region Capture Micro-C (RCMC) protocol. Cells are chemically fixed, nuclei are digested with micrococcal nuclease (MNase), and fragments are biotinylated, proximity ligated, dinucleosomes gel extracted and purified, library prepped, PCR amplified, and region-captured to create a sequencing library. After sequencing, mapping, and normalization, the data is visualized as a contact matrix. (**b**) Benchmarking comparison of RCMC against the highest resolution Tiled-Micro-Capture-C (TMCC)[17], Micro-C[13], and Hi-C[31] mESC datasets. Region-averaged calculations are shown for RCMC, TMCC, Micro-C, and Hi-C, and calculations for individual captured regions are also shown for RCMC and TMCC. The x-axis shows the fraction of all reads that uniquely map to the target region (both read mates fall within the Captured region) that are structurally informative (*cis* contacts >=1 kb). The y-axis shows the fraction of all contact bins separated by 10 kb that contain at least one read at 100 bp resolution.

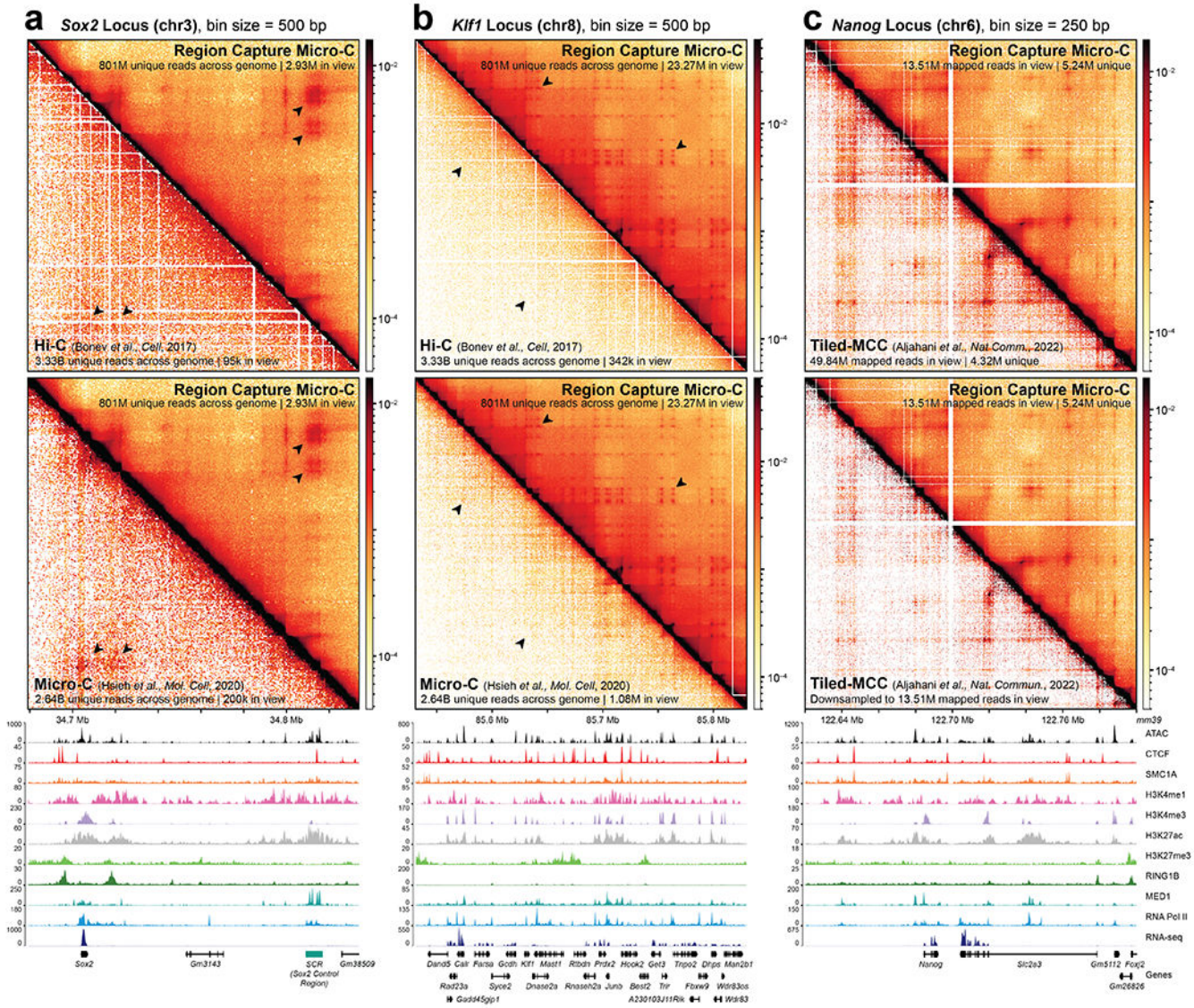


Figure 2. RCMC generates deep contact maps, reveals previously unresolved aspects of 3D genome structure, and outperforms other 3C methods.

(**a-b**) Contact map comparison of RCMC against the deepest available mESC Hi-C (top; Bonev 2017[31]) and Micro-C (middle; Hsieh 2020[13]) datasets at the (**a**) *Sox2* and (**b**) *Klf1* regions at 500 bp resolution. Gene annotations and ATAC, ChIP-seq, and RNA-seq (see Supplementary Table 1) signal tracks are shown below the contact maps, while the contact intensity scale is shown to the right. The RCMC data shown throughout this manuscript were pooled from two biological replicates in wild-type mESCs. (**c**) Contact map comparison of RCMC against Tiled-Micro-Capture-C (TMCC)[17] at the *Nanog* locus at 250 bp resolution. Full datasets are visualized in the top contact map, and TMCC has been downsampled to match the total number of RCMC sequencing reads in view in the bottom contact map.

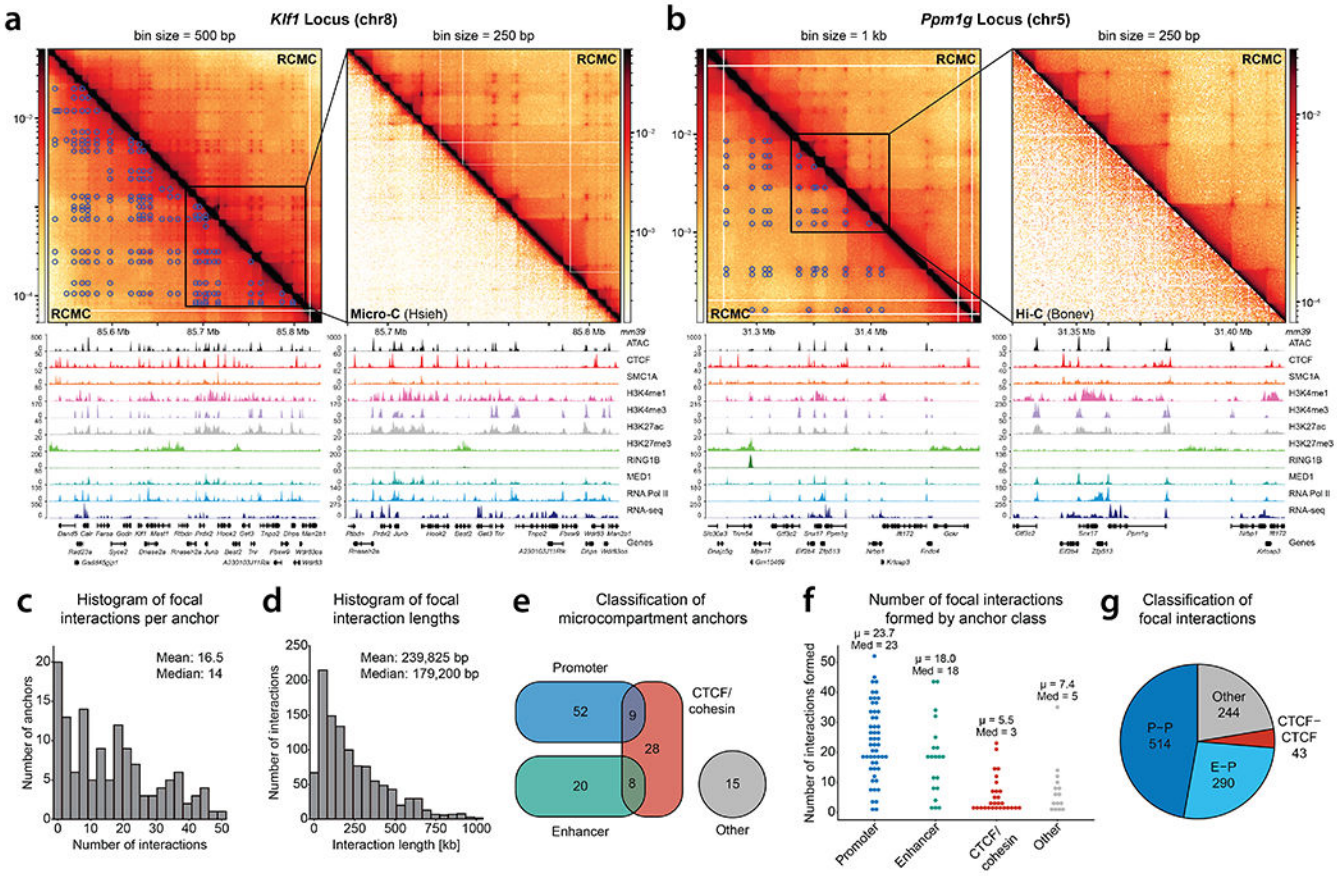


Figure 3. RCMC identifies highly nested, focal interactions called microcompartments which frequently connect enhancers and promoters.

(**a-b**) Contact map visualization of RCMC data and called microcompartments at the *Klf1* (**a**) and *Ppm1g* (**b**) locus at 500 bp (**a**) and 1 kb (**b**) resolution (left) and 250 bp resolution (zoom in, right). Manually annotated microcompartment contacts are shown below the contact map diagonal on the left, while comparisons against genome-wide Micro-C[13] (**a**) and Hi-C[31] (**b**) are shown on the right. (**c-d**) Histograms showing distributions of (**c**) the number of focal interactions formed by microcompartment anchors and (**d**) the lengths spanned by focal interactions in kb. (**e**) Venn diagram of microcompartment anchor categories according to chromatin features overlapped by the anchor ±1 kb. Promoters were defined as regions around annotated transcription start sites[51] ±2 kb, enhancers as regions with overlapping peaks of H3K4me1 (ENCFF282RLA) and H3K27ac (GSE90893) in ChIP-seq data which did not overlap promoters, and CTCF/cohesin as regions with overlapping peaks of CTCF (GSE90994) and SMC1A (GSE123636) in ChIP-seq data. Other regions are those not overlapping any of these features. (**f**) Swarm plot of the number of focal interactions formed by individual microcompartment anchors divided according to categories in (**e**), including the mean (μ) and median (Med) for each distribution. Anchors fitting into more than one category were excluded. (**g**) Fractions of loops classified into different categories: P-P (promoter-promoter), E-P (enhancer-promoter), CTCF-CTCF (CTCF/cohesin-CTCF/cohesin), other (Other-Other interactions, or any other combinations).

CTCF-CTCF interactions do not include any anchors which overlap promoter or enhancer regions.

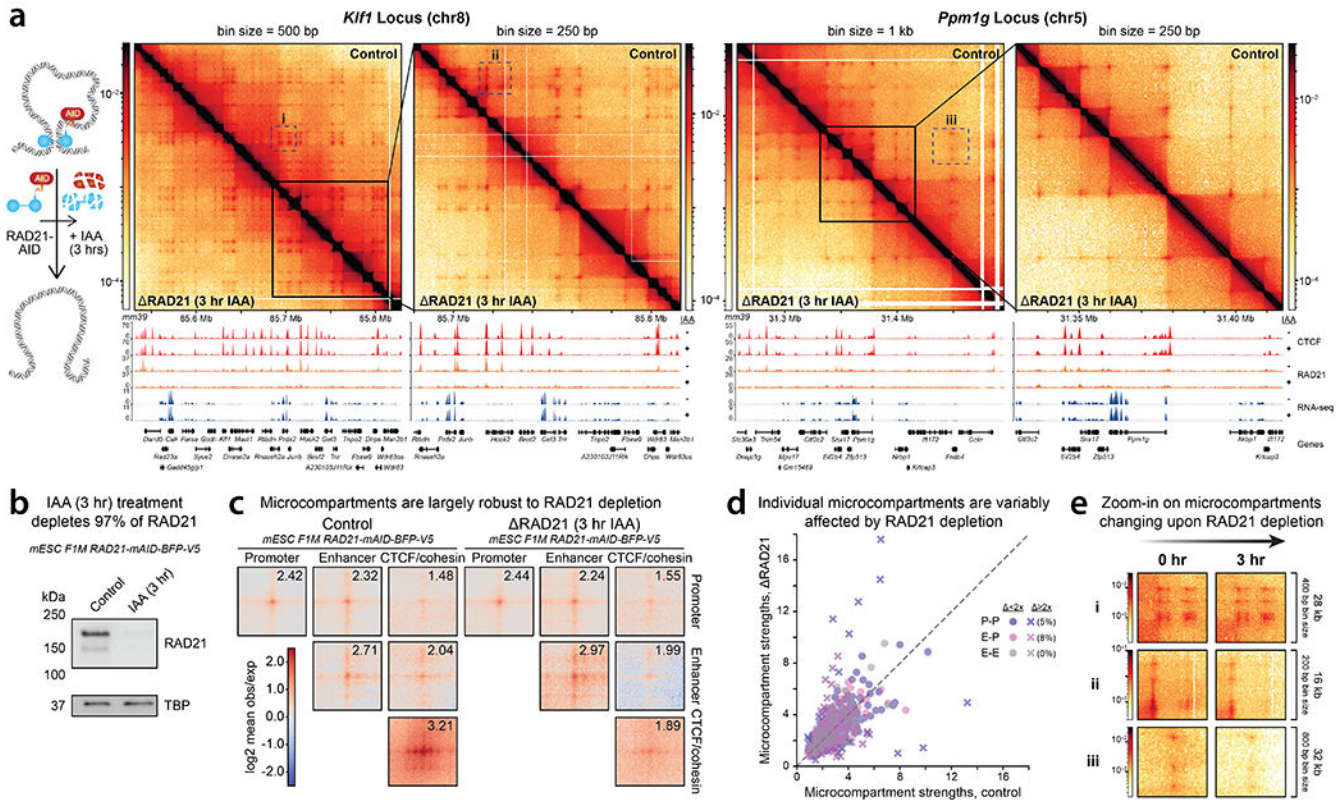


Figure 4. Most microcompartments are robust to the loss of loop extrusion.

(**a**) Cohesin (RAD21) depletion does not strongly perturb most microcompartments. Left: Treatment paradigm for rapid depletion of RAD21 upon IAA treatment in clone F1M RAD21-mAID-BFP-V5 mESCs[12,38]. Right: Contact maps comparing DMSO-treated control (above) and RAD21-depleted (below) samples are shown for the *Klf1* and *Ppm1g* loci. (**b**) Western blot showing near-complete (97%) depletion of RAD21 following 3 hours of IAA treatment. This Western blot was performed once using cells collected simultaneously for RCMC. (**c**) Aggregate peak analysis matrix of called microcompartmental contacts after RAD21 depletion compared to their respective controls, separated by the identity of each contact's constituent anchors. Plots show a 20 kb window centered on the loop at 250 bp resolution. The background-normalized intensity for a 1250x1250 bp box around the central dot for each aggregate peak is shown in the upper right of each plot as a quantification of aggregate dot strength. (**d**) Plot of individual microcompartment strengths (as quantified in (c)) in the RAD21-depleted (y-axis) and control (x-axis) conditions, shown for P-P (purple, n=418), E-P (pink, n=238), and E-E (gray, n=40) loops. Interactions changing in strength by two-fold or more are visualized as x's, while interactions below the threshold are visualized as circles and percentages are noted. (**e**) Zoomed-in contact maps of microcompartment examples in (a) that strengthen (i) or weaken (ii,iii) relative to the control treatment and the background in response to RAD21 depletion.

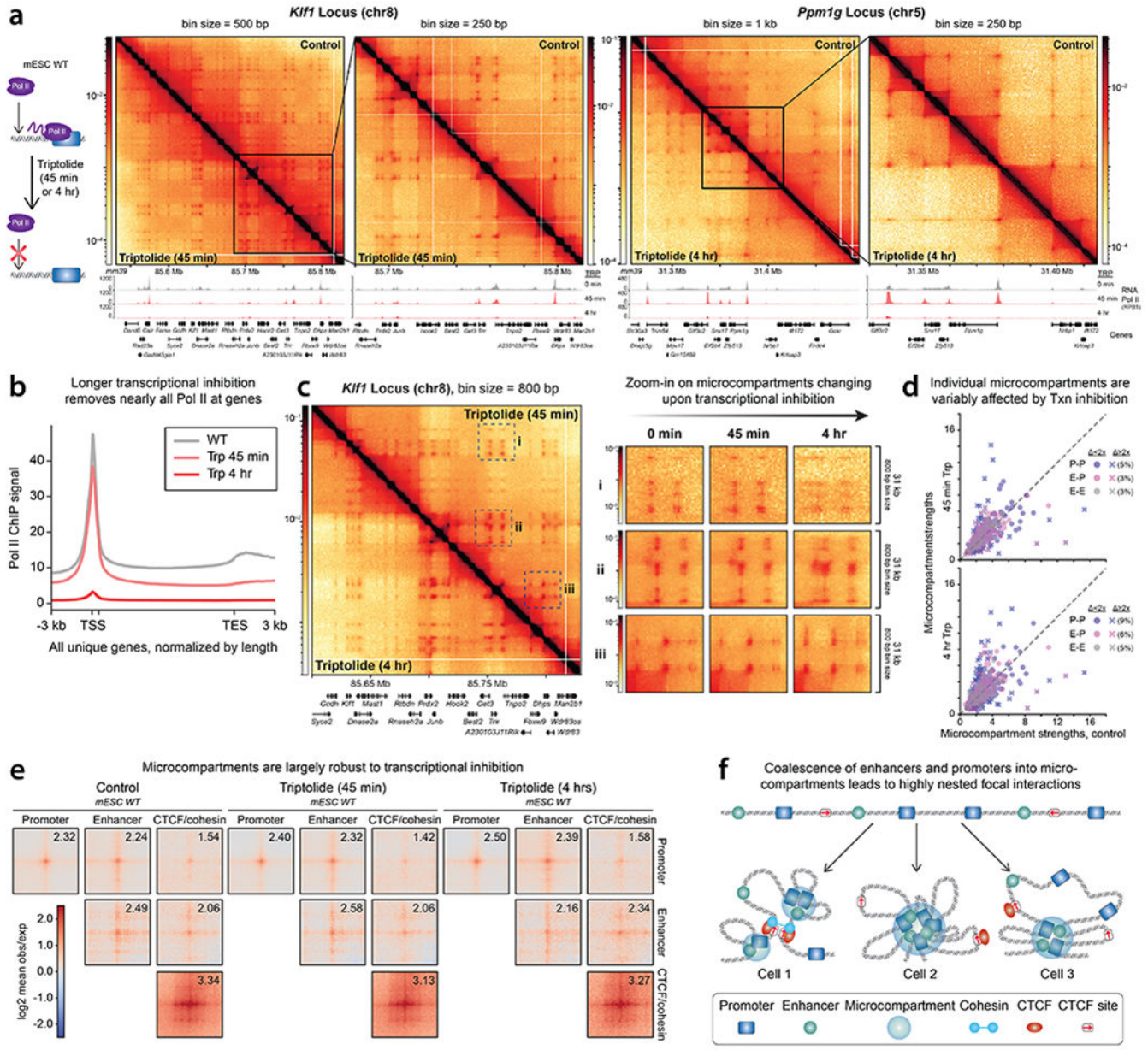


Figure 5. Most microcompartments are robust to the inhibition of transcription.

(**a**) Inhibition of transcription initiation with triptolide does not strongly affect most microcompartments. Left: Overview of triptolide treatment for WT mESCs (45 min or 4 hr). Right: Contact maps comparing WT control (above) and transcriptionally-inhibited (below) samples are shown for the *Klf1* locus (45 min timepoint shown vs control) and the *Ppm1g* locus (4 hr timepoint shown vs control). RNA Pol II ChIP-seq data (RPB1) is shown below. (**b**) Aggregate RPB1 RNA Pol II ChIP-seq signal at genes after triptolide treatment (45 min and 4 hr) and a control (WT). The x-axis depicts all unique mouse genes normalized by length and flanked by 3 kb upstream and downstream of their TSS and TES, respectively. The first 500 bp downstream of the TSS (marked by the second x-axis tick mark) are not normalized to avoid normalizing the core promoter against variable gene body

lengths. (**c**) Left: Contact maps comparing the transcriptional inhibition timepoints (45 min treatment above, 4 hr treatment below) are shown for the *Klf1* locus. Right: Zoomed-in contact maps of microcompartments across the control and triptolide treatment timepoints that weaken (i) or strengthen (ii,iii) in response to transcriptional inhibition. (**d**) Plot of individual microcompartment strengths in the transcriptionally inhibited (y-axis) and control (x-axis) conditions, shown for P-P (purple, n=418), E-P (pink, n=238), and E-E (gray, n=40) loops. Interactions changing in strength by two-fold or more are visualized as x's (percentages noted), and as circles otherwise. (**e**) Aggregate peak analysis matrix of called microcompartmental contacts across the two transcriptional inhibition timepoints compared to the control, separated by the identity of each contact's constituent anchors. Plots show a 20 kb window centered on the loop at 250 bp resolution, with background-normalized dot intensities shown in the upper right of each plot. (**f**) Proposed model for the formation of microcompartments. Coalescence of multiple promoters and enhancer elements in a gene-dense region may occur through A/B-block copolymer microphase separation, resulting in variable combinations of multiway interactions being present in different cells and giving rise to tessellated focal interactions in population-averaged RCMC data.