

Genome reorganization and its functional impact during breast cancer progression

Kathleen S Metz Reed¹, Andrew Fritz², Haley Greenyer², Kerstin Heselmeyer-Haddad³, Seth Fietze⁴, Janet Stein², Gary Stein^{2*}, Tom Misteli^{1*}

¹National Cancer Institute, National Institutes of Health, Bethesda, United States; ²Department of Biochemistry and University of Vermont Cancer Center, University of Vermont Larner College of Medicine, Burlington, United States; ³OMICS Technology Facility, Genetics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, United States; ⁴College of Nursing and Health Sciences, University of Vermont Cancer Center, Burlington, United States

eLife Assessment

The study by Reed et al. provides **fundamental** findings defining the topological changes that occur during tumorigenesis. These **compelling** findings enhance the understanding of stable long-range connections among genes that reprogram cancer-related functions.

*For correspondence:

Gary.Stein@med.uvm.edu (GS);
mistelit@mail.nih.gov (TM)

Competing interest: The authors declare that no competing interests exist.

Funding: See page 25

Preprint posted

08 July 2025

Sent for Review

08 July 2025

Reviewed preprint posted

10 September 2025

Reviewed preprint revised

21 May 2026

Version of Record published

16 June 2026

Reviewing Editor: Yu Zhao, Institute of Radiation Medicine, Chinese Academy of Medical Sciences and Peking Union Medical College, China

© This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0 public domain dedication](https://creativecommons.org/licenses/by/4.0/).

Abstract Cancer progression involves extensive alterations in epigenetic and gene expression programs, but the accompanying changes in higher-order genome organization remain less well understood. Using high-resolution Micro-C mapping in the MCF10 cell model of breast cancer, we profiled chromatin compartments, topologically associated domains, and chromatin loops. We find large-scale compartmental shifts occur predominantly in early stages of cancer development, with more fine-scale structural changes in topologically associating domains and loops accumulating during the later transition to metastasis. Relating these chromatin features to gene expression and enhancer-associated histone marks revealed that many differentially expressed genes are physically connected to distal regulatory elements. While enhancer–promoter contact frequency and distal enhancer activity correlated with gene expression, strong changes in chromatin looping were relatively infrequent during progression, suggesting that alterations in chromatin contacts are not globally necessary, but may facilitate gene regulation at a subset of genes. These results elucidate the connection between gene regulation and genome remodeling in a cell-based cancer progression model.

Introduction

The eukaryotic genome is highly organized in the cell nucleus. Among the most prominent structural features are kilobase-sized chromatin loops, medium-scale topologically associating domains (TADs), and higher-order compartments (Dekker and Mirny, 2016; Jerkovic and Cavalli, 2021; Misteli, 2020; Pombo and Dillon, 2015; Rowley and Corces, 2018). How chromatin organization contributes to epigenetic control of gene regulation, including in physiological and pathological settings, such as cancer, remains only partially understood (Corces et al., 2016; Peng et al., 2022; Zhang et al., 2021; Amodeo et al., 2025).

A prominent mechanism to generate higher-order chromatin structures is loop extrusion (Davidson and Peters, 2021; de Wit and Nora, 2023; Sanborn et al., 2015). During this process,

the multi-component cohesin complex is loaded onto chromatin and, using its intrinsic molecular motor activity, extrudes chromatin bidirectionally along the genome to form a loop or a domain until it encounters the major chromatin architectural protein CTCF bound to convergently oriented binding sites. The encounter of cohesin with bound CTCF stalls the extrusion process and generates a chromatin loop or TAD. While loop extrusion has universally been implicated in the formation of chromatin loops and domains, the formation of chromatin features by other mechanisms has also been observed, especially at a smaller scale (Goel et al., 2023; Hsieh et al., 2020; Oberbeckmann et al., 2024; Rowley and Corces, 2018).

A common property of higher-order chromatin folding is that the resulting loops, domains, and compartments bring distal genome elements into spatial proximity which has been implicated in gene regulation (Jerkovic and Cavalli, 2021). For example, it has been suggested that one function of TADs is to facilitate the interaction of regulatory elements, particularly gene enhancers, with their target genes located within the same TAD (Chakraborty et al., 2024; Chakraborty et al., 2025; da Costa-Nunes and Noordermeer, 2023; Xiao et al., 2021; Sood and Misteli, 2022). Similarly, long-range interactions via chromatin loops are thought to be essential at some loci to bring gene enhancers into proximity to their target promoters (D'Ippolito et al., 2018; Kane et al., 2022; Schoenfelder and Fraser, 2019; Stolper et al., 2023; Popay and Dixon, 2022). However, regulation of many gene loci also appears independent of chromatin organization, and enhancer–promoter proximity often does not correlate with gene activity (Benabdallah et al., 2019; Despang et al., 2019; Hansen et al., 2024; Hsieh et al., 2022). In fact, acute depletion of cohesin revealed genome-wide disruption of chromatin organization but surprisingly limited impact on gene expression (Rao et al., 2017). A possible explanation for these divergent findings is that chromatin organization may be functionally more relevant to bring about changes in gene activity rather than maintenance of gene expression, as suggested by several cohesin depletion studies (Hsieh et al., 2022; Cuartero et al., 2018; Popay et al., 2024).

Genome organization is likely relevant for cancer and its progression. Mutations in loop extrusion machinery, such as cohesin and the cohesin processivity factor NIPBL or at CTCF-binding sites have been reported in many cancers (Corces et al., 2016; Peng et al., 2022; Wang et al., 2022; Guo et al., 2018; Katainen et al., 2015). In addition, many structural variants (SVs) such as deletions, duplications, and translocations have been documented in various cancer subtypes where SVs and the ensuing reorganization around them can lead to aberrant gene regulation (Cosenza et al., 2022; Mitelman et al., 2007; Della Chiara et al., 2023).

Despite these observations, the full extent of genome reorganization during cancer progression and its functional consequences remains largely unknown. Several studies primarily focused on large-scale reorganizations have found changes in higher-order chromatin organization such as chromosome clustering and dynamic compartments, some of which correlated with changes in differentially expressed oncogenes and enhancers (Barutcu et al., 2015; Fritz et al., 2018; Johnstone et al., 2020; Li et al., 2021; Mourad et al., 2014; Rickman et al., 2012; Taberlay et al., 2016; Yang et al., 2020; Zhou et al., 2019). Analysis of TADs in various cancers has found mixed results, with some studies pointing to increased TADs and gained boundaries and others observing more stable TAD organization or weakened boundaries (Johnstone et al., 2020; Taberlay et al., 2016; Yang et al., 2020; Kim et al., 2022; Rossini et al., 2024). Furthermore, cancer-associated SVs, such as chromosomal translocations, have been related to altered gene expression, for example via enhancer-hijacking (Mortenson et al., 2024; Wang et al., 2021). However, how local chromatin loops and TADs are restructured during oncogenic reprogramming and how these changes relate to cancer-associated gene expression has not been well documented.

We address the question of how local and global chromatin organization changes during cancer progression and how these events relate to cancer gene expression programs by generating high-resolution Micro-C maps in the well-established MCF10 breast cancer progression model (Puleo and Polyak, 2021). This cancer model consists of three epithelial cell lines that were sequentially derived from a non-cancerous individual (Soule et al., 1990). MCF10A is an adherent epithelial non-cancerous cell line that spontaneously immortalized from the initial sample. Pre-malignant MCF10AT1 cells were derived from MCF10A cells by overexpressing mutant Ha-Ras oncogene followed by long-term passage as mouse xenografts (Dawson et al., 1996). In immunocompromised mice, MCF10AT1 cells form precancerous lesions, and approximately 25% progress to invasive carcinoma over time. The

MCF10CA1a cell line is derived from metastatic tumors generated from xenografted MCF10AT1 cells and is considered fully malignant, metastatic, and highly aggressive, forming tumors in 100% of xenografted mice which can metastasize rapidly to the lung (Santner et al., 2001). The MCF10 progression series represents a spectrum of cells that share a similar genetic background but are increasingly more cancerous, and they have been widely used to study the genetic and epigenetic changes that occur during cancer progression and epithelial–mesenchymal transition (Fritz et al., 2018; Bhardwaj et al., 2017; Buchanan et al., 2007; Dorgham et al., 2023; Kadota et al., 2010; Lin et al., 2019; Maguire et al., 2016; Marella et al., 2009; Rhee et al., 2008; So et al., 2012).

In this study, comparing fine-scale chromatin organization and other epigenetic features in the MCF10 cancer progression model has allowed us to identify changes in genome reorganization and relate them to changes in gene expression, including of cancer progression-associated genes. Our results provide novel insights into the principles of chromatin-mediated gene regulation and into the dynamic structure–function relationship contributing to genome regulation derived from an in vitro cancer progression model.

Results

Mapping global and local genome organization across breast cancer progression

To understand how genome organization changes during cancer progression, we generated high-resolution (5 kb) genome-wide maps of chromatin contacts using Micro-C in the MCF10A, MCF10AT1, and MCF10CA1a cancer progression series (Figure 1A, Figure 1—figure supplement 1A). We obtained high-quality data with at least 1 billion Micro-C contacts per cell line, spread across two biological replicates with four technical replicates each (Supplementary file 1).

We identified features of chromatin organization at several levels, including large-scale reorganizations of compartments, medium-scale changes in TADs, and fine-scale changes in chromatin loops (Figure 1B). We assigned A and B compartments using CALDER at a resolution of 10 kb, TADs using SpectralTAD combined with FAN-C boundary insulation score (IS) calculations at 10 kb, and loops using SIP at 5- and 10-kb resolution (see Methods for details). Each cell type had similar percentages of the genome assigned to the active A (47.1–50.2%) or inactive B (49.7–52.9%) compartment, with the two cancer cell types MCF10AT1 and MCF10CA1a having a slightly higher proportion of A compartment designations (Figure 1C). Similarly, we detected a similar number of TADs (between 7459 and 7825) and chromatin loops (between 15,713 and 17,332) in each cell line (Figure 1D, E). Although the three cell lines are karyotypically similar due to their shared genetic background, they contain large-scale chromosomal duplications and translocations which were identified by SKY karyotyping analysis (Figure 1—figure supplement 1B), and numerical chromosome aberrations based on SKY and Micro-C sequencing depth analysis were included in the analysis of all chromatin features (see Methods; Figure 1—figure supplement 1C, Supplementary file 2). After this correction, chromatin loop counts showed a high degree of reproducibility between technical and biological replicates (Figure 1—figure supplement 1D). We used this deep dataset to characterize structural features that are reshaped during cancer progression.

Cancer progression reorganizes compartments, TADs, and chromatin loops

Comparative analysis across the three cell types identified significant changes in all major chromatin features during cancer progression (Figure 1, Figure 1—figure supplement 2A, B, Supplementary files 3–5).

At a large-scale, we detected changes in compartmentalization. We observed a general shift towards the more active A compartment in early cancer progression, with a larger portion of genomic regions becoming more A-like (31.0%) compared to more B-like (26.0%) in the transition from MCF10A to MCF10AT1, while these changes are more balanced in the later transition from MCF10AT1 to MCF10CA1a (30.0% and 30.6%, respectively). Interactions within the most A- and B-like compartments were predominant in the pre-cancerous MCF10A cells (Figure 1F). However, in MCF10AT1 and MCF10CA1a, stronger interactions appear between more intermediate regions, suggesting a greater degree of intermixing that is consistent with increased compartmental heterogeneity which appears

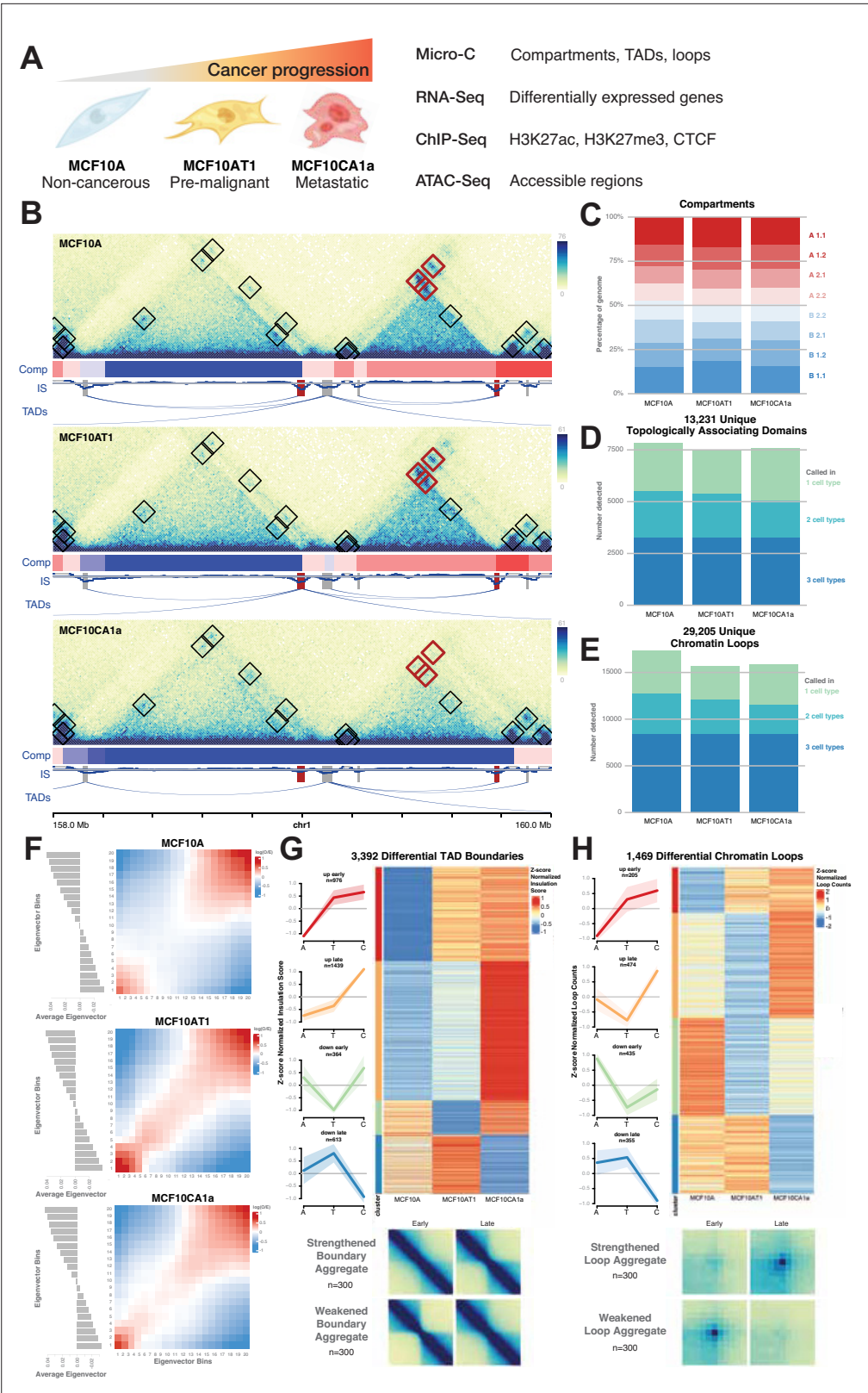


Figure 1. Reorganization of compartments, topologically associating domains (TADs), and loops during breast cancer progression. **(A)** A diagram of the experimental design. Three epithelial cell lines represent various stages of breast cancer progression; MCF10A are non-cancerous, MCF10AT1 are pre-malignant, and MCF10CA1a are metastatic. In each cell line, we generated 5-kb resolution, Micro-C to identify features such as compartments, *Figure 1 continued on next page*

Figure 1 continued

TADs, and chromatin loops. We overlapped these features with functional changes in gene expression from RNA-Seq, histone modifications, and CTCF binding from ChIP-Seq, and chromatin accessibility from ATAC-Seq. **(B)** Micro-C maps of a 2-Mb region of chromosome 1 in MCF10A (non-cancerous), MCF10AT1 (pre-cancerous), and MCF10CA1a (metastatic) cells at 5-kb resolution. Each map has annotations for loop calls, both static (black boxes) and differential (red boxes). Below each map is a track indicating compartment calls from CALDER (dark red is most A-like, dark blue is most B-like) as well as insulation scores tracks with static (gray) and differential (red) boundaries marked. Ribbons indicate TAD calls for each cell type. **(C)** Lengths of the genome assigned to each compartment in each cell type. **(D)** Numbers of TAD and **(E)** loop calls from each cell type, colored by the number of cell types in which they were initially detected. **(F)** Saddle plots of interactions between regions of different compartments in MCF10A, MCF10AT1, and MCF10CA1a. Bottom plots indicate the average eigenvector value for each compartment ventile. Plots shown are for chromosomes 2, 12, and 17 (see Methods). **(G)** Left: Differential TAD boundaries clustered by their timing of change, depicted in line plots and heatmap. Right: Aggregate plots of weakened and strengthened TAD boundaries ($n = 100$). **(H)** Left: Differential chromatin loops clustered by their timing of change, depicted in line plots and heatmap. Right: Aggregate plots of weakened and strengthened loops ($n = 100$).

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Karyotypic analysis and Micro-C loop reproducibility.

Figure supplement 2. Differential loop and topologically associating domain (TAD) feature details.

Figure supplement 3. Differential MCF10 loops are conserved in cancer cell lines.

Figure supplement 4. Differential topologically associating domain (TAD) boundaries are conserved in cancer cell lines and tissues.

to occur early during cancer progression (**Figure 1F**; *Johnstone et al., 2020*; *Sen et al., 2024*; *Liu et al., 2025*).

TAD boundaries represent genomic regions where upstream and downstream sequences are partially insulated from one another, with fewer contacts between them than within (*Dekker and Mirny, 2016*; *Rowley and Corces, 2018*; *da Costa-Nunes and Noordermeer, 2023*). We detected a total of 13,231 TADs across all three cell types, with 17,097 unique boundaries. TADs detected range in size from 190 kb to as large as 3.8 Mb, with a mean of 663 kb and a median of 460 kb (**Figure 1—figure supplement 2C**). Assessing changes in IS at TAD boundaries revealed 3392 (19.8%) boundaries where the degree of insulation changed significantly over the course of cancer progression. Because individual boundaries may be simultaneously used by multiple TADs, the total number of TADs which changed during progression is 5084 (38.5%) (**Figure 1G**). There are nearly three times as many boundary changes between later stages in cancer progression (1693 differential boundaries between MCF10AT1 and MCF10CA1a) than early stages (567 between MCF10A and MCF10AT1). Interestingly, TAD boundaries that gained or lost insulation during progression showed a significant enrichment for weakened boundaries (71.2%) with far fewer boundaries exhibiting increased insulation strength as cancer progressed (28.8%; p -value 0, permutation test, **Figure 1—figure supplement 2D**). This late-stage weakening of boundaries may reflect a more heterogeneous cell population as cancer progresses (*Corces et al., 2016*; *Bose et al., 2024*; *Dagogo-Jack and Shaw, 2018*; *van den Brand et al., 2025*; *Choppavarapu et al., 2025*).

Chromatin loops are formed by two distal genomic regions that are in more frequent contact than their surrounding or intervening sequences, indicated by higher contact frequency (*Dekker and Mirny, 2016*; *Misteli, 2020*; *Davidson and Peters, 2021*; *de Wit and Nora, 2023*). We found 29,205 chromatin loops across all three cell lines, ranging in size from 50 kb to 2 Mb, with a mean of 402 kb and median of 270 kb in length (**Figure 1—figure supplement 2C**). 77.6% of loop anchors coincided with CTCF peaks, representing 95.0% of loops with at least one anchor bound by CTCF, and CTCF-bound loops were stronger and longer than non-CTCF loops (**Figure 1—figure supplement 2E, F**). Loop boundaries often overlapped with TAD boundaries, with 52.4% of TADs consisting of loop domains across all cell lines. However, a majority (73.0%) of chromatin loops did not include TADs (**Figure 1—figure supplement 2G**). TADs without loop interactions at their boundaries tended to be larger, while loops without TADs can be found at all sizes but are enriched for shorter loops (**Figure 1—figure supplement 2H**).

To identify loops that changed significantly during cancer progression, we assessed changes in contact frequency among every loop in each cell type, correcting for karyotypic differences that result in differences in coverage between cell lines (see Methods). We identified 1469 chromatin loops that change significantly over the course of cancer progression, including both weakened and strengthened contacts (**Figure 1H**), representing 5.0% of all identified loops. Differential loops were defined by a ≥ 1.5 fold-change between contact frequencies in any two MCF10 cell lines, and an adjusted p-value of < 0.025 when considering variation across biological and technical replicates. Unlike TADs, there was a more balanced number of changes between early (1004 differential loops between MCF10A and MCF10AT1) and late (1204 between MCF10AT1 and MCF10CA1a) progression stages, as well as between strengthened (679 loops, 46.2% of all differential loops) and weakened loops (790 loops, 53.8% of all differential loops). Interestingly, only a small portion (19.0%) of differential loops were accompanied by changes in CTCF binding (**Figure 1—figure supplement 2I**). Motif analysis of differential loop anchors revealed only weak motifs of various transcription factors enriched at the boundaries of gained and lost loops, although occupancy did not appear high enough to explain most of the changes we observe (**Figure 1—figure supplement 2J**). Weakened loops were often associated with a decrease in H3K27ac, a mark of active enhancers, consistent with the notion that active enhancers can help recruit loop extrusion machinery (**Figure 1—figure supplement 2K**; see below) (**Georgiades et al., 2023; Rinzema et al., 2022**).

Taken together, these results demonstrate significant global changes in genome organization during cancer progression across multiple scales from chromatin compartments to loops.

Comparison to breast cancer patient data

To assess how generalizable the static and dynamic structures detected in the MCF10 progression series are to human tumors, we examined the chromatin loops and TAD boundaries from the MCF10 progression series in non-cancerous mammary epithelial cells (HMEC), five cell lines representing distinct cancer subtypes ranging from less (luminal, HER2+) to more aggressive (triple negative), as well as tissue samples from triple negative breast cancer (TNBC) patients with contralateral normal controls (**Figure 1—figure supplements 3 and 4; Kim et al., 2022**).

We found that most loops and TAD boundaries detected in MCF10 cells had conserved signatures in each of the other cell types, with chromatin loops generally showing high observed-versus-expected contact frequencies and boundaries showing strong dips in IS (**Figure 1—figure supplements 3A and 4A, B**). When we compared the profiles of differential MCF10 features relative to static MCF10 features within each cell type, we found some cell-line specific changes. For example, loops in the 'up-early' cluster, which are weaker in MCF10A and stronger in both MCF10AT1 and MCF10CA1a, were significantly stronger than static loops among patient TNBC cells (**Figure 1—figure supplement 3B**; p-value ≤ 0.05 , Wilcoxon rank sum test). This loop cluster was also stronger in HER2+ (HCC1954) and TNBC A (HCC70) cell lines, and the 'down-early' loop cluster is significantly weaker than static loops. Among differential TAD boundaries, those that have strong insulation in MCF10A and MCF10AT1 cells, but reduced insulation in MCF10CA1a, showed stronger insulation among all other examined cancer cell types when compared to static loop boundaries, suggesting that the boundaries that weaken in MCF10CA1a are not necessarily consistently weakened in breast cancer cell lines and tissues (**Figure 1—figure supplement 4C**; p-value ≤ 0.05 , Wilcoxon rank sum test). These findings highlight that different model systems indeed have distinct profiles of structural change, just as they have distinct gene expression profiles.

Chromatin loops connect gene promoters to distal regulatory features

We then sought to explore how long-range chromatin interactions relate to gene expression. We identified 17,185 expressed genes across any of the three MCF10 cell types using RNA-Seq and 8840 differentially expressed genes across all pairwise comparisons in the MCF10 cancer progression (see Methods; **Figure 2—figure supplement 1A, B, Supplementary file 6**). A similar number of genes changed in later stages of cancer development (4968 between MCF10AT1 and MCF10CA1a) compared to early progression (4773 between MCF10A and MCF10AT1). Reassuringly, as expected from previous studies (**Fritz et al., 2018; Rhee et al., 2008**), genes associated with epithelial morphogenesis and cell adhesion were upregulated early during progression, whereas regulation of differentiation, tissue development, metabolism, and signal transduction genes was observed during later

stages of progression (**Figure 2—figure supplement 1C**). These changes are consistent with the development of an intermediate and diverse pre-cancerous state early on during progression, while late changes are known to facilitate metastasis and support the epithelial-to-mesenchymal-like transition observed phenotypically among the progression series (**Figure 2—figure supplement 1D**; **Ferlier and Coulouarn, 2022**; **Gavish et al., 2023**; **Liberzon et al., 2015**; **Ma et al., 2003**).

To assess clinical significance of the differentially expressed genes in the MCF10 progression system, we overlapped this gene set with data from the Cancer Genome Atlas breast cancer cohort (TCGA-BRCA) (**Ciriello et al., 2015**; **Koboldt et al., 2012**). A total of 1884 genes that are differentially expressed between any stages of the MCF10 progression series were also differentially expressed between normal and tumor samples from breast cancer patients, but only roughly half change in the same direction (see Methods; **Figure 2—figure supplement 1E**). Interestingly, we found a higher degree of directional agreement between late changing genes (i.e. genes that change in MCF10CA1a compared to MCF10A and MCF10AT1) than early changing genes (i.e. genes that change in MCF10AT1 and MCF10CA1a compared to MCF10A). Of note, several looped genes which are differentially expressed in the MCF10 series showed an effect on overall patient survival (**Figure 2—figure supplement 2A–D**), and all but one of these genes exhibit significant differences in expression between normal and tumor TCGA-BRCA samples (**Figure 1—figure supplement 2E**; two-sampled Wilcoxon test of normalized counts, see Methods).

To explore the functional role of chromatin loops, we next overlapped chromatin loops with expressed and differential gene promoters and potential regulatory regions. We identified 52,953 potential enhancers as defined by overlapping histone H3K27ac ChIP-Seq and ATAC-Seq accessibility peaks, commonly used to identify enhancer regions (see Methods) (**Creyghton et al., 2010**; **Glossary, 2025**; **Heintzman et al., 2007**). Approximately 66.8% of chromatin loops featured some combination of active gene promoters and enhancers within 10 kb of loop anchors (**Figure 2A**). We found that chromatin loops that connect two features (either enhancers or promoters; mean length 260–311 kb) are typically shorter than those that contain only one feature (mean length 376–435 kb) or none (**Figure 2B**; mean length 532 kb; p-value for all comparisons $<2.2 \times 10^{-16}$, Wilcoxon rank sum test), with promoter–promoter loops being the smallest on average (mean length 260 kb). Interestingly, enhancer–promoter and enhancer–enhancer loops (mean counts 7.9 and 7.9, respectively) were stronger than promoter–promoter loops (mean counts 7.2) despite being longer on average, suggesting that epigenetic signatures associated with active enhancers may support stronger contacts (**Figure 2C**; p-value for both comparisons $<1.3 \times 10^{-5}$, Wilcoxon rank sum test).

Differentially expressed genes show proximal and distal epigenetic changes at persistent chromatin loops

To understand the regulatory modes of action of genes which were differentially expressed during cancer progression, we determined if they had gained or lost the activity-associated H3K27ac mark at their promoters or at distally looped enhancers. We found that while many genes only featured a corresponding change in H3K27ac at their promoter (57.9% of upregulated and 34.1% of downregulated genes), a large percentage also showed changes in distal enhancer activity (26.5% of upregulated and 15.7% of downregulated genes), suggesting that enhancer loops may be playing an important functional role in control of these genes (**Figure 2D, F**). Comparing the direction of fold-change for genes and promoter H3K27ac, distal H3K27ac, or contact frequency with distal enhancers using Fisher's Exact test revealed odds ratios significantly higher than 1 for all comparisons (8.8, 2.1, and 1.2, respectively, for changes between MCF10A and MCF10CA1a), but that there was a stronger association with promoters than enhancers or loop strength (**Figure 2—figure supplement 3A, B**). This trend was similar for genes that were differentially regulated both early and late, suggesting that the role of chromatin loops is consistent across all stages of cancer progression. In support of the finding that both distal regulatory changes and changes in contact frequency appear to contribute to changes in gene expression, looped genes that are regulated similarly in MCF10 progression and patient data include both up- and downregulated genes anchored by both static (i.e. *RRM2* and *FERMT2*) and dynamic chromatin loops (i.e. *INHBA* and *PCDH9*; **Figure 2—figure supplement 3C–F**).

To further assess the contribution of looped regulatory regions, we explored the degree of change in epigenetic marks and contact frequency between proximal and distal regions. Comparing the changes in acetylation at all gene promoters and distal regulatory regions revealed that upregulated

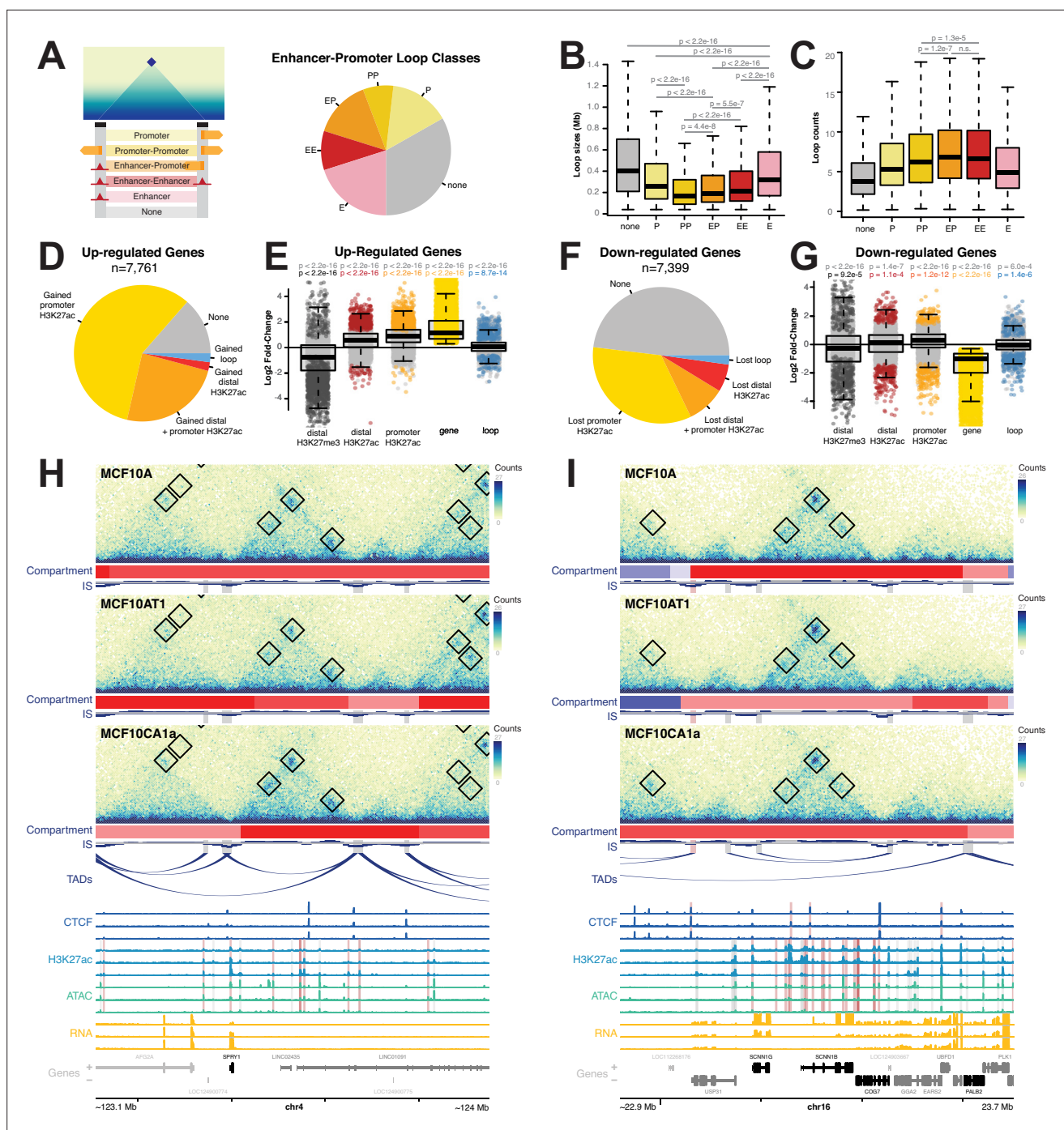


Figure 2. Persistent chromatin loops connect differentially expressed genes to distal enhancers. (A) Percentages of loops designated as either promoter–promoter, enhancer–promoter, enhancer–enhancer, or single-sided promoter or enhancer loops. (B) Distributions of loop sizes by enhancer/promoter designations. p-values represent Wilcoxon tests comparing the means of different loop classes. Boxplots show the median (middle line), 25th and 75th quartiles (box perimeters), and range excluding outliers (dashed line whiskers). Outliers are defined as values that are over 1.5 times the interquartile range beyond the box bounds and are excluded from these plots. p-values represent Wilcoxon tests comparing the means of various loop sets. Non-significant (n.s.) represents p-values above 0.05. (C) Distributions of loop strength by enhancer/promoter designations. p-values represent Wilcoxon tests comparing the means of various loop sets. Non-significant (n.s.) represents p-values above 0.05. (D) Percentages of upregulated genes that have gained H3K27ac at promoters, distal enhancers, both, or gained loops. (E) Log₂ fold-change of distal H3K27me3 (gray), distal H3K27ac (red), promoter H3K27ac (orange), gene expression (yellow), and loop strength (blue), when overlapped. Gray dots indicate features that do not change significantly, while colored points are significantly differential features. Boxplots are defined as in (B). p-values represent Wilcoxon tests comparing the means of each class to 0. Non-significant (n.s.) represents p-values above 0.01. (F) Percentages of downregulated genes that have gained H3K27ac at

Figure 2 continued on next page

Figure 2 continued

promoters, distal enhancers, both, or gained loops. (G) Log₂ fold-change of distal H3K27me3 (gray), distal H3K27ac (red), promoter H3K27ac (orange), gene expression (yellow), and loop strength (blue), when overlapped. Boxplot details as defined in (E). (H) An example of an upregulated gene (SPRY1) connected to gained enhancers by static loops. Black boxes show loop annotations. Red compartment tracks indicate A compartments, while blue indicates B compartments. In CTCF signal tracks, red highlights indicate differential CTCF peaks. In H3K27ac and ATAC-Seq signal tracks, red highlights indicate differential enhancers as determined by changes in H3K27ac. Genes highlighted in black are differentially expressed. (I) An example of downregulated genes (SCNN1G and SCNN1B) connected to lost enhancers by static loops. Plot annotations are as described in (H).

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Differential gene expression patterns across breast cancer progression include cancer-relevant pathways.

Figure supplement 2. Clinical relevance of select genes misregulated in MCF10.

Figure supplement 3. Differentially expressed genes overlap differential and static loops with differential distal regulatory regions.

Figure supplement 4. Relationship between differential gene expression, chromatin loops, and distal enhancers is consistent based on expression levels and cancer stages.

Figure supplement 5. Differential topologically associating domain (TAD) boundaries and subcompartment shifts have a subtle relationship to gene expression.

genes exhibit a significant increase in H3K27ac at distally looped enhancers, as well as a significant loss of repressive H3K27me3 marks (**Figure 2E**; p-value <2.2e−16, one-sample Wilcoxon test). This trend is weaker for downregulated genes, which feature a closer balance of gained and lost H3K27ac at both enhancers and promoters, as well as both gained and lost H3K27me3 at distal regions (**Figure 2G**). These results suggest that the static chromatin structures observed during the cancer progression process contribute to the control of differentially regulated genes, particularly among upregulated genes.

We observed that genes that are more strongly expressed are more likely to overlap with the anchor of a chromatin loop, ranging from roughly 33.4% overlap for genes with 0 read counts, 47.0% for genes with 10–1000 counts, and 56.2% for genes with 5000 and more counts (**Figure 2—figure supplement 4A**). To explore whether differential loops are more prominent among genes that change from a low active state to higher expression bins, we analyzed 108 genes that went from an unexpressed or ‘off’ state (2 or fewer read counts) in one cell line to an expressed ‘on’ state (100 or more read counts) in another (**Figure 2—figure supplement 4B–E**). While these genes were not enriched for differential loops, over 40% overlap with static loops. Similarly, genes that change from a modest ‘on’ state to high expression levels (1000 or more read counts) are not enriched for differential loops; however, they do exhibit a higher static loop overlap (61.8%) consistent with higher total gene expression levels (**Figure 2—figure supplement 4F–I**). For all gene sets examined, looped genes showed strong and similar trends at distal regulatory regions.

Given that only 5% of loops changed significantly during progression (see **Figure 1**), it is not surprising that only a small percentage of differentially expressed genes exhibited significant changes in chromatin contacts with distal enhancers (2.1% of upregulated and 2.2% of downregulated genes; **Figure 2D, F**). This trend was similar between both early- and late-regulated genes (**Figure 2—figure supplement 4J**). On average, there is a slight but significant change in contact frequency between gene promoters and distal enhancers that corresponds to the change in gene expression (**Figure 2E, F**). However, most differentially expressed genes are in regions where chromatin structure is essentially stable, reinforcing that persistent structural changes are not universally required for gene regulation.

For example, the *SPRY1* gene, which regulates cell growth and differentiation and has been shown to be upregulated in triple-negative breast cancer tumors (He et al., 2016), is upregulated between MCF10AT1 and MCF10CA1a, and is statically looped to distal enhancers that gain H3K27ac (**Figure 2H**). Similarly, the *SCNN1G* gene, which encodes for a subunit of a sodium channel and is suppressed in head and neck squamous cell cancer (Yang et al., 2022), is downregulated between MCF10AT1 and MCF10CA1a, and is statically looped to distal enhancers that lose H3K27ac (**Figure 2I**). In both cases, the contact frequency remains relatively constant despite changes in distal enhancer acetylation. Many additional examples of statically looped, differentially expressed genes were found, including *RRM2* and *FERMT2* (**Figure 2—figure supplement 3E, F**). Taken together, our results show that changes in gene expression are not necessarily accompanied by structural changes, and they suggest that stable chromatin loops may facilitate functionally relevant gene expression programs by

providing a pre-existing structure through which differentially regulated distal enhancers can act on target genes.

Changes in TAD boundary insulation and subcompartments have subtle impacts on gene expression

To explore the impact of changes in insulation at domain boundaries on gene expression, we next examined genes within 50 kb of differential boundaries. We found that genes close to weakened boundaries were not enriched for differentially expressed genes, but those near strengthened boundaries were (**Figure 2—figure supplement 5A**; p-values of 0.141 for early and 0 for late strengthened boundaries, and p-values of 1 for weakened boundaries, permutation test). While strengthened boundaries featured both upregulated and downregulated genes, there was a small but significant correlation between the strength of changing boundaries and fold-change of expression of nearby genes when compared to genes at static boundaries, suggesting an average positive impact on gene expression (**Figure 2—figure supplement 5B**; p-value 5.9×10^{-3} , Wilcoxon rank sum test). This suggests that changes in TAD boundary insulation have small but noticeable impacts on gene expression.

We also examined how subcompartments change genome-wide and at gene promoters. We see that between any two cell types, a majority of subcompartment changes are small changes, for example from A.2.2 to A.2.1 (1 step more A-like) or B.1.1 (1 step more B-like), with larger shifts being rarer (**Figure 2—figure supplement 5C**). The promoters of active genes are enriched for A.1.1 and A.1.2 subcompartments but depleted for all B subcompartments, while inactive gene promoters closely resemble genome-wide distributions (**Figure 2—figure supplement 5D, E**). The promoters of differentially expressed genes have similar subcompartments at lower and higher expression levels, but these changes are more drastic for genes that shift from on to off or on to high, as defined above (**Figure 2—figure supplement 5F–H**). Differentially expressed genes with promoters that shift to more B-like by 2 or more subcompartments or more A-like by 3 or more subcompartments have significant impacts on fold-change, but smaller shifts have minimal impacts on gene expression (**Figure 2—figure supplement 5I**; p-value ≤ 0.01 , Wilcoxon rank sum test). In summary, small changes in subcompartments are very common but appear to have little impact on gene expression, while larger changes correlate more strongly with changes in gene expression during breast cancer progression but are comparatively rare.

Changes in enhancer acetylation and enhancer–promoter contact are associated with changes in gene expression

To begin to distinguish the effects of enhancer–promoter contact and chromatin looping from enhancer activity effects, we compared gene expression changes at looped and non-looped enhancer–promoter pairs. To do so, we used the activity-by-contact (ABC) model to predict functional enhancer–promoter pairs. ABC combines estimates of enhancer accessibility and activity from ATAC-Seq and H3K27ac ChIP-Seq with enhancer–promoter contact frequency from Micro-C data to generate an estimate of the likelihood of functional enhancer–gene interactions (see Methods) (**Fulco et al., 2019**). This method allowed us to identify distal regulatory regions that are functionally linked to gene promoters without specifically requiring overlap with chromatin loops. For example, an enhancer and promoter may be in high contact as measured by Micro-C because they overlap with loop anchors, or because they are at close genomic distance. For this analysis, potential enhancers were defined as any overlapping H3K27ac ChIP-Seq and ATAC-Seq peaks 750 bp or farther from the transcription start sites of the target gene.

Applying the ABC model to our data identified 150,056 potential enhancer–promoter pairs across all three cell types. Of these, 53.4% are also promoters of other genes, 23.7% are within the bodies of other genes, and 22.9% are intergenic, and range in distance from 750 to 5 M base pairs away from target gene promoters (**Figure 3—figure supplement 1A**). To better understand the relationship between contact frequency, enhancer activity, and gene expression, we asked how changes in enhancer activity or contact relate to gene expression at target promoters.

Pairwise comparison of the MCF10 progression lines indicated that changes in both contact frequency and enhancer activity appear to drive changes in enhancer–promoter networks predicted by ABC (**Figure 3**). For example, observing potential enhancers with changes in H3K27ac between MCF10CA1a and MCF10A reveals that these enhancers also exhibit a change in contact frequency

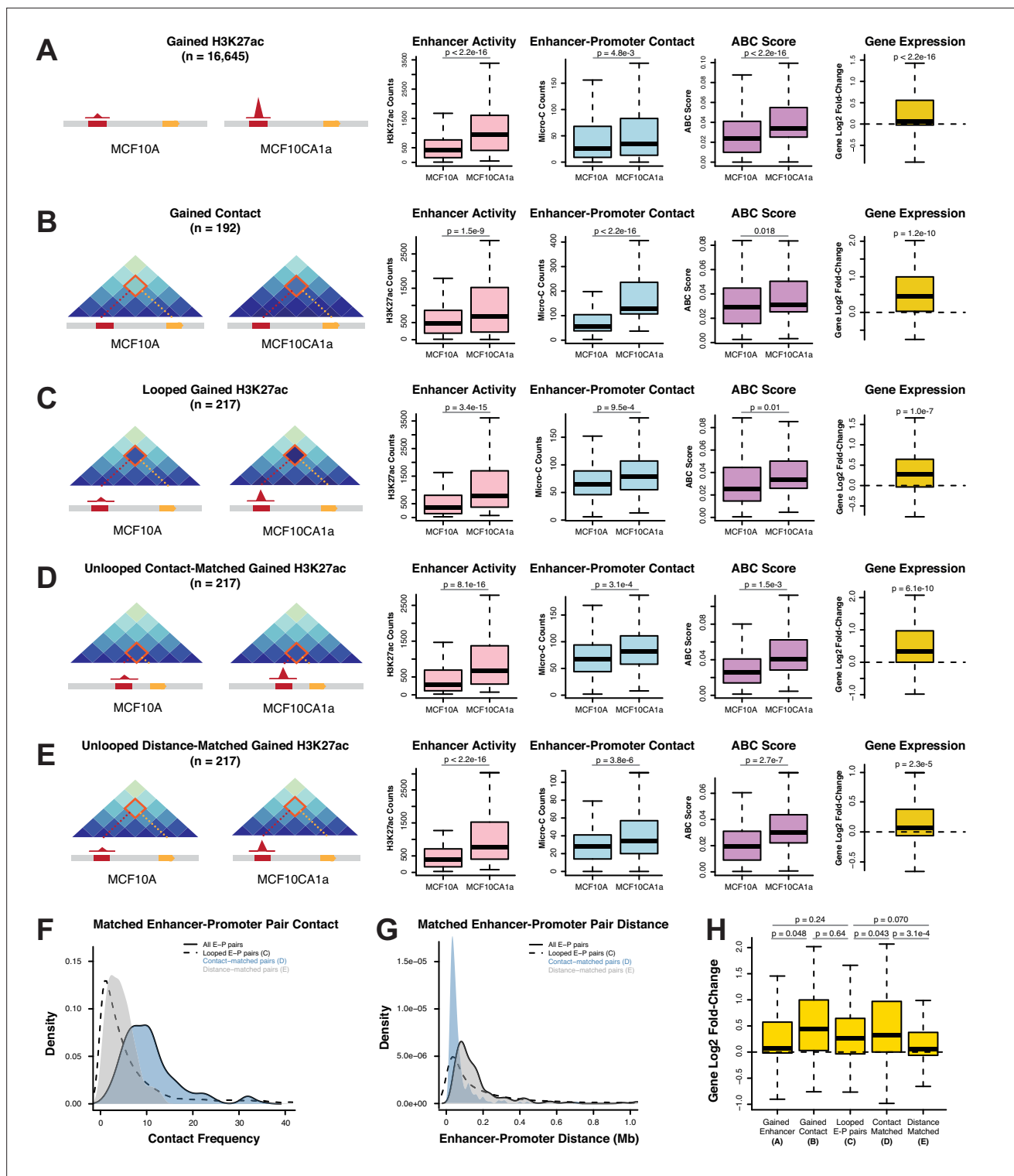


Figure 3. Changes in enhancer acetylation or enhancer–promoter contact are associated with changes in gene expression. Boxplots of distal enhancer H3K27ac (pink), enhancer–promoter contact (blue), and activity-by-contact (ABC) score (purple), as well as gene log₂ fold-change (yellow) for enhancer promoter pairs that feature (A) differential H3K27ac among enhancers, (B) differential enhancer–promoter contact frequency, and (C) differential H3K27ac for enhancer–promoter pairs supported by a chromatin loop. Boxplots in (D) and (E) represent sets of non-looped enhancer–promoter pairs with differential H3K27ac that are matched to the looped set in (C) by contact and distance, respectively. Boxplots show the median (middle line), 25th and 75th quartiles (box perimeters), and range excluding outliers (dashed line whiskers). Outliers are defined as values that are over 1.5 times the interquartile range beyond the box bounds and are excluded from these plots. p-values represent T-tests comparing the means of values in MCF10A

Figure 3 continued

and MCF10CA1a for enhancer activity, enhancer–promoter contact, and ABC score, and *T*-tests comparing the mean of the value to 0 for gene log₂ fold-change. (F) Contact distribution of all enhancer–promoter pairs (dashed line), compared to the looped enhancer–promoter pairs in (C, solid line), the contact-matched pairs in (D, blue shade), and the distance-matched pairs in (E, gray shade). (F) Distance distribution of all enhancer–promoter pairs (dashed line), compared to the looped enhancer–promoter pairs in (C, solid line), the contact-matched pairs in (D, blue shade), and the distance-matched pairs in (E, gray shade). (G) Summary boxplot of the gene log₂ fold-change for the enhancer–promoter pairs previously shown in figures (A–E). *p*-values represent *T*-tests comparing the means of average gene log₂ fold-changes values for different sets of enhancer–promoter pairs.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Enhancer–promoter pair details.

and are associated with upregulation of target genes (**Figure 3A**). We also found that changes in contact frequency are associated with increases in H3K27ac and correlate with higher gene expression (**Figure 3B**). These results show that not only changes in either contact frequency and enhancer activity correlate with increased gene expression, but they also correlate with each other, suggesting a potentially linked functional role during enhancer–promoter communication.

To then relate enhancer–promoter pairs to chromatin loops and to orthogonally assess whether chromatin loops are acting as a functional bridge for active enhancers, we compared looped and non-looped enhancer–promoter pairs. Enhancer–promoter pairs that have changes in distal H3K27ac and are supported by chromatin loops correlated with changes in gene regulation (**Figure 3C**). This effect was stronger than distance-matched non-looped enhancer–promoter pairs, but similar to contact-matched non-looped pairs, suggesting that increased contact frequency caused by loop extrusion may contribute to the stronger correlation with gene expression (**Figure 3D, E**). Contact-matched non-looped pairs were closer on average to the looped pairs of similar contact frequency, while distance-matched non-looped pairs were in less-frequent contact than looped pairs of similar genomic distance (**Figure 3F, G**).

Comparing the distributions of target gene fold-change for these various sets of enhancer–promoter pairs revealed several trends (**Figure 3H**). First, pairs with significant changes in contact have a larger mean gene fold-change than pairs with significant changes in activity, suggesting that either can contribute to changes in enhancer–promoter functional pairing but that contact may have a particularly strong impact. Second, looped enhancer–promoter pairs have a comparable or larger mean gene fold-change to pairs with changes in activity or contact, suggesting again that chromatin loops may support functional enhancer–promoter pairs. Lastly, looped pairs have a similar mean gene fold-change as contact-matched pairs, which in turn have a higher mean gene fold-change than distance-matched pairs, suggesting that the increased contact frequency that chromatin loops provide to enhancer–promoter pairs may be a driving force for the functional connection. These trends hold true for all tested pairwise comparisons between cell types (**Figure 3—figure supplement 1B, C**).

Taken together, these findings demonstrate that not all gene regulatory changes are accompanied by chromatin reorganization, but when it occurs, reorganization appears to facilitate functional changes.

Changes in chromatin looping are enriched for progression-associated differentially expressed genes

We next explored how changes in chromatin looping may functionally contribute to gene regulation during cancer progression. To do so, we compared changes in chromatin loop contacts to alterations in expression of progression-associated genes. Overall, while only a small subset of gene promoters overlaps with the anchors of differential chromatin loops (507 genes, 3.0% of expressed genes), those that do are significantly enriched for genes that are differentially expressed during cancer progression based on permutation analysis (331 differentially expressed genes, 65.3% of all differentially looped genes; **Figure 4—figure supplement 1A**).

We asked whether there was a relationship between the formation or loss of loops and differential expression of these loop-associated genes. We indeed found that differential loops are more likely to change in the same direction as the gene (i.e. increased contact frequency with distal regions associated with increased gene expression) (**Figure 4A, B**). The fold-change of differentially expressed genes which also showed differential looping were significantly higher than a random sampling of differentially expressed genes (**Figure 4A**). For example, of 3261 genes that were differentially upregulated

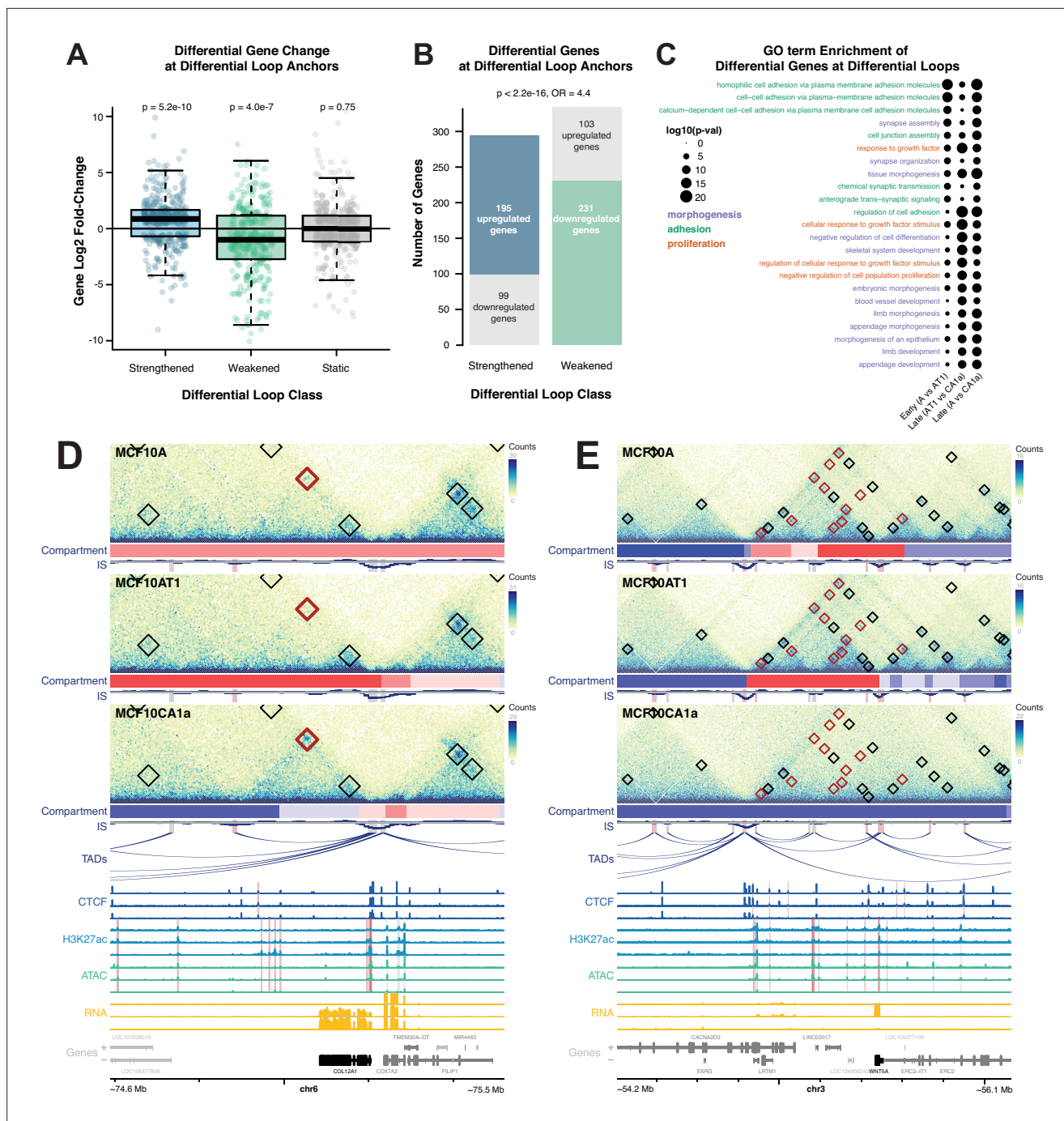


Figure 4. Differential loops are enriched for cancer-relevant differentially expressed genes. **(A)** Log₂ fold-change of differentially expressed genes at the anchors of gained (blue), weakened (green), or static (gray) loops. Boxplots show the median (middle line), 25th and 75th quartiles (box perimeters), and range excluding outliers (dashed line whiskers). Outliers are defined as values that are over 1.5 times the interquartile range beyond the box bounds and are excluded from these plots. p-values represent Wilcoxon tests comparing the mean of each set to 0. **(B)** Bar plot showing the number of differentially expressed genes at strengthened or weakened loop anchors. Bar segments are colored by whether the gene is changing the same (blue for upregulated in strengthened loops, green for downregulated in weakened loops) or opposite (gray) direction as the loop. p-value represents a Fisher's exact test for whether the odds ratio (OR) is greater than 1. **(C)** Gene ontology (GO) term enrichments for genes upregulated in MCF10A, MCF10AT1, or MCF10CA1a. Size indicates p-value. Terms are color-coded based on gene type; morphogenesis (purple), proliferation (orange), and cell adhesion (teal). **(D)** An example of an upregulated gene (COL12A1) with a promoter that overlaps a strengthened loop with distal enhancers. Black boxes show loop annotations, while red boxes indicate differential loops. Red compartment tracks indicate A compartments, while blue indicates B compartments. In CTCF signal tracks, red highlights indicate differential CTCF peaks. In H3K27ac and ATAC-Seq signal tracks, red highlights indicate differential

Figure 4 continued on next page

Figure 4 continued

enhancers as determined by changes in H3K27ac. Genes highlighted in black are differentially expressed. (E) An example of a downregulated gene (WNT5A) with a promoter that overlaps with several weakened loops containing distal enhancers that lose H3K27ac. Plots are annotated as in (A).

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Differential genes within differential loops.

Figure supplement 2. Examples of differentially expressed genes at differential loop anchors.

between MCF10A and MCF10CA1a, loops were significantly strengthened at 98 of these genes and significantly weakened at 31. Similarly, of 3088 downregulated genes, 65 genes overlap weakened loop anchors and 41 genes overlap strengthened loop anchors. In contrast, the number of expected genes at strengthened or weakened loops for a random sampling of genes this size is 38 and 32, respectively (Figure 4—figure supplement 1B, C). We also found a subset of chromatin loops and genes changed in opposite directions (Figure 4B). The genes whose changes in expression correlate with changes in looping are enriched for several cancer-relevant pathways, such as morphogenesis, differentiation, and proliferation (Figure 4C; Hanahan and Weinberg, 2011).

In total, we identified 127 unique genes upregulated in areas that experience increased chromatin contacts, either at loop anchors or within existing structures. As an example, the promoter of the *COL12A1* gene, which encodes a collagen protein known to be associated with aggressive and mesenchymal phenotypes, overlaps a loop boundary that is very weak in MCF10A cells where *COL12A1* is not expressed (Hu et al., 2022; Papanicolaou et al., 2022; Song et al., 2023). As *COL12A1* gene expression is upregulated during progression, contacts increase at a distal region 310 kb away, and H3K27ac and accessibility also increase at these likely distal regulatory regions (Figure 4D).

Similarly, we observe 123 unique genes that are downregulated during oncogenic progression. One example is *WNT5A*, which encodes for an important signaling protein and is downregulated in breast cancer and across MCF10A progression (Borcherding et al., 2015; Roarty et al., 2009; Zeng et al., 2016). Similar to *COL12A1*, the promoter of *WNT5A* is involved in many differential distal regulatory contacts, ranging in distance from 240 to 640 kb (Figure 4E). Unlike *COL12A1*, these contacts are strongest in MCF10A cells and severely weaken in MCF10AT1 and MCF10CA1a cells. Accompanying these changes in contact, the distal regulatory regions that appear to support *WNT5A* in MCF10A cells become deacetylated and decrease in accessibility. Another example of an upregulated gene that is looped to a distal enhancer is *RASA1*, which gains contact frequency as gene expression increases, while *ZEB2* and *SDC3* are downregulated genes that lose enhancer contact frequency similar to *WNT5A* (Figure 4—figure supplement 2A–C). Other loci showed the opposite behavior; for example, *TNFRSF21* and *HS3ST3A1* are upregulated genes at the anchors of weakened loops, while *NNMT* is a downregulated gene at a weakened loop anchor (Figure 4—figure supplement 2D–F).

Genome reorganization occurs at cancer-relevant genes and is accompanied by proximal and distal changes in histone marks

Finally, we aimed to comprehensively explore the genes that are differentially regulated in areas that also have strong changes in chromatin looping, to better understand the potential regulatory mechanisms at play.

A locus-by-locus view of gene and loop fold-change allows us to view the relationship between changes in expression and structure among each pairwise comparison of cells (Figure 5A–C). While we see that a majority of genes have a corresponding change in looping (i.e. upregulated genes overlapping strengthened loops), we observed that the percentage of corresponding changes increases in the later stages of cancer progression. For example, the percentage of differential loop-gene pairs where the gene overlaps at least one gained loop is 47.5% and 69.9% among genes upregulated in MCF10A compared to MCF10AT1 and MCF10CA1a, respectively, 66.7% and 79.6% among genes upregulated in MCF10AT1 compared to MCF10A and MCF10AT1, respectively, and 70.7% and 58.0% among genes upregulated in MCF10CA1a compared to MCF10A and MCF10AT1, respectively. This may indicate that the regulatory impacts of changes in chromatin looping occur over longer timescales, or that genes impacted by changes in chromatin structure may be more common in metastatic

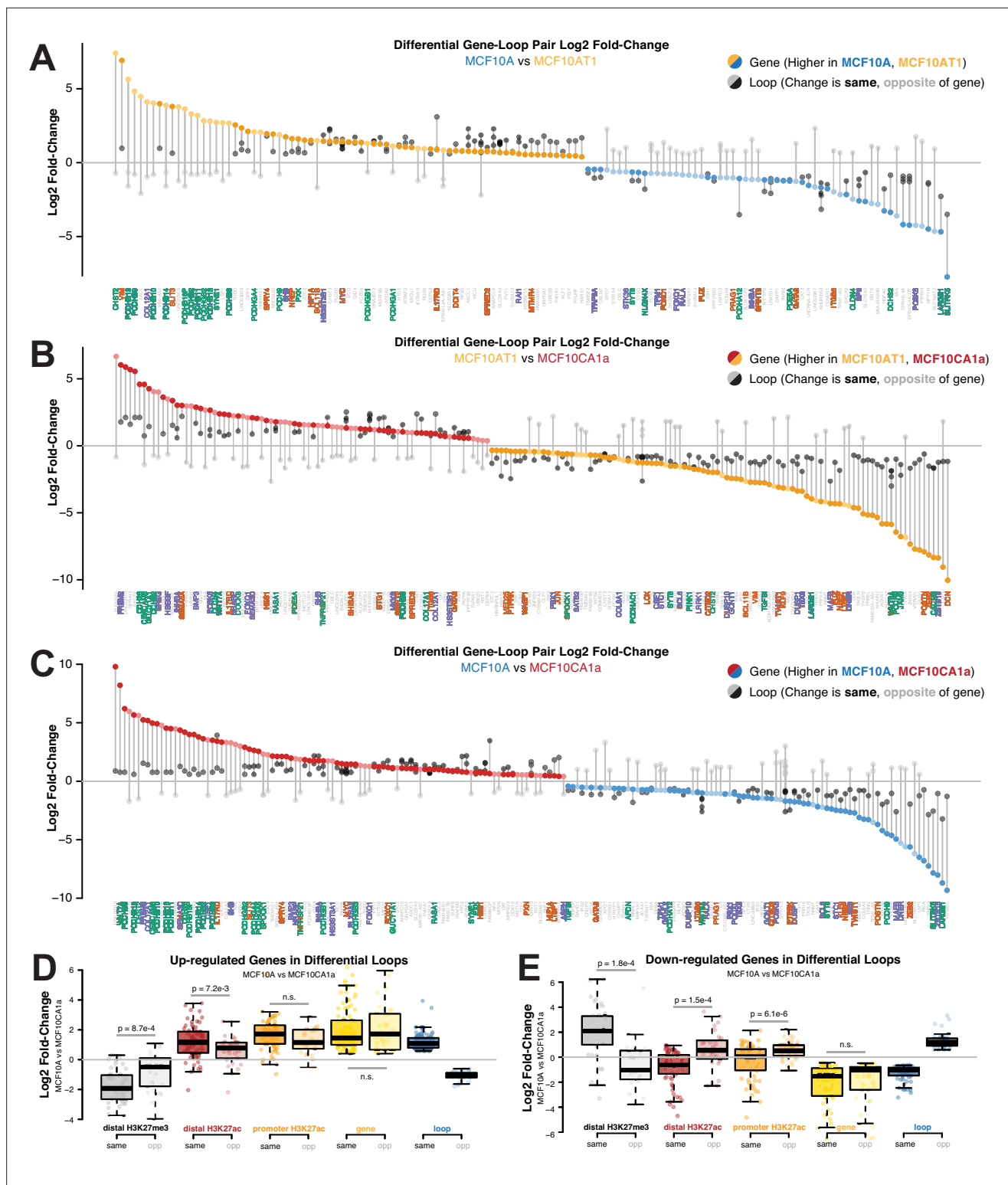


Figure 5. Progression-associated differentially expressed genes exhibit local and distal epigenetic changes at differential loops. Log₂ fold-change of genes (colored dots) and the differential loops they overlap with (black/gray dots) for genes and loops that change between (A) MCF10A and MCF10AT1, (B) MCF10AT1 and MCF10CA1a, and (C) MCF10A and MCF10CA1a. Gene labels are below. (D) Log₂ fold-change between MCF10A and MCF10CA1a of distal H3K27me3 (gray), distal H3K27ac (red), promoter H3K27ac (orange), gene expression (yellow), and loops (blue) among upregulated genes that overlap with gained loops (darker colors) or lost loops (lighter colors). Boxplots are defined as in (A). p-values represent T-tests comparing the mean values of the features at loops that change in the same and those that change in opposite directions from the differential genes at

Figure 5 continued on next page

Figure 5 continued

their anchors. Non-significant (n.s.) p-values are any p-values above 0.01. (E) Log₂ fold-change of distal H3K27me3 (gray), distal H3K27ac (red), promoter H3K27ac (orange), gene expression (yellow), and loops (blue) among downregulated genes that overlap with gained loops (darker colors) or lost loops (lighter colors). p-values represent T-tests comparing the mean values of the features at loops that change in the same and those that change in opposite directions from the differential genes at their anchors. Non-significant (n.s.) p-values are any p-values above 0.01.

cells. We did not, however, find any correlation between the magnitude of loop fold-change and gene fold-change.

To better understand how changes in looping are related to gene expression, we compared patterns of gene expression, promoter H3K27 acetylation, and distal enhancer H3K27 acetylation or trimethylation at looped genes that change in either the same or opposite direction as the loops they overlap (**Figure 5D, E, Figure 4—figure supplement 1D, E**). These two histone H3 modifications are mutually exclusive and have opposite effects on gene expression, marking enhancers and silencers, respectively (**Pasini et al., 2010**). Both modifications are able to impact distal gene expression via chromatin interactions (**Cai et al., 2021**).

Upregulated genes between MCF10A and MCF10CA1a showed similar epigenetic signatures at both proximal and distal regions, regardless of whether the loop they overlap gets stronger or weaker (**Figure 5D**). Promoter regions gained H3K27ac, while distal regions gained H3K27ac and lost H3K27me3. There is, however, a significant difference in the extent of distal changes depending on the loop direction, with strengthened loops exhibiting a significantly higher increase in distal H3K27ac and a decrease in H3K37me3 marks. This behavior further supports the notion that these distally looped regulatory regions are important functional elements. Downregulated genes showed distinctly different signatures (**Figure 5E**). Genes that overlap with weakened loop anchors showed decreased promoter H3K27ac and distal H3K27ac and increased distal H3K27me3, consistent with signatures typically associated with reduced gene expression (**Pasini et al., 2010**). Interestingly, genes that overlap strengthened loop anchors showed different patterns, with a gain in promoter H3K27ac and loss of distal H3K27me3 repressive marks. We conclude that expression of a subset of progression-associated genes is strongly correlated with loop formation. These trends are similar but weaker for genes that change between different cell types (**Figure 4—figure supplement 1D, E**).

Taken together, our genome-wide analysis of structural and regulatory changes during MCF10A cancer progression has highlighted hundreds of restructured regions where cancer-relevant genes are differentially regulated. These findings suggest that, while relatively rare, changes in chromatin looping may be associated with regulatory changes that drive expression of hundreds of oncogenes during cancer progression.

Discussion

Multiple levels of chromatin organization and integration with epigenetic parameters contribute to regulation of gene expression. We have identified dynamic chromatin organizational changes on multiple scales during breast cancer progression using the well-characterized MCF10A model system. By comparing both a pre-malignant and metastatic cell line to non-cancerous epithelial cells we were able to detect both early- and late-stage changes as well as similarities in genome structure and relate them to gene expression changes, including many cancer-relevant genes.

We found that compartmental shifts occur more often in the early stages of cancer development. This behavior is consistent with previous studies that have shown intermixing of chromatin A and B compartments in cancer cells (**Johnstone et al., 2020; Rossini et al., 2024**). The general shift towards more active compartments during cancer progression may reflect a broader epigenetic landscape and greater heterogeneity in gene expression within cancer cell populations (**Corces et al., 2016; Liu et al., 2025; Guo et al., 2019; Carter and Zhao, 2021; Flavahan et al., 2017**).

On a finer scale, structural changes in TADs occurred more often during later stages of metastasis. We also found an abundance of weakened TAD boundaries in MCF10 breast cancer progression compared with boundaries that gain insulation. This finding is in line with previous studies, which have shown TAD weakening in triple-negative breast cancer (**Kim et al., 2022; van den Brand et al., 2025**), but is contrary to observations in prostate cancer, which have suggested the formation of many additional TADs (**Taberlay et al., 2016**), or findings of minimal changes in

TAD structure in colorectal or breast cancer (*Johnstone et al., 2020; Rossini et al., 2024*). The differences in TAD changes across cancer studies may reflect differences in the types of cancer samples analyzed, or the analysis methods used. For example, we detected subtle changes in IS at TAD boundaries in the MCF10 model, while other methods tailored toward detecting more drastic restructuring events, such as the formation of TAD cliques or TAD fusions, found different behaviors (*Rossini et al., 2024; Choppavarapu et al., 2025*). Recent studies comparing chromatin structures across patient samples have also revealed a high degree of TAD and loop heterogeneity among tumors, even of the same cancer type, and it is thus not surprising that patterns observed between different cancer types and in a cell culture model would vary even further (*Choppavarapu et al., 2025*). The weakened TAD boundaries in the MCF10 progression system are typically strong boundaries in cancer types, suggesting that the same boundaries are not weakened universally in cancer. The weakening of TAD boundaries is intriguing in that it may point to higher heterogeneity in chromatin structure in more aggressive cancer stages, possibly contributing to more extensive gene misregulation (*Liu et al., 2025; van den Brand et al., 2025; Choppavarapu et al., 2025*). Further studies using single-cell analysis in the MCF10CA1a population will be necessary to confirm if this is the case.

Comparing the structural and functional features of the MCF10 progression system to other breast cancer datasets revealed many common chromatin loops, TADs, and differentially expressed genes, as well as differences between various cancer progression models and patient data. There are hundreds of genes that change over the course of MCF10 progression that are also significantly different in breast cancer tumors from patients compared to healthy controls, including many that overlap with chromatin loops; however, this represents only roughly a quarter of all differential genes in the MCF10 progression model. Despite this transcriptional diversity, chromatin loops and TAD boundaries detected in MCF10 cell lines showed similar average signatures in mammary epithelial cell lines and triple-negative breast cancer patients. However, cell type-specific differences do appear among differential features based on their timing and direction of change in the MCF10 progression model. For example, loops that are stronger in metastatic MCF10CA1a cells were also significantly stronger than static MCF10 loops among TNBC patients, making them a particularly interesting subset of regions to explore further.

Chromatin loops functionally connect gene promoters to distal regulatory elements. In the MCF10 model, many genes differentially regulated during cancer progression are associated with chromatin loops shared between all three cell lines, but which show changes in distal enhancer H3K27ac and H3K27me3. These trends were stronger with upregulated genes, suggesting that we may need to explore different epigenetic signatures to better understand how chromatin structure may influence repression. We also observed that more highly expressed genes are more likely to overlap with chromatin loops. ABC analysis showed that looped enhancer–promoter pairs also exhibit greater correlation between distal enhancer H3K27ac and gene expression than non-looped enhancers due to their increased contact frequency. These findings suggest that persistent chromatin loops that do not change during cancer progression nevertheless have functional relevance and that they do so by bridging enhancers to target gene promoters.

ABC analysis also revealed that subtle changes in chromatin contact can contribute to the rewiring of enhancer–promoter regulatory connections. Enhancer–promoter pairs that exhibit changes in contact correlated with stronger changes in target gene expression than those with only changes in activity, in line with the concept that contact with distal regulatory elements is an important component of gene regulation (*Yang and Hansen, 2024*). We also found that changes in chromatin contact are associated with more modest changes in activity, and vice versa. This correlation between enhancer activity and enhancer–promoter contact further points to a functional link between the two. Together, these results suggest that both contact frequency and activity contribute to enhancer–promoter connectivity.

Strong changes in contacts at chromatin loops are relatively rare across cancer progression, but many clear examples do exist, and they are notably enriched for differentially expressed genes. Interestingly, only a small portion of differential loops can be explained by changes in CTCF binding at anchors, suggesting other forces may be influencing their contact frequency. Loss of H3K27ac can be seen at the anchors of lost loops, consistent with the idea that active enhancers can help recruit cohesin to chromatin (*Georgiades et al., 2023; Rinzema et al., 2022*). Therefore, some weakened loops might be explained by a loss of H3K27ac, which leads to a loss of cohesin at that region.

We also found that genes at differential loops are more likely to be differentially regulated in the same direction as the change in loop strength. This supports the notion that a subset of genes may be regulated in a structure-dependent manner (Choppavarapu et al., 2025; Reyes-Gopar et al., 2025; Yost et al., 2025). This interpretation is in line with previous observations, which have shown a subset of genes to be sensitive to cohesin or NIPBL depletion, which disrupts chromatin loops (de Wit and Nora, 2023; Chakraborty et al., 2024; Xiao et al., 2021; Stolper et al., 2023; Cuartero et al., 2018; Guckelberger et al., 2024; Kubo et al., 2021; Thiecke et al., 2020). Importantly, these findings suggest this subset of structure-sensitive genes may include many with relevance to breast cancer progression.

Epigenetic signatures at gene promoters and distal regions differed based on the direction of gene change. Upregulated genes consistently showed a gain of active-associated H3K27ac marks at both promoters and distal regions, and a loss of distal repressive H3K27me3 marks. These changes were shared between static, strengthened, and weakened loops, although upregulated genes at strengthened loops had stronger distal changes. These findings are consistent with loops functionally supporting interactions with distal enhancers via increased contact. However, downregulated genes showed more complex patterns. Fewer downregulated genes could be explained by changes in H3K27ac or H3K27me3 at the promoter or distal regions, with over half of downregulated genes having no clear epigenetic driver, compared to only 15% of upregulated genes. This suggests that these histone marks are not sufficient to explain downregulated genes as well as they can explain upregulated genes. Downregulated genes at differential loops also showed opposite patterns based on the direction of loop change; weakened loops showed loss of distal H3K27ac and gain of H3K27me3 consistent with an inactivated enhancer, while strengthened loops showed the opposite. Additional studies will be required to fully understand the repressive mechanisms in this system and how they relate to chromatin structure.

Our study has several limitations. While the MCF10 progression system is a well-established and -characterized model, it does not reflect the complexities of in vivo tumor formation. In addition, the MCF10 cell lines used here do not fully represent all stages of cancer progression as they encompasses normal, pre-malignant, and metastatic states, but lack a separate malignant but non-metastatic state such as MCF10DCIS. Furthermore, our analysis is correlative in nature and does not demonstrate causality of chromatin structure changes at the loci of differentially expressed genes. Follow-up studies exploring any of the gene loci identified here as possibly being contact-dependent could help elucidate these relationships. Additionally, as a population-level assay, Micro-C is only able to provide aggregate data across an entire population of cells. To address how cell-to-cell heterogeneity contributes to some of the functional relationships we observe, and whether that heterogeneity is occurring at a cellular or population level, we would need to apply single-cell sequencing or imaging-based approaches. These questions will be the subjects of future studies.

In summary, using the MCF10 breast cancer model, we have generated a rich genomic dataset of structural and functional changes in the genome during breast cancer progression. Our data uncovers new insights into the structure-function relationship in gene regulation and into the role of genome organization during malignant breast cancer progression.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line (<i>Homo sapiens</i>)	Cell line MCF10A	ATCC	Cat#:CRL-10317; RRID:CVCL_0598	Non-cancerous mammary epithelial cells
Cell line (<i>Homo sapiens</i>)	Cell line MCF10AT1	Karmanos Cancer Institute	ID:MCF10AneoT; RRID:CVCL_5554	Pre-malignant mammary epithelial cells
Cell line (<i>Homo sapiens</i>)	Cell line MCF10CA1a	Karmanos Cancer Institute	ID:MCF10CA1a.cl1; RRID:CVCL_6676	Metastatic mammary epithelial cells
Antibody	anti-human CTCF (rabbit monoclonal)	Cell Signaling Technology	Cat#:3418	ChIP-Seq (20 µl per 150 µg chromatin)

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	anti-human H3K27ac (rabbit polyclonal)	AbCam	Cat#:ab4729; RRID:AB_2118291	ChIP-Seq (2 µg per 10 µg chromatin)
Antibody	anti-human H3K27me3 (mouse monoclonal)	AbCam	Cat#:ab6002; RRID:AB_305237	ChIP-Seq (4 µg per 15 µg chromatin)
Commercial assay or kit	Dovetail Micro-C Kit & Assay Service	Dovetail Genomics	Cat#:20101E	Loop level project
Commercial assay or kit	Direct-zol RNA Miniprep	Zymo Research	Cat#:R2050	
Commercial assay or kit	NEBNext rRNA Depletion Kit v2	New England Biolabs	Cat#:7400	
Commercial assay or kit	NEBNext Ultra II RNA-Seq Library Prep Kit	New England Biolabs	Cat#:E7770	
Software, algorithm	BWA mem	PMID:19451168	RRID:SCR_010910	Version 0.7.17
Software, algorithm	Pairtools	https://doi.org/10.1101/2023.02.13.528389	RRID:SCR_023038	Version 0.3.0
Software, algorithm	samtools	PMID:33590861	RRID:SCR_002105	Version 1.17
Software, algorithm	Juicer tools	PMID:27467249	RRID:SCR_017226	Version 1.22.01
Software, algorithm	NeoLoopFinder	PMID:34092790		
Software, algorithm	CALDER	PMID:33972523		Version 2.0
Software, algorithm	FAN-C	PMID:33334380		Version 0.9.21
Software, algorithm	SpectralTAD	PMID:32689928		Version 1.18.0
Software, algorithm	SIP	PMID:32127418		Version 1.6.2
Software, algorithm	STAR	PMID:23104886	RRID:SCR_004463	Version 2.4
Software, algorithm	Trim Galore	https://doi.org/10.5281/zenodo.5127898	RRID:SCR_011847	
Software, algorithm	Bowtie2	PMID:22388286	RRID:SCR_016368	
Software, algorithm	MACS2	PMID:18798982	RRID:SCR_013291	
Software, algorithm	DESeq2	PMID:25516281	RRID:SCR_015687	Version 1.42.0
Software, algorithm	CSAW	PMID:26578583	RRID:SCR_026738	
Software, algorithm	GEPIA	PMID:40396370	RRID:SCR_028242	
Software, algorithm	GenomicRanges	PMID:23950696	RRID:SCR_000025	
Software, algorithm	mariner	PMID:38814811		
Software, algorithm	plotgardener	PMID:35134826		
Software, algorithm	Activity-by-contact	PMID:31784727		
Software, algorithm	nullranges	PMID:37084270		
Software, algorithm	GSEA	PMID:16199517	RRID:SCR_003199	
Software, algorithm	HOMER	PMID:20513432	RRID:SCR_010881	

Cell culture

MCF10A cells were obtained from AATC (CRL-10317). MCF10AT1 (MCF-10AneoT) and MCF10CA1a (MCF10CA1a.cl1) cells were obtained from the Karmanos Cancer Institute. All cells were received frozen, thawed in media, and grown to a confluence of 80% after two to five passages. Stock solutions were frozen down to be used for all experiments. The identity of each cell line has been authenticated using short tandem repeat profiling and cell lines were tested periodically for the absence of mycoplasma.

MCF10A and MCF10AT1 cells were cultured in standard high calcium medium with growth factors, consisting of DMEM/F12 (Invitrogen, 21041025) supplemented with 1.05 mM CaCl₂, 10 mM HEPES, 10 µg/ml insulin (Sigma, I1882), 20 ng/ml EGF (PeproTech, AF-100-15), 0.5 µg/ml Hydrocortisone (Sigma, H0135), 100 ng/ml cholera toxin (Sigma, C8052), and 5% horse serum (Invitrogen, 16050122). MCF10CA1a cells were cultured in standard high calcium medium without growth factors, consisting of DMEM/F12, 1.05 mM CaCl₂, 10 mM HEPES, and 5% horse serum.

Karyotypic analysis

Metaphase chromosomes were prepared by incubating cells with 100 ng/ml Colcemid (Roche, Brighton, MA) for 2 hr, followed by mitotic shake-off. The mitotic cells were then treated with a hypotonic solution (0.075 M KCl) for 20 min at 37°C. After this treatment, the cells were centrifuged, the supernatant was extracted, and the cells were initially fixed with a methanol/acetic acid solution (3:1). This step was repeated three times. Finally, the cells were fixed onto a slide using a humidity-controlled Thermotron (Thermotron, Holland MI).

SKY probes were purchased from Applied Spectral Imaging (Carlsbad, CA) and hybridized onto slides that were aged at 37°C for 3 days. Detection was then carried out according to the protocol provided by Applied Spectral Imaging, using the following purchased antibodies: Avidin Cy5 (Rockland Immunochemicals, Limerick, PA), Mouse Anti-Digoxin antibody (Sigma-Aldrich), and a goat anti-mouse antibody conjugated to Cy5.5 (Rockland Immunochemicals). Slides were then mounted and counterstained with DAPI (Vector Laboratories, Newark, CA). Hybridization occurred over a period of two days at 37°C. Approximately 20–25 metaphases were imaged and karyotyped using the ASI GenASIS 8.2.2 software on an Olympus BX63 microscope (Evident, Tokyo, Japan) equipped with a Spectral Cube (Applied Spectral Imaging, Carlsbad, CA).

Composite karyotypes for the three cell lines, listing abnormalities found in three or more metaphases, are as follows:

MCF10A: 41–53, XX, idic(1)(q10), +1, der(3)t(3;9)(p21;p22), i(8)(q10), der(9)t(3;5;9)(p14;p23;q34;p21)

MCF10AT1: 42–49, XX, der(3)t(3;9)(p21;p22), t(3;17)(p14;p11.2), t(6;19)(p25;q12q13.3), der(9)t(3;5;9)(p14;p23;q34;p21), t(19;9)(q34;p11), +20

MCF10CA1a: 43–63, XX, der(3)t(3;9)(p21;p22), +der(3)t(3;9)(p21;p22), t(3;17)(p14;p11.2), der(9)t(3;5;9)(p14;p23;q34;p21), +10

Micro-C library preparation

Cells were frozen and pellets of 1 M cells were used for Micro-C library preparation. The Micro-C library was prepared using the Dovetail Micro-C Kit according to the manufacturer's protocol. Briefly, the chromatin was fixed with 0.3 M disuccinimidyl glutarate and 37% formaldehyde in the nucleus. The cross-linked chromatin was then digested in situ with micrococcal nuclease (MNase). Following digestion, cells were lysed with 20% SDS to extract the chromatin fragments, and the chromatin fragments were bound to Chromatin Capture Beads. Next, the chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of adapter-containing ends. After proximity ligation, the crosslinks were reversed, the associated proteins were degraded, and the DNA was purified, then converted into a sequencing library using Illumina-compatible adaptors. Biotin-containing fragments were isolated using streptavidin beads prior to PCR amplification. Each library was sequenced on an Illumina Novaseq platform to generate 240–340 million 2 × 150 bp read pairs (average depth of 280 M reads). Micro-C libraries were prepared by Dovetail Genomics (Scotts Valley, CA). Four pellets and subsequent Micro-C libraries (technical replicates) were prepared per culture of cells (biological replicates), with two biological replicates grown per cell line on separate weeks.

RNA-Seq library preparation

Total RNA was isolated from cells using Trizol (Life Technologies) and purified using the Direct-zol RNA kit (Zymo Research, Irvine, CA, USA: R2050). RNA quality and quantity were assessed using the RNA 6000 Nano Kit with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA quantity was further assessed using a Nanodrop2000 (Thermo Scientific, Lafayette, CO) and Qubit HS RNA assay (Thermo Fisher Scientific). Total RNA was depleted of ribosomal RNA (New England Biolabs,

NEB #7400), reverse transcribed using a random hexamer strategy, and strand-specific adapters were added following the NEBNext Ultra II RNA-seq library prep kit (New England Biolabs, E7770). Paired-end sequencing was used to generate high depth coverage ranging from 80 to 120 million paired-end reads. Three biological replicates were generated per cell line from independent cultures and library preparations.

ChIP-Seq library preparation

ChIPseq for CTCF (Cell Signaling Technology, catalog number 3418) and histone marks H3K27ac (Abcam, ab4729) and H3K27Me3 (Abcam ab6002). Independent biological replicates of each cell line (MCF10A, MCF10AT1, and MCF10CA1a) were performed as described, including the optional step of snap freezing of chromatin and excluding the optional third washing step (O'Geen *et al.*, 2010). Additionally, the chromatin was precleared with protein A/G beads for 2 hr at 4°C prior to incubation with antibodies. For CTCF, we used 20 µl of antibody (Cell Signaling Technologies, 3418S) and 150 µg of chromatin for each sample. For H3K27ac, 10 µg of chromatin was used and 2 µg of antibody, while for H3K27me3, 15 µg of chromatin was used with 4 µg of antibody. Two biological replicates were generated per cell line from independent cultures and library preparations.

ATAC-Seq library preparation

The OMNI-ATACseq protocol was followed as previously described, with an optimized 5 min of nuclear extraction rather than 3 min (Maor-Nof *et al.*, 2021; Corces *et al.*, 2017). Two biological replicates were generated per cell line from independent cultures and library preparations.

Micro-C processing

Micro-C data from each technical replicate (library) was processed from raw reads into contact maps using guidelines outlined by Dovetail Genomics (Dovetail, 2025). Paired reads were aligned to the hg38 human genome assembly (NCBI GRCh38) using BWA mem (version 0.7.17; settings: -5SP -T0) (Li and Durbin, 2009). Pairtools (version 0.3.0) was then used to parse ligations, sort reads, and remove PCR duplicates using the parse (settings: --min-mapq 40 --walks-policy 5unique --max-inter-align-gap 30), sort, and dedup (settings: --mark-dups) commands (Abdennur *et al.*, 2023). Alignment files were generated using pairtools split to create .pairs and .sam files, and samtools view (version 1.17; settings: -bS) to create .bam files (Danecek *et al.*, 2021). Final .bam alignment files were sorted and indexed using samtools sort and index commands. The .bam files were used to extract QC metrics using a script from Dovetail Genomics, and calculate complexity using preseq lc_extrap (settings: -bam -pe -extrap 2.1e9 -step 1e7 -seg_len 1000000000). Pairs files were used to generate contact maps using juicer_tools pre (version 1.22.01), and normalized using SCALE (Durand *et al.*, 2016).

For merged contact maps, we first merged pairs files using pairtools merge, and then ran juicer_tools pre command on the resulting .pairs files. In total, we generated contact maps for each library (technical replicates), each sample (biological replicates), each cell type, and we created one fully merged map including all reads for a precise map of shared features. Micro-C maps for individual technical replicates and cell types are available on GEO (accession GSE254045).

Micro-C copy number variation analysis

SV and copy number variations were detected from contact maps using NeoLoopFinder calculate-cnv at 1-Mb resolution (-e uniform; *Supplementary file 2; Wang et al.*, 2021). These values were averaged over the course of large-scale variations and used as correction factors for differential loop analysis (see Differential loop, TAD, gene, and peak analysis).

Regions excluded in analysis

Denylists of regions to ignore for feature calling were generated based on regions where SCALE normalization factors were unable to be calculated at 5 or 10 kb, ignoring single bins and merging continuous areas within 100 kb. ABC analysis also combined this denylist with the ENCODE denylist for hg38 (Amemiya *et al.*, 2019).

Micro-C compartment calling

Compartments were called using CALDER (version 2.0; R version 4.2.1) at 10-kb resolution and SCALE normalization (*Liu et al., 2021*). We also used FAN-C (version 0.9.21) to calculate eigenvectors at 100 kb using SCALE normalization to build saddle plots, and oriented eigenvectors manually based on overlap with active genes (*Kruse et al., 2020*). A table of subcompartments for each cell type is available on GEO (accession GSE254045).

Micro-C topologically associated domain calling

Topologically associated domains (TADs) were called using SpectralTAD (version 1.18.0) at 10-kb resolution with SCALE normalization, a window size of 300, and a level of 3, with a quality filter applied (*Cresswell et al., 2020*). We then created a unified TAD list by merging TADs that overlap at both ends between cell types, with a maximum gap of 10 kb (1 pixel). We then combined this analysis with continuous IS calculations from FAN-C insulation command at 10-kb resolution with SCALE normalization. We then used the FAN-C boundaries command to detect IS boundaries and only kept TADs that also overlapped with an IS boundary. Tables of TADs and boundaries are included in **Supplementary files 4 and 5**. IS tracks for each cell type are available on GEO (accession GSE254045).

Micro-C chromatin loop calling

Chromatin loops were called using SIP (version 1.6.2), run at 5 and 10 kb and then merged to 10 kb (-g 2 -fdr 0.05). Loops were called in each cell type map using SCALE normalization, in addition to the combined map, and then merged similarly to TADs to create a unified loop list (**Supplementary file 3; Rowley et al., 2020**). Cell type loops are available on GEO (accession GSE254045).

RNA-Seq processing

All RNA-Seq data was analyzed using the nf-core/rnaseq pipeline (*Patel et al., 2020*). Adapter and quality trimming was performed with Trim Galore (*Babraham Bioinformatics, 2017*). Reads were aligned to the hg38 reference genome using STAR (*Dobin et al., 2013*), and gene expression was quantified using Salmon (*Patro et al., 2017*). Differentially expressed genes were called using DESeq2 (*Love et al., 2014*). An adjusted p-value of 0.05 and a log₂ fold change of 1 were used as thresholds to select significant differential expression. A table of all differentially expressed genes and their DESeq2 metrics are listed in **Supplementary file 6**. A full table of expression data for all genes can be found on GEO (accession GSE320216).

ChIP-Seq processing and peak calling

ChIP-Seq data for CTCF, H3K27ac, and H3K27me3 was processed as detailed previously (*Fritz et al., 2018; Ghule et al., 2023*). In summary, adapters were cut (cutadapt v1.11), and low-quality reads trimmed (Galaxy FASTQ Quality Trimmer 1.0.0; window 10, step 1, minimum quality 20). Reads were mapped to the human genome (hg38 canonical) using STAR version 2.4 (*Dobin et al., 2013*) with splicing disabled (-alignIntronMax 1). Enriched regions (narrowPeak calls for CTCF and H3K27me3, broadPeak calls for H3K27ac) for each replicate were generated using MACS2 (*Feng et al., 2012*), and replicates were then evaluated using deepTools (*Ramírez et al., 2016*) to correlate alignments and IDR (*Li et al., 2011*) to evaluate peak call reproducibility. After pooling replicates, MACS2 (*Zhang et al., 2008*) was used to call peaks at high stringency (p-value <10e-5); these peaks were further filtered according to IDR cutoffs. ChIP-Seq data is available on GEO (accession GSE98551 for CTCF, and GSE229295 for H3K27ac and H3K27me3).

ATAC-Seq processing and peak calling

Read trimming and quality filtering was performed using FastQC (*Bioinformatics, 2015*) and TrimGalore (*Babraham Bioinformatics, 2017*). The filtered fastq were then downsampled to approximately 50 million reads per replicate. Reads were aligned to the hg38 reference genome using Bowtie2 (*Langmead and Salzberg, 2012*). Mitochondrial, multi-mapped, and low-quality reads, and duplicates were removed using samtools (*Danecek et al., 2021*). MACS2 (*Zhang et al., 2008*) was used to call narrowPeaks, followed by IDR (*Li et al., 2011*) to generate sample level peak sets. ATAC-Seq data is available on GEO (accession GSE320215).

Enhancer and promoter definitions

Gene promoters were defined as regions between 2000 bp upstream and 500 bp downstream of gene transcription start sites.

Potential enhancer regions were identified based on regions that contained both an ATAC-Seq and H3K27ac ChIP-Seq peak. For ABC analysis, potential enhancers were defined as 150,000 ATAC-Seq peaks with the highest levels of H3K27ac signal, but were subset for regions with H3K27ac peaks after running (see below).

Compartmental saddle plots

Saddle plots were made manually in R. We selected three chromosomes that had no major karyotypic differences between cell lines and had high correlation of eigenvectors between replicates with the same signage (chr2, chr12, and chr17). For each chromosome and cell type, we sorted eigenvectors into 20 bins. We then calculated the mean observed/expected values (using SCALE normalization) between regions belonging to different bins and plotted it as a heatmap.

Differential loop, TAD, gene, and peak analysis

Differential genes were calculated using DESeq2 (version 1.42.0) (Love *et al.*, 2014). Each cell type had three replicates, and a design of ~cellType was used. No fold-change cutoff was applied. Genes with an adjusted p-value below 0.01 were considered significant.

Differential H3K27ac within ATAC-Seq peaks were calculated using a similar design, but with a p-value cutoff of 0.05.

Differential CTCF peaks were called using the CSAW package following the workflow described in the original publication (Fritz *et al.*, 2018), and an adjusted p-value cutoff of 0.1.

Differential loops were also identified using DESeq2, but with additional adjustments. Raw, un-normalized loop counts were pulled from each technical replicate map, resulting in 8 values per cell type (4 technical replicates for each of 2 biological replicates per cell type). An LRT design was used of ~technicalRep + biologicalRep + cellType, with ~technicalRep + biologicalRep as the comparison. Size factors were provided manually based on NeoLoopFinder output (Supplementary file 2). Regions of duplication were identified based on NeoLoopFinder CNV output and spectral karyotypes. For each duplicated region, the CNVs across the region were averaged and used as custom size factors to correct for differences in chromosomal copy number between cell lines. Differential loops were identified based on a fold-change cutoff of 1.5 and an adjusted p-value cutoff of 0.025.

Differential TAD boundaries were detected using an alternative method. ISs were pulled from all TAD boundaries at the technical replicate level (8 values per cell type). A T-test was then applied for each pairwise comparison of cell types, and p-values were adjusted using FDR. Boundaries with an adjusted p-value below 0.01 were considered significant.

Differential loops, TAD boundaries, and genes were clustered based on their patterns of change across the three cell types using k-means clustering of centered and scaled normalized counts.

Integration of Cancer Genome Atlas data

RNA-Seq data from the breast cancer (BRCA) cohort of the Cancer Genome Atlas was used to analyze whether genes differentially expressed in the MCF10 progression model are also differentially expressed in breast cancer tumor tissues ($n = 1102$) from patients compared to normal control tissues ($n = 113$). Normalized gene transcript counts from RSEM were downloaded from Xena (Goldman *et al.*, 2020). Protein coding genes that are differentially expressed between tumors and paired peritumors in TCGA-BRCA data were identified using GEPIA3 using DESeq2 with a $\log_2$ fold-change cutoff of 1 and an adjusted p-value cutoff of 0.05. A total of 5289 differential genes were identified from the TCGA-BRCA data: 3168 with significantly higher expression in tumor samples, and 2121 with significantly higher expression in normal samples (Tang *et al.*, 2017). Overall survival curves for select genes were generated using GEPIA (Tang *et al.*, 2017).

Integration of additional breast cancer structural data

Chromatin loops and TAD boundaries identified in the MCF10 progression model were also examined in other triple-negative breast cancer cell lines and patient samples using data from GEO (accession GSE167150) (Kim *et al.*, 2022). IS tracks were calculated as for MCF10 cell lines, using FAN-C at

10-kb resolution and KR normalization. Chromatin loop counts were extracted using KR normalization at 10-kb resolution.

Micro-C feature overlap and aggregate analysis

Micro-C feature overlap and analysis was conducted in R primarily using the *GenomicRanges*, *InteractionSet*, and *mariner* packages (Lawrence et al., 2013; Lun et al., 2016; Davis et al., 2024). To overlap chromatin loops with other features such as promoters and enhancers, we used loop anchors at 10-kb resolution and allowed features within 10 kb (± 1 Hi-C pixel) of the loop anchors to be considered overlapping. We used this broad definition for loop overlap to account for the fact that loops are non-punctate and show increased contact frequency within several bins on either side of the called loop pixel, as evidenced by both individual loci and aggregate loop analysis. Aggregate matrices of loop pixels and TAD boundaries were generated using *mariner* and visualized using *plotgardener* (Davis et al., 2024; Kramer et al., 2022).

Activity-by-contact

The ABC was used based on Fulco et al., 2019 with slight adjustments (Fulco et al., 2019). To allow for direct comparison of all enhancer–promoter pairs across cell lines, we manually defined potential enhancer regions and used the same input for each cell type. These potential enhancer regions were defined as they typically are in the pipeline, by finding 150,000 ATAC-Seq peaks with the highest H3K27ac levels. The output was filtered using a suggested ABC score cutoff of 0.025 to identify likely enhancer–promoter pairs. To allow for direct comparison with our other enhancer analysis, we filtered the output based on the enhancer regions also overlapping H3K27ac, which still represented a majority of the valid pairs identified.

Matched enhancer–promoter sets

Covariate-matched subset selection among non-looped enhancer–promoter pairs was performed using the *matchRanges* function from the *nullranges* package (Love M Davis et al., 2022; Kolberg et al., 2020). Enhancer–promoter pair distance or mean Micro-C contact frequency were used as covariates. Matching was done with the stratified matching method without replacement.

Gene ontology and gene set enrichment analysis

Gene ontology term enrichment was performed in R using the *gprofiler2* package (Kolberg et al., 2020). Gene set enrichment analysis (GSEA) was performed with the GSEA software, using size factor normalized RNA-Seq counts as input (Subramanian et al., 2005) and the Hallmark H1 gene set.

ATAC-Seq motif analysis

Motif analysis of ATAC-Seq peaks within strengthened and weakened loop anchors was conducted using the HOMER suite findMotifsGenome.pl script with size given (Heinz et al., 2010). ATAC-Seq peaks within the anchors of static loop anchors were used as background.

H3K27ac peak pileup analysis

H3K27ac ChIP-Seq analysis in the anchors of gained, lost, and static loops was conducted using deep-tools (Ramírez et al., 2016). Alignment files were normalized using RPGC with the bamCoverage function, and adjusted using scale factors generated from edgeR TMM normalization factors of counts from overlapping H3K27ac and ATAC-Seq peaks (Robinson et al., 2010). Aggregate profile plots were then created using the plotProfile command.

Genomic data visualization

Micro-C contact frequency maps, aggregate analysis plots, gene annotations, and genomic signal tracks (RNA-Seq, ChIP-Seq, and ATAC-Seq) were visualized and plotted in R using the *plotgardener* package (Kramer et al., 2022).

Acknowledgements

We thank Kathleen Quinn and Joseph Boyd (University of Vermont) for their work building the ChIP-Seq libraries and processing the ChIP-Seq data, respectively. We thank Darawalee Wangsa and Danny Wangsa in the CCR Genetics Branch and OMICS Technology Facility at the NCI for their expert SKY analysis. We thank Alquassem Abuorquob (University of Vermont) for his assistance in ChIP-Seq library preparation. We thank Jordan Zhang, Misha Gattengo, and Sierra Wilson at Dovetail Genomics for coordinating Micro-C library preparation and preliminary analysis. STR analyses and Next Generation Sequencing was in part done with the assistance of the Vermont Integrated Genomics Resource (VIGR) at the Vermont Cancer Center, University of Vermont and the Genomics Sequencing Facility (GSF) at Greehey Children's Cancer Research Institute UT Health San Antonio. The results published here are in part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>. We thank the members of the Misteli and Stein labs for feedback. This work was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), National Cancer Institute, Center for Cancer Research (grant no. ZIABC010309-24 to TM), the Northern New England Clinical Translation Network (grant no. GM115516 to GS), funding from the National Cancer Institute (grant no. P01CA240685 to GS, JS, and SF; grant no. 1F32CA220935-01A1 to AF), and funding from the Charlotte Perelman and Arthur Jason Perelman Fund for Cancer Research (GS and JS). The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Additional information

Funding

Funder	Grant reference number	Author
National Cancer Institute	ZIABC010309-24	Tom Misteli
Northern New England Clinical Translation Network	GM115516	Gary Stein
National Cancer Institute	P01CA240685	Andrew Fritz Janet Stein Gary Stein
National Cancer Institute	1F32CA220935-01A1	Andrew Fritz
Charlotte Perelman and Arthur Jason Perelman Fund for Cancer Research		Janet Stein Gary Stein

The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Author contributions

Kathleen S Metz Reed, Conceptualization, Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Andrew Fritz, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review and editing; Haley Greenyer, Conceptualization, Resources, Data curation, Software, Formal analysis, Investigation, Methodology, Writing – review and editing; Kerstin Heselmeyer-Haddad, Formal analysis, Methodology, Writing – review and editing; Seth Fietze, Resources, Data curation, Funding acquisition; Janet Stein, Gary Stein, Conceptualization, Resources, Supervision, Funding acquisition, Project administration, Writing – review and editing; Tom Misteli, Conceptualization, Resources, Supervision, Funding acquisition, Investigation, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Kathleen S Metz Reed  <https://orcid.org/0000-0003-1466-8258>
Gary Stein  <https://orcid.org/0000-0001-8762-5422>

Tom Misteli  <https://orcid.org/0000-0003-3530-3020>

Peer review material

Reviewer #1 (Public review): <https://doi.org/10.7554/eLife.108135.3.sa1>
Reviewer #2 (Public review): <https://doi.org/10.7554/eLife.108135.3.sa2>
Reviewer #3 (Public review): <https://doi.org/10.7554/eLife.108135.3.sa3>
Author response <https://doi.org/10.7554/eLife.108135.3.sa4>

Additional files

Supplementary files

Supplementary file 1. Micro-C library quality metrics for individual technical replicates, biological replicates, and cell types.

Supplementary file 2. CNVs from NeoLoopFinder (100-kb resolution). Columns indicate bin coordinates (A–C), and CNV values for MCF10A (D), MCF10AT1 (E), and MCF10CA1a (F).

Supplementary file 3. Chromatin loop summary table for total loop set ($n = 29,205$). Columns include loop anchor coordinates (A–J), loop call status by cell type (K–N), SIP AP score (O), loop name (P), loop span length (Q), average and maximum unnormalized counts (R–S), $\log_2$ fold-change (T–V), adjusted p-value (W), differential status by pairwise comparison (X–Z), unnormalized counts by technical replicate (AA–AX), maximum $\log_2$ fold-change (AY), pairwise comparison with greatest fold-change (AZ), differential cluster (BA), Z-score normalized counts (BB–BE), and variance stabilized counts (BF–BH).

Supplementary file 4. TAD boundary summary table for total boundary set ($n = 17,097$). Columns include boundary coordinates (A–C), insulation scores by technical replicate (D–AA), adjusted p-values, difference in insulation score, and differential status for each pairwise comparison (AB–AJ), differential status across all comparisons (AK), differential cluster (AL), average insulation score per cell type (AM–AO), and Z-score normalized insulation scores (AP–AS).

Supplementary file 5. TAD summary table for total set of TADs ($n = 13,231$). Columns include TAD boundary coordinates (A–J), and whether the TAD was called in each cell type (K–M).

Supplementary file 6. Differentially expressed genes from the MCF10 progression series ($n = 8840$). Columns indicate gene coordinates (A–E), gene identifiers (F–I), base mean expression (J), pairwise $\log_2$ fold-change (K–M), adjusted p-values (N–P), and differential status (Q–S), differential cluster (T), unnormalized counts by replicate (U–AC), Z-score normalized counts (AD–AF), and variance stabilized counts (AG–AI).

MDAR checklist

Data availability

The Micro-C, ATAC-Seq, and RNA-Seq datasets generated and analyzed during the current study are available on GEO (SuperSeries GSE320319; SubSeries GSE254045, GSE320215, GSE320216). CTCF ChIP-Seq data was previously published under GEO accession GSE98551. H3K27ac and H3K27me3 ChIP-Seq data were previously published under GEO accession GSE229295. The code used to generate figures from the current study is available on GitHub at the following repository: <https://github.com/ksmetz/MCF10-MicroC> copy archived at [Reed, 2025](#).

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Metz Reed KS, Misteli T, Fritz A, Greenyer H, Fietze S, Stein J, Stein G	2026	Genome reorganization and its functional impact during breast cancer progression	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE320319	NCBI Gene Expression Omnibus, GSE320319
Metz RKS, Misteli T, Fritz A, Greenyer H, Fietze S, Stein J, Stein G	2025	Genome reorganization and its functional impact during breast cancer progression [Micro-C]	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE254045	NCBI Gene Expression Omnibus, GSE254045

Continued on next page

Continued

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Reed KS, Misteli T, Fritz A, Greenyer H, Fietze S, Stein J, Stein G	2026	Genome reorganization and its functional impact during breast cancer progression [ATAC-Seq]	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE320215	NCBI Gene Expression Omnibus, GSE320215
Reed KS, Misteli T, Fritz A, Greenyer H, Fietze S, Stein J, Stein G	2026	Genome reorganization and its functional impact during breast cancer progression [RNA-Seq]	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE320216	NCBI Gene Expression Omnibus, GSE320216

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Fritz A, Boyd J, Tye C, Lian J, Stein G	2017	Chromatin structure and CTCF across the MCF10 breast cancer progression series (ChIP-Seq)	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98551	NCBI Gene Expression Omnibus, GSE98551
Stein G, Boyd JR, Ghule P, Fietze S	2023	Long-range genomic contacts and spatiotemporal chromatin landscape of human histone gene clusters at Histone Locus Bodies during the cell cycle in breast cancer [ChIP-seq]	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE229295	NCBI Gene Expression Omnibus, GSE229295

References

- Abdennur N**, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, Venev SV, Open2C. 2023. Pairtools: from sequencing data to chromosome contacts. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.02.13.528389>, PMID: 36824968
- Amemiya HM**, Kundaje A, Boyle AP. 2019. The ENCODE blacklist: identification of problematic regions of the genome. *Scientific Reports* **9**:9354. DOI: <https://doi.org/10.1038/s41598-019-45839-z>, PMID: 31249361
- Amodeo ME**, Eyler CE, Johnstone SE. 2025. Rewiring cancer: 3D genome determinants of cancer hallmarks. *Current Opinion in Genetics & Development* **91**:102307. DOI: <https://doi.org/10.1016/j.gde.2024.102307>, PMID: 39862605
- Babraham Bioinformatics**. 2017. Trim Galore!. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed February 15, 2017].
- Barutcu AR**, Lajoie BR, McCord RP, Tye CE, Hong D, Messier TL, Browne G, van Wijnen AJ, Lian JB, Stein JL, Dekker J, Imbalzano AN, Stein GS. 2015. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biology* **16**:214. DOI: <https://doi.org/10.1186/s13059-015-0768-0>, PMID: 26415882
- Benabdallah NS**, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, Grimes GR, Therizols P, Bickmore WA. 2019. Decreased enhancer-promoter proximity accompanying enhancer activation. *Molecular Cell* **76**:473–484. DOI: <https://doi.org/10.1016/j.molcel.2019.07.038>, PMID: 31494034
- Bhardwaj A**, Singh H, Rajapakse K, Tachibana K, Ganesan N, Pan Y, Gunaratne PH, Coarfa C, Bedrosian I. 2017. Regulation of miRNA-29c and its downstream pathways in preneoplastic progression of triple-negative breast cancer. *Oncotarget* **8**:19645–19660. DOI: <https://doi.org/10.18632/oncotarget.14902>, PMID: 28160548
- Bioinformatics B**. 2015. FastQC: A quality control tool for high throughput sequence data [Internet]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [Accessed February 15, 2017].
- Borcherding N**, Kusner D, Kolb R, Xie Q, Li W, Yuan F, Velez G, Askeland R, Weigel RJ, Zhang W. 2015. Paracrine WNT5A signaling inhibits expansion of tumor-initiating cells. *Cancer Research* **75**:1972–1982. DOI: <https://doi.org/10.1158/0008-5472.CAN-14-2761>, PMID: 25769722
- Bose A**, Datta S, Mandal R, Ray U, Dhar R. 2024. Increased heterogeneity in expression of genes associated with cancer progression and drug resistance. *Translational Oncology* **41**:101879. DOI: <https://doi.org/10.1016/j.tranon.2024.101879>, PMID: 38262110
- Buchanan NS**, Zhao J, Zhu K, Patwa TH, Miller FR, Lubman DM. 2007. Differential expression of acidic proteins with progression in the MCF10 model of human breast disease. *International Journal of Oncology* **31**:941–949. DOI: <https://doi.org/10.3892/IJO.31.4.941/HTML>, PMID: 17786328
- Cai Y**, Zhang Y, Loh YP, Tng JQ, Lim MC, Cao Z, Raju A, Lieberman Aiden E, Li S, Manikandan L, Tergaonkar V, Tucker-Kellogg G, Fullwood MJ. 2021. H3K27me3-rich genomic regions can function as silencers to repress

- gene expression via chromatin interactions. *Nature Communications* **12**:719. DOI: <https://doi.org/10.1038/s41467-021-20940-y>, PMID: 33514712
- Carter B**, Zhao K. 2021. The epigenetic basis of cellular heterogeneity. *Nature Reviews. Genetics* **22**:235–250. DOI: <https://doi.org/10.1038/s41576-020-00300-0>, PMID: 33244170
- Chakraborty S**, Wenzlitschke N, Anderson MJ, Eraso A, Baudic M, Thompson JJ, Evans AA, Shatford-Adams LM, Chari R, Awasthi P, Dale RK, Lewandoski M, Petros TJ, Rocha PP. 2024. Structural perturbation of chromatin domains with multiple developmental regulators can severely impact gene regulation and development. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.08.03.606480>, PMID: 39372737
- Chakraborty S**, Wenzlitschke N, Anderson MJ, Eraso A, Baudic M, Thompson JJ, Evans AA, Shatford-Adams LM, Chari R, Awasthi P, Dale RK, Lewandoski M, Petros TJ, Rocha PP. 2025. Deletion of a single CTCF motif at the boundary of a chromatin domain with three FGF genes disrupts gene expression and embryonic development. *Developmental Cell* **60**:1838–1853. DOI: <https://doi.org/10.1016/j.devcel.2025.02.002>, PMID: 40015278
- Choppavarapu L**, Fang K, Liu T, Ohihoin AG, Jin VX. 2025. Hi-C profiling in tissues reveals 3D chromatin-regulated breast tumor heterogeneity informing a looping-mediated therapeutic avenue. *Cell Reports* **44**:115450. DOI: <https://doi.org/10.1016/j.celrep.2025.115450>, PMID: 40112000
- Ciriello G**, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, Bowlby R, Shen H, Hayat S, Fieldhouse R, Lester SC, Tse GMK, Factor RE, Collins LC, Allison KH, Chen Y-Y, et al. 2015. Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* **163**:506–519. DOI: <https://doi.org/10.1016/j.cell.2015.09.033>, PMID: 26451490
- Corces R**, Corces VG, Di CL, Shilatifard A. 2016. The three-dimensional cancer genome. *Current Opinion in Genetics & Development* **36**:1–7. DOI: <https://doi.org/10.1016/j.gde.2016.01.002>, PMID: 26855137
- Corces MR**, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, Kathiria A, Cho SW, Mumbach MR, Carter AC, Kasowski M, Orloff LA, Risca VI, Kundaje A, Khavari PA, Montine TJ, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature Methods* **14**:959–962. DOI: <https://doi.org/10.1038/nmeth.4396>, PMID: 28846090
- Cosenza MR**, Rodriguez-Martin B, Korbel JO. 2022. Structural variation in cancer: role, prevalence, and mechanisms. *Annual Review of Genomics and Human Genetics* **23**:123–152. DOI: <https://doi.org/10.1146/annurev-genom-120121-101149>, PMID: 35655332
- Cresswell KG**, Stansfield JC, Dozmorov MG. 2020. SpectralTAD: an R package for defining a hierarchy of topologically associated domains using spectral clustering. *BMC Bioinformatics* **21**:319. DOI: <https://doi.org/10.1186/s12859-020-03652-w>, PMID: 32689928
- Creighton MP**, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *PNAS* **107**:21931–21936. DOI: https://doi.org/10.1073/PNAS.1016071107/SUPPL_FILE/ST01.XLSX, PMID: 21106759
- Cuartero S**, Weiss FD, Dharmalingam G, Guo Y, Ing-Simmons E, Masella S, Robles-Rebollo I, Xiao X, Wang Y-F, Barozzi I, Djeghloul D, Amano MT, Niskanen H, Petretto E, Dowell RD, Tachibana K, Kaikkonen MU, Nasmyth KA, Lenhard B, Natoli G, et al. 2018. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nature Immunology* **19**:932–941. DOI: <https://doi.org/10.1038/s41590-018-0184-1>, PMID: 30127433
- da Costa-Nunes JA**, Noordermeer D. 2023. TADs: Dynamic structures to create stable regulatory functions. *Current Opinion in Structural Biology* **81**:102622. DOI: <https://doi.org/10.1016/j.sbi.2023.102622>, PMID: 37302180
- Dagogo-Jack I**, Shaw AT. 2018. Tumour heterogeneity and resistance to cancer therapies. *Nature Reviews. Clinical Oncology* **15**:81–94. DOI: <https://doi.org/10.1038/nrclinonc.2017.166>, PMID: 29115304
- Danecek P**, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* **10**:1–4. DOI: <https://doi.org/10.1093/gigascience/giab008>, PMID: 33590861
- Davidson IF**, Peters JM. 2021. Genome folding through loop extrusion by SMC complexes. *Nature Reviews. Molecular Cell Biology* **22**:445–464. DOI: <https://doi.org/10.1038/s41580-021-00349-7>, PMID: 33767413
- Davis ES**, Parker SM, Kramer NE, Flores JP, Kiran M, Phanstiel DH. 2024. Mariner: explore the Hi-Cs. *Bioinformatics* **40**:38814811. DOI: <https://doi.org/10.1093/bioinformatics/btae352>, PMID: 38814811
- Dawson PJ**, Wolman SR, Tait L, Heppner GH, Miller FR. 1996. MCF10AT: a model for the evolution of cancer from proliferative breast disease. *The American Journal of Pathology* **148**:313–319. PMID: 8546221.
- Dekker J**, Mirny L. 2016. The 3D genome as moderator of chromosomal communication. *Cell* **164**:1110–1121. DOI: <https://doi.org/10.1016/j.cell.2016.02.007>, PMID: 26967279
- Della Chiara G**, Jiménez C, Viridi M, Crosetto N, Bienko M. 2023. Enhancers dysfunction in the 3D genome of cancer cells. *Frontiers in Cell and Developmental Biology* **11**:1303862. DOI: <https://doi.org/10.3389/fcell.2023.1303862/BIBTEX>
- Despang A**, Schöpflin R, Franke M, Ali S, Jerković I, Paliou C, Chan W-L, Timmermann B, Wittler L, Vingron M, Mundlos S, Ibrahim DM. 2019. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nature Genetics* **51**:1263–1271. DOI: <https://doi.org/10.1038/s41588-019-0466-z>, PMID: 31358994
- de Wit E**, Nora EP. 2023. New insights into genome folding by loop extrusion from inducible degron technologies. *Nature Reviews. Genetics* **24**:73–85. DOI: <https://doi.org/10.1038/s41576-022-00530-4>, PMID: 36180596

- D'Ippolito AM**, McDowell IC, Barrera A, Hong LK, Leichter SM, Bartelt LC, Vockley CM, Majoros WH, Safi A, Song L, Gersbach CA, Crawford GE, Reddy TE. 2018. Pre-established chromatin interactions mediate the genomic response to glucocorticoids. *Cell Systems* **7**:146–160. DOI: <https://doi.org/10.1016/j.cels.2018.06.007>, PMID: 30031775
- Dobin A**, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21. DOI: <https://doi.org/10.1093/bioinformatics/bts635>, PMID: 23104886
- Dorgham MG**, Elliott BA, Holley CL, Mansfield KD. 2023. m6A regulates breast cancer proliferation and migration through stage-dependent changes in Epithelial to Mesenchymal Transition gene expression. *Frontiers in Oncology* **13**:1268977. DOI: <https://doi.org/10.3389/FONC.2023.1268977/FULL>
- Dovetail**. 2025. Welcome to the Dovetail Linked-Read Analysis Page. <https://dovetail-analysis.readthedocs.io/en/latest/index.html> [Accessed March 27, 2025].
- Durand NC**, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. 2016. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems* **3**:95–98. DOI: <https://doi.org/10.1016/j.cels.2016.07.002>, PMID: 27467249
- Feng J**, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. *Nature Protocols* **7**:1728–1740. DOI: <https://doi.org/10.1038/nprot.2012.101>, PMID: 22936215
- Ferlier T**, Coulouarn C. 2022. Regulation of gene expression in cancer-an overview. *Cells* **11**:4058. DOI: <https://doi.org/10.3390/cells11244058>, PMID: 36552821
- Flavahan WA**, Gaskell E, Bernstein BE. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**:eaal2380. DOI: <https://doi.org/10.1126/science.aal2380>, PMID: 28729483
- Fritz AJ**, Ghule PN, Boyd JR, Tye CE, Page NA, Hong D, Shirley DJ, Weinheimer AS, Barutcu AR, Gerrard DL, Fietze S, van Wijnen AJ, Zaidi SK, Imbalzano AN, Lian JB, Stein JL, Stein GS. 2018. Intranuclear and higher-order chromatin organization of the major histone gene cluster in breast cancer. *Journal of Cellular Physiology* **233**:1278–1290. DOI: <https://doi.org/10.1002/jcp.25996>, PMID: 28504305
- Fulco CP**, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR, Anyoha R, Doughty BR, Patwardhan TA, Nguyen TH, Kane M, Perez EM, Durand NC, Lareau CA, Stamenova EK, Aiden EL, Lander ES, Engreitz JM. 2019. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nature Genetics* **51**:1664–1669. DOI: <https://doi.org/10.1038/s41588-019-0538-0>, PMID: 31784727
- Gavish A**, Tyler M, Greenwald AC, Hoefflin R, Simkin D, Tschernichovsky R, Galili Darnell N, Somech E, Barbolin C, Antman T, Kovarsky D, Barrett T, Gonzalez Castro LN, Halder D, Chanoch-Myers R, Laffy J, Mints M, Wider A, Tal R, Spitzer A, et al. 2023. Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. *Nature* **618**:598–606. DOI: <https://doi.org/10.1038/s41586-023-06130-4>, PMID: 37258682
- Georgiades E**, Harrold CL, Roberts N, Kassouf M, Riva SG, Sanders E, Francis HS, Blayney J, Oudelaar AM, Milne TA, Higgs DR, Hughes J. 2023. Active regulatory elements recruit cohesin to establish cell-specific chromatin domains. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.10.13.562171>
- Ghule PN**, Boyd JR, Kabala F, Fritz AJ, Bouffard NA, Gao C, Bright K, Macfarlane J, Seward DJ, Pegoraro G, Misteli T, Lian JB, Fietze S, Stein JL, van Wijnen AJ, Stein GS. 2023. Spatiotemporal higher-order chromatin landscape of human histone gene clusters at histone locus bodies during the cell cycle in breast cancer progression. *Gene* **872**:147441. DOI: <https://doi.org/10.1016/j.gene.2023.147441>, PMID: 37094694
- Glossary**. 2025. ENCODE [Internet]. <https://www.encodeproject.org/glossary/> [Accessed March 8, 2025].
- Goel VY**, Huseyin MK, Hansen AS. 2023. Region Capture Micro-C reveals coalescence of enhancers and promoters into nested microcompartments. *Nature Genetics* **55**:1048–1056. DOI: <https://doi.org/10.1038/s41588-023-01391-1>, PMID: 37157000
- Goldman MJ**, Craft B, Hastie M, Repčeka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J, Haussler D. 2020. Visualizing and interpreting cancer genomics data via the Xena platform. *Nature Biotechnology* **38**:675–678. DOI: <https://doi.org/10.1038/s41587-020-0546-8>, PMID: 32444850
- Guckelberger P**, Doughty BR, Munson G, Rao SSP, Tan Y, Cai XS, Fulco CP, Nasser J, Mualim KS, Bergman DT, Ray J, Jagoda E, Munger CJ, Gschwind AR, Sheth MU, Tan AS, Pulido SG, Mitra N, Weisz D, Shamim MS, et al. 2024. Cohesin-mediated 3D contacts tune enhancer-promoter regulation. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.07.12.603288>, PMID: 39026740
- Guo YA**, Chang MM, Huang W, Ooi WF, Xing M, Tan P, Skanderup AJ. 2018. Mutation hotspots at CTCF binding sites coupled to chromosomal instability in gastrointestinal cancers. *Nature Communications* **9**:1520. DOI: <https://doi.org/10.1038/s41467-018-03828-2>, PMID: 29670109
- Guo M**, Peng Y, Gao A, Du C, Herman JG. 2019. Epigenetic heterogeneity in cancer. *Biomarker Research* **7**:23. DOI: <https://doi.org/10.1186/s40364-019-0174-y>, PMID: 31695915
- Hanahan D**, Weinberg RA. 2011. Hallmarks of cancer: The next generation. *Cell* **144**:646–674. DOI: <https://doi.org/10.1016/J.CELL.2011.02.013/ASSET/2067D218-2368-451A-B4DA-CF9EF4807B23/MAIN.ASSETS/GR1.JPG>, PMID: 21376230
- Hansen KL**, Adachi AS, Braccioli L, Kadvani S, Boileau RM, Pokorny B, Shah R, Anderson EC, Martinovic M, Zhang K, Carel I, Bonitto K, Billelloch R, Fudenberg G, de Wit E, Nora EP. 2024. Synergy between Cis -regulatory elements can render cohesin dispensable for distal enhancer function. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.10.04.615095>
- He Q**, Jing H, Liaw L, Gower L, Vary C, Hua S, Yang X. 2016. Suppression of Spry1 inhibits triple-negative breast cancer malignancy by decreasing EGF/EGFR mediated mesenchymal phenotype. *Scientific Reports* **6**:23216. DOI: <https://doi.org/10.1038/srep23216>, PMID: 26976794

- Heintzman ND**, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* **39**:311–318. DOI: <https://doi.org/10.1038/ng1966>, PMID: 17277777
- Heinz S**, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* **38**:576–589. DOI: <https://doi.org/10.1016/j.molcel.2010.05.004>, PMID: 20513432
- Hsieh THS**, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R, Darzacq X. 2020. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Molecular Cell* **78**:539–553. DOI: <https://doi.org/10.1016/j.molcel.2020.03.002>, PMID: 32213323
- Hsieh THS**, Cattoglio C, Slobodyanyuk E, Hansen AS, Darzacq X, Tjian R. 2022. Enhancer-promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nature Genetics* **54**:1919–1932. DOI: <https://doi.org/10.1038/s41588-022-01223-8>, PMID: 36471071
- Hu Y**, Huang X, Lu X, Lin S. 2022. COL12A1 as a prognostic biomarker in HER2-enriched breast cancer and its association with immune infiltration. *European Journal of Gynaecological Oncology* **43**:85–90. DOI: <https://doi.org/10.22514/ejgo.2022.045>
- Jerkovic I**, Cavalli G. 2021. Understanding 3D genome organization by multidisciplinary methods. *Nature Reviews Molecular Cell Biology* **22**:511–528. DOI: <https://doi.org/10.1038/s41580-021-00362-w>, PMID: 33953379
- Johnstone SE**, Reyes A, Qi Y, Adriaens C, Hegazi E, Pelka K, Chen JH, Zou LS, Drier Y, Hecht V, Shores N, Selig MK, Lareau CA, Iyer S, Nguyen SC, Joyce EF, Hacohen N, Irizarry RA, Zhang B, Aryee MJ, et al. 2020. Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **182**:1474–1489. DOI: <https://doi.org/10.1016/j.cell.2020.07.030>, PMID: 32841603
- Kadota M**, Yang HH, Gomez B, Sato M, Clifford RJ, Meerzaman D, Dunn BK, Wakefield LM, Lee MP. 2010. Delineating genetic alterations for tumor progression in the MCF10A series of breast cancer cell lines. *PLOS ONE* **5**:e9201. DOI: <https://doi.org/10.1371/journal.pone.0009201>, PMID: 20169162
- Kane L**, Williamson I, Flyamer IM, Kumar Y, Hill RE, Lettice LA, Bickmore WA. 2022. Cohesin is required for long-range enhancer action at the Shh locus. *Nature Structural & Molecular Biology* **29**:891–897. DOI: <https://doi.org/10.1038/s41594-022-00821-8>, PMID: 36097291
- Katainen R**, Dave K, Pitkänen E, Palin K, Kivioja T, Välimäki N, Gylfe AE, Ristolainen H, Hänninen UA, Cajuso T, Kondelin J, Tanskanen T, Mecklin J-P, Järvinen H, Renkonen-Sinisalo L, Lepistö A, Kaasinen E, Kilpivaara O, Tuupanen S, Enge M, et al. 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nature Genetics* **47**:818–821. DOI: <https://doi.org/10.1038/ng.3335>, PMID: 26053496
- Kim T**, Han S, Chun Y, Yang H, Min H, Jeon SY, Kim J, Moon H-G, Lee D. 2022. Comparative characterization of 3D chromatin organization in triple-negative breast cancers. *Experimental & Molecular Medicine* **54**:585–600. DOI: <https://doi.org/10.1038/s12276-022-00768-2>, PMID: 35513575
- Koboldt DC**, Fulton RS, McLellan MD, Schmidt H, Kalicki-Verizer J, McMichael JF. 2012. Comprehensive molecular portraits of human breast tumours. *Nature* **490**:61–70. DOI: <https://doi.org/10.1038/nature11412>, PMID: 23000897
- Kolberg L**, Raudvere U, Kuzmin I, Vilo J, Peterson H. 2020. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research* **9**:ELIXIR-709. DOI: <https://doi.org/10.12688/f1000research.24956.2>, PMID: 33564394
- Kramer NE**, Davis ES, Wenger CD, Deoudes EM, Parker SM, Love MI, Phanstiel DH. 2022. Plotgardener: cultivating precise multi-panel figures in R. *Bioinformatics* **38**:2042–2045. DOI: <https://doi.org/10.1093/bioinformatics/btac057>, PMID: 35134826
- Kruse K**, Hug CB, Vaquerizas JM. 2020. FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biology* **21**:303. DOI: <https://doi.org/10.1186/s13059-020-02215-9>, PMID: 33334380
- Kubo N**, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, Hocker JD, Conte M, Gorkin D, Yu M, Li B, Dixon JR, Hu M, Nicodemi M, Zhao H, Ren B. 2021. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nature Structural & Molecular Biology* **28**:152–161. DOI: <https://doi.org/10.1038/s41594-020-00539-5>, PMID: 33398174
- Langmead B**, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**:357–359. DOI: <https://doi.org/10.1038/nmeth.1923>, PMID: 22388286
- Lawrence M**, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLOS Computational Biology* **9**:e1003118. DOI: <https://doi.org/10.1371/journal.pcbi.1003118>, PMID: 23950696
- Li H**, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–1760. DOI: <https://doi.org/10.1093/bioinformatics/btp324>, PMID: 19451168
- Li Q**, Brown JB, Huang H, Bickel PJ. 2011. Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics* **5**:1752–1779. DOI: <https://doi.org/10.1214/11-AOAS466>, PMID: 22905077
- Li J**, Fang K, Choppavarapu L, Yang K, Yang Y, Wang J. 2021. Hi-C profiling of cancer spheroids identifies 3D-growth-specific chromatin interactions in breast cancer endocrine resistance. *Clinical Epigenetics* **13**:1–13. DOI: <https://doi.org/10.1186/S13148-021-01167-6/FIGURES/6>, PMID: 34535185

- Liberzon A**, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Systems* **1**:417–425. DOI: <https://doi.org/10.1016/j.cels.2015.12.004>, PMID: 26771021
- Lin H-H**, Qraitem M, Lian Y, Taylor SR, Farkas ME. 2019. Analyses of BMAL1 and PER2 oscillations in a model of breast cancer progression reveal changes with malignancy. *Integrative Cancer Therapies* **18**:1534735419836494. DOI: <https://doi.org/10.1177/1534735419836494>, PMID: 30943793
- Liu Y**, Nanni L, Sungalee S, Zufferey M, Tavernari D, Mina M, Ceri S, Oricchio E, Ciriello G. 2021. Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. *Nature Communications* **12**:1. DOI: <https://doi.org/10.1038/s41467-021-22666-3>, PMID: 33972523
- Liu M**, Jin S, Agabiti SS, Jensen TB, Yang T, Radda JSD, Ruiz CF, Baldissera G, Rajaei M, Li F-Y, Townsend JP, Muzumdar MD, Wang S. 2025. Tracing the evolution of single-cell 3D genomes in Kras-driven cancers. *Nature Genetics* **57**:3075–3087. DOI: <https://doi.org/10.1038/s41588-025-02297-w>, PMID: 40825871
- Love MI**, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:1–21. DOI: <https://doi.org/10.1186/S13059-014-0550-8/FIGURES/9>, PMID: 25516281
- Love M Davis E**, Phanstiel D, Lee S. 2022. Nullranges: generation of null ranges via bootstrapping or covariate matching [internet]. ffc311b. Github. <https://github.com/nullranges/nullranges>
- Lun ATL**, Perry M, Ing-Simmons E. 2016. Infrastructure for genomic interactions: Bioconductor classes for Hi-C, ChIA-PET and related experiments. *F1000Research* **5**:950. DOI: <https://doi.org/10.12688/f1000research.8759.2>, PMID: 27303634
- Ma X-J**, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, Payette T, Pistone M, Stecker K, Zhang BM, Zhou Y-X, Varnholt H, Smith B, Gadd M, Chatfield E, Kessler J, Baer TM, Erlander MG, Sgroi DC. 2003. Gene expression profiles of human breast cancer progression. *PNAS* **100**:5974–5979. DOI: <https://doi.org/10.1073/pnas.0931261100>, PMID: 12714683
- Maguire SL**, Peck B, Wai PT, Campbell J, Barker H, Gulati A, Daley F, Vyse S, Huang P, Lord CJ, Farnie G, Brennan K, Natrajan R. 2016. Three-dimensional modelling identifies novel genetic dependencies associated with breast cancer progression in the isogenic MCF10 model. *The Journal of Pathology* **240**:315–328. DOI: <https://doi.org/10.1002/path.4778>, PMID: 27512948
- Maor-Nof M**, Shipony Z, Marinov GK, Greenleaf WJ, Gitler AD. 2021. An optimized ATAC-seq protocol for genome-wide mapping of active regulatory elements in primary mouse cortical neurons. *STAR Protocols* **2**:100854. DOI: <https://doi.org/10.1016/j.xpro.2021.100854>, PMID: 34647036
- Marella NV**, Malyavantham KS, Wang J, Matsui S, Liang P, Berezney R. 2009. Cytogenetic and cDNA microarray expression analysis of MCF10 human breast cancer progression cell lines. *Cancer Research* **69**:5946–5953. DOI: <https://doi.org/10.1158/0008-5472.CAN-09-0420>, PMID: 19584277
- Misteli T**. 2020. The self-organizing genome: principles of genome architecture and function. *Cell* **183**:28–45. DOI: <https://doi.org/10.1016/j.cell.2020.09.014>, PMID: 32976797
- Mitelman F**, Johansson B, Mertens F. 2007. The impact of translocations and gene fusions on cancer causation. *Nature Reviews. Cancer* **7**:233–245. DOI: <https://doi.org/10.1038/nrc2091>, PMID: 17361217
- Mortenson KL**, Dawes C, Wilson ER, Patchen NE, Johnson HE, Gertz J, Bailey SD, Liu Y, Varley KE, Zhang X. 2024. 3D genomic analysis reveals novel enhancer-hijacking caused by complex structural alterations that drive oncogene overexpression. *Nature Communications* **15**:6130. DOI: <https://doi.org/10.1038/s41467-024-50387-w>, PMID: 39033128
- Mourad R**, Hsu P-Y, Juan L, Shen C, Koneru P, Lin H, Liu Y, Nephew K, Huang TH, Li L. 2014. Estrogen induces global reorganization of chromatin structure in human breast cancer cells. *PLOS ONE* **9**:e113354. DOI: <https://doi.org/10.1371/journal.pone.0113354>, PMID: 25470140
- Oberbeckmann E**, Quililan K, Cramer P, Oudelaar AM. 2024. In vitro reconstitution of chromatin domains shows a role for nucleosome positioning in 3D genome organization. *Nature Genetics* **56**:483–492. DOI: <https://doi.org/10.1038/s41588-023-01649-8>, PMID: 38291333
- O’Geen H**, Fietze S, Farnham PJ. 2010. Using ChIP-seq technology to identify targets of zinc finger transcription factors. *Methods in Molecular Biology* **649**:437–455. DOI: https://doi.org/10.1007/978-1-60761-753-2_27, PMID: 20680851
- Papanicolaou M**, Parker AL, Yam M, Filipe EC, Wu SZ, Chitty JL, Wyllie K, Tran E, Mok E, Nadalini A, Skhinas JN, Lucas MC, Herrmann D, Nobis M, Pereira BA, Law AMK, Castillo L, Murphy KJ, Zaratzian A, Hastings JF, et al. 2022. Temporal profiling of the breast tumour microenvironment reveals collagen XII as a driver of metastasis. *Nature Communications* **13**:4587. DOI: <https://doi.org/10.1038/s41467-022-32255-7>, PMID: 35933466
- Pasini D**, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, Skotte J, Wutz A, Porse B, Jensen ON, Helin K. 2010. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Research* **38**:4958–4969. DOI: <https://doi.org/10.1093/nar/gkq244>, PMID: 20385584
- Patel H**, Ewels P, Peltzer A, Hammarén R, Botvinnik O, Sturm G. 2020. Nf-core/rnaseq: nf-core/rnaseq v3.0 - silver shark [internet]. 3.0. Zenodo. <https://doi.org/10.5281/zenodo.4323183>
- Patro R**, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**:417–419. DOI: <https://doi.org/10.1038/nmeth.4197>, PMID: 28263959

- Peng A, Peng W, Wang R, Zhao H, Yu X, Sun Y. 2022. Regulation of 3D organization and its role in cancer biology. *Frontiers in Cell and Developmental Biology* **10**:879465. DOI: <https://doi.org/10.3389/fcell.2022.879465>, PMID: 35757006
- Pombo A, Dillon N. 2015. Three-dimensional genome architecture: players and mechanisms. *Nature Reviews. Molecular Cell Biology* **16**:245–257. DOI: <https://doi.org/10.1038/nrm3965>, PMID: 25757416
- Popay TM, Dixon JR. 2022. Coming full circle: On the origin and evolution of the looping model for enhancer-promoter communication. *The Journal of Biological Chemistry* **298**:102117. DOI: <https://doi.org/10.1016/j.jbc.2022.102117>, PMID: 35691341
- Popay TM, Pant A, Munting F, Black ME, Haghani N, Dixon JR. 2024. Genome-wide *in Vivo* dynamics of cohesin-mediated loop extrusion and its role in transcription activation. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.10.02.616323>
- Puleo J, Polyak K. 2021. The MCF10 model of breast tumor progression. *Cancer Research* **81**:4183–4185. DOI: <https://doi.org/10.1158/0008-5472.CAN-21-1939>, PMID: 34400468
- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**:W160–W165. DOI: <https://doi.org/10.1093/nar/gkw257>, PMID: 27079975
- Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, Huang X, Shamim MS, Shin J, Turner D, Ye Z, Omer AD, Robinson JT, Schlick T, Bernstein BE, Casellas R, et al. 2017. Cohesin loss eliminates all loop domains. *Cell* **171**:305–320. DOI: <https://doi.org/10.1016/j.cell.2017.09.026>, PMID: 28985562
- Reed KSM. 2025. MCF10-microc. swl:1:rev:513532d48f436b1cbbdfda894601909166bc2cdd. Software Heritage. <https://archive.softwareheritage.org/swl:1:dir:e662a6b50726f3f993a41cb0bbe858c937879b12;origin=https://github.com/ksmetz/MCF10-MicroC;visit=swl:1:snp:90dafbab49f06c1b52242cd1775c51ed94088905;anchor=swl:1:rev:513532d48f436b1cbbdfda894601909166bc2cdd>
- Reyes-Gopar H, Pérez-Fuentes KA, Bendall ML, Hernández-Lemus E. 2025. Integration of chromosome conformation and gene expression networks reveals regulatory mechanisms in triple negative breast cancer. *Frontiers in Cell and Developmental Biology* **13**:1597245. DOI: <https://doi.org/10.3389/fcell.2025.1597245>, PMID: 40688355
- Rhee DK, Park SH, Jang YK. 2008. Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model. *Genomics* **92**:419–428. DOI: <https://doi.org/10.1016/j.ygeno.2008.08.005>, PMID: 18804527
- Rickman DS, Soong TD, Moss B, Mosquera JM, Dlabal J, Terry S. 2012. Oncogene-mediated alterations in chromatin conformation. *PNAS* **109**:9083–9088. DOI: <https://doi.org/10.1073/PNAS.1112570109/-/DCSUPPLEMENTAL>, PMID: 22615383
- Rinzema NJ, Sofiadis K, Tjalsma SJD, Verstegen M, Oz Y, Valdes-Quezada C, Felder AK, Filipovska T, van der Elst S, de Andrade dos Ramos Z, Han R, Krijger PHL, de Laat W. 2022. Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nature Structural & Molecular Biology* **29**:563–574. DOI: <https://doi.org/10.1038/s41594-022-00787-7>, PMID: 35710842
- Roarty K, Baxley SE, Crowley MR, Frost AR, Serra R. 2009. Loss of TGF-beta or Wnt5a results in an increase in Wnt/beta-catenin activity and redirects mammary tumour phenotype. *Breast Cancer Research* **11**:R19. DOI: <https://doi.org/10.1186/bcr2244>, PMID: 19344510
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–140. DOI: <https://doi.org/10.1093/bioinformatics/btp616>, PMID: 19910308
- Rossini R, Oshaghi M, Nekrasov M, Bellanger A, Domaschens R, Dijkwel Y, Abdelhalim M, Collas P, Tremethick D, Paulsen J. 2024. Loss of multi-level 3D genome organization during breast cancer progression. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.11.26.568711>, PMID: 38076897
- Rowley MJ, Corces VG. 2018. Organizational principles of 3D genome architecture. *Nature Reviews. Genetics* **19**:789–800. DOI: <https://doi.org/10.1038/s41576-018-0060-8>, PMID: 30367165
- Rowley MJ, Poulet A, Nichols MH, Bixler BJ, Sanborn AL, Brouhard EA, Hermetz K, Linsenbaum H, Csankovszki G, Lieberman Aiden E, Corces VG. 2020. Analysis of Hi-C data using SIP effectively identifies loops in organisms from *C. elegans* to mammals. *Genome Research* **30**:447–458. DOI: <https://doi.org/10.1101/gr.257832.119>, PMID: 32127418
- Sanborn AL, Rao SSP, Huang SC, Durand NC, Huntley MH, Jewett AI. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *PNAS* **112**:E6456–E6465. DOI: https://doi.org/10.1073/PNAS.1518552112/SUPPL_FILE/PNAS.1518552112.ST01.XLSX, PMID: 26499245
- Santner SJ, Dawson PJ, Tait L, Soule HD, Eliason J, Mohamed AN, Wolman SR, Heppner GH, Miller FR. 2001. Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. *Breast Cancer Research and Treatment* **65**:101–110. DOI: <https://doi.org/10.1023/a:1006461422273>, PMID: 11261825
- Schoenfelder S, Fraser P. 2019. Long-range enhancer-promoter contacts in gene expression control. *Nature Reviews. Genetics* **20**:437–455. DOI: <https://doi.org/10.1038/s41576-019-0128-0>, PMID: 31086298
- Sen A, Xu Z, Tyndale ST, Yasis J, Cho CY, Bump R, Chandran S, Luo L, Fu Y, Petersen LK, Shokhirev M, Kuo DJ, McVicker G, Dixon JR. 2024. Heterogeneity in chromatin structure drives core regulatory pathways in B-cell acute lymphoblastic leukemia. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.10.04.616668>

- So JY, Lee HJ, Kramata P, Minden A, Suh N. 2012. Differential expression of key signaling proteins in MCF10 cell lines, a human breast cancer progression model. *Molecular and Cellular Pharmacology* **4**:31–40 PMID: 24558516.
- Song Y, Wang L, Wang K, Lu Y, Zhou P. 2023. COL12A1 acts as a novel prognosis biomarker and activates cancer-associated fibroblasts in pancreatic cancer through bioinformatics and experimental validation. *Cancers* **15**:1480. DOI: <https://doi.org/10.3390/cancers15051480>, PMID: 36900272
- Sood V, Misteli T. 2022. The stochastic nature of genome organization and function. *Current Opinion in Genetics & Development* **72**:45–52. DOI: <https://doi.org/10.1016/j.gde.2021.10.004>, PMID: 34808408
- Soule HD, Maloney TM, Wolman SR, Peterson WD Jr, Brenz R, McGrath CM, Russo J, Pauley RJ, Jones RF, Brooks SC. 1990. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Research* **50**:6075–6086 PMID: 1975513.
- Stolper RJ, Tsang FH, Georgiades E, Hansen LLP, Downes DJ, Harrold CL, Hughes JR, Beagrie RA, Davies B, Kassouf MT, Higgs DR. 2023. Loop extrusion by cohesin plays a key role in enhancer-activated gene expression during differentiation. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.09.07.556660>
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* **102**:15545–15550. DOI: <https://doi.org/10.1073/pnas.0506580102>, PMID: 16199517
- Taberlay PC, Achinger-Kawecka J, Lun ATL, Buske FA, Sabir K, Gould CM, Zotenko E, Bert SA, Giles KA, Bauer DC, Smyth GK, Stirzaker C, O'Donoghue SI, Clark SJ. 2016. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Research* **26**:719–731. DOI: <https://doi.org/10.1101/gr.201517.115>, PMID: 27053337
- Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. 2017. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research* **45**:W98–W102. DOI: <https://doi.org/10.1093/nar/gkx247>, PMID: 28407145
- Thiecke MJ, Wutz G, Muhar M, Tang W, Bevan S, Malysheva V, Stocsits R, Neumann T, Zuber J, Fraser P, Schoenfelder S, Peters J-M, Spivakov M. 2020. Cohesin-Dependent and -independent mechanisms mediate chromosomal contacts between promoters and enhancers. *Cell Reports* **32**:107929. DOI: <https://doi.org/10.1016/j.celrep.2020.107929>, PMID: 32698000
- van den Brand T, Collier MD, Flach KD, Gregoricchio S, Mayayo-Peralta I, Dauvey Z, Schuurman K, Teunissen H, Zwart W, de Wit E. 2025. A 3D genome compendium of breast cancer progression. *iScience* **28**:113268. DOI: <https://doi.org/10.1016/j.isci.2025.113268>, PMID: 40894868
- Wang X, Xu J, Zhang B, Hou Y, Song F, Lyu H, Yue F. 2021. Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. *Nature Methods* **18**:661–668. DOI: <https://doi.org/10.1038/s41592-021-01164-w>, PMID: 34092790
- Wang M, Sunkel BD, Ray WC, Stanton BZ. 2022. Chromatin structure in cancer. *BMC Mol Cell Biol* **23**:1–10. DOI: <https://doi.org/10.1186/S12860-022-00433-6>, PMID: 35902807
- Xiao JY, Hafner A, Boettiger AN. 2021. How subtle changes in 3D structure can create large changes in transcription. *eLife* **10**:e64320. DOI: <https://doi.org/10.7554/eLife.64320>, PMID: 34240703
- Yang Y, Choppavarapu L, Fang K, Naeini AS, Nosirov B, Li J, Yang K, He Z, Zhou Y, Schiff R, Li R, Hu Y, Wang J, Jin VX. 2020. The 3D genomic landscape of differential response to EGFR/HER2 inhibition in endocrine-resistant breast cancer cells. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1863**:194631. DOI: <https://doi.org/10.1016/j.bbaggm.2020.194631>, PMID: 32956836
- Yang Y, Zhong X, Lan Y, Liang P, Huang Y, Wang Y, Zhou X, Zhang Z, Liang Y, Xiao X. 2022. Aberrant inactivation of SCNN1G promotes the motility of head and neck squamous cell carcinoma. *Pathology, Research and Practice* **240**:154175. DOI: <https://doi.org/10.1016/j.prp.2022.154175>, PMID: 36327816
- Yang JH, Hansen AS. 2024. Enhancer selectivity in space and time: from enhancer-promoter interactions to promoter activation. *Nature Reviews. Molecular Cell Biology* **25**:574–591. DOI: <https://doi.org/10.1038/s41580-024-00710-6>, PMID: 38413840
- Yost KE, Zhao Y, Hung KL, Zhu K, Xu D, Corces MR, Shams S, Louie BH, Sarmashghi S, Sundaram L, Luebeck J, Clarke S, Doane AS, Granja JM, Choudhry H, Imieliński M, Cherniack AD, Khurana E, Bafna V, Felau I, et al. 2025. Three-dimensional genome landscape of primary human cancers. *Nature Genetics* **57**:1189–1200. DOI: <https://doi.org/10.1038/s41588-025-02188-0>, PMID: 40355593
- Zeng R, Huang J, Zhong M-Z, Li L, Yang G, Liu L, Wu Y, Yao X, Shi J, Wu Z. 2016. Multiple roles of WNT5A in breast cancer. *Medical Science Monitor* **22**:5058–5067. DOI: <https://doi.org/10.12659/msm.902022>, PMID: 28005837
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9**:R137. DOI: <https://doi.org/10.1186/gb-2008-9-9-r137>, PMID: 18798982
- Zhang L, Liu F, Wu L, Fu S, Xing L, Zhang L, Wang X, Wang Y. 2021. Disorder of three-dimensional chromosome structure plays a role in carcinogenesis. *Clinical and Translational Discovery* **1**:e17. DOI: <https://doi.org/10.1002/ctd2.17>
- Zhou Y, Gerrard DL, Wang J, Li T, Yang Y, Fritz AJ, Rajendran M, Fu X, Stein G, Schiff R, Lin S, Fietze S, Jin VX. 2019. Temporal dynamic reorganization of 3D chromatin architecture in hormone-induced breast cancer and endocrine resistance. *Nature Communications* **10**:1. DOI: <https://doi.org/10.1038/s41467-019-09320-9>, PMID: 30944316