


REVIEW ARTICLE **OPEN**


3C methods in cancer research: recent advances and future prospects

 Insoo Yoon^{1,3}, Uijin Kim^{1,3}, Kyung Oh Jung², Yousuk Song^{1,3}, Taesoo Park^{1,3} and Dong-Sung Lee¹  [✉]

© The Author(s) 2024, corrected publication 2024

In recent years, Hi-C technology has revolutionized cancer research by elucidating the mystery of three-dimensional chromatin organization and its role in gene regulation. This paper explored the impact of Hi-C advancements on cancer research by delving into high-resolution techniques, such as chromatin loops, structural variants, haplotype phasing, and extrachromosomal DNA (ecDNA). Distant regulatory elements interact with their target genes through chromatin loops. Structural variants contribute to the development and progression of cancer. Haplotype phasing is crucial for understanding allele-specific genomic rearrangements and somatic clonal evolution in cancer. The role of ecDNA in driving oncogene amplification and drug resistance in cancer cells has also been revealed. These innovations offer a deeper understanding of cancer biology and the potential for personalized therapies. Despite these advancements, challenges, such as the accurate mapping of repetitive sequences and precise identification of structural variants, persist. Integrating Hi-C with multiomics data is key to overcoming these challenges and comprehensively understanding complex cancer genomes. Thus, Hi-C is a powerful tool for guiding precision medicine in cancer research and treatment.

Experimental & Molecular Medicine (2024) 56:788–798; <https://doi.org/10.1038/s12276-024-01236-9>

INTRODUCTION

Three-dimensional (3D) chromatin organization within the nucleus is a dynamic and intricate phenomenon that substantially influences various cellular processes, including gene regulation, DNA replication, and development¹. Understanding the spatial arrangement and interactions of genomic elements is crucial for revealing the complexities of diseases, with cancer being a prominent focus of contemporary research^{2–4}. Over the past decade, the emergence of chromatin conformation capture (3C) techniques and the subsequent development of Hi-C technology have enhanced our ability to dissect the 3D chromatin landscape, offering unparalleled insights into genomic structure and function^{5,6}. In this study, we explore recent advancements in Hi-C technology, shedding light on its applications in deciphering the multifaceted realm of 3D chromatin interactions, with a primary emphasis on cancer research.

The cornerstone of Hi-C methodology lies in its ability to capture and quantify the spatial proximity of genomic regions within the nucleus. This powerful approach enables the comprehensive study of critical chromatin structures, including chromatin loops, topologically associating domains (TADs), and the roles of architectural proteins, such as CCCTC-binding factor (CTCF) and cohesin, in shaping these structures^{7–9}. Moreover, the application of Hi-C can extend beyond primary chromatin structures to explore higher-order chromatin organization, thereby providing a holistic view of how the genome functions as an integrated system.

Advanced tools that can delve into the genetic intricacies of malignant transformation can immensely benefit the study of cancer, which is characterized by genomic instability and aberrant gene regulation³. Recent advancements in Hi-C are revolutionizing cancer research by elucidating the roles of oncogene activation, structural variations (SVs), haplotype phasing, and extrachromosomal DNA (ecDNA) in the occurrence and progression of cancer^{10–13}.

This paper aims to reveal the importance of these cutting-edge Hi-C technologies and their specific contributions to our understanding of cancer biology. From the precise identification of structural variants that activate oncogenes to the elucidation of allelic variations at the haplotype level, Hi-C is a powerful tool for studying the often elusive mechanisms of cancer initiation and progression. Notably, the detection of ecDNA and its effects on gene regulation and amplification in cancer adds another dimension to our understanding of this complex disease.

However, it is essential to acknowledge that Hi-C research is challenging. The existence of repetitive sequences and the need for optimized methods to map SVs are some of the hurdles that researchers must overcome. Nevertheless, the integration of Hi-C technology with multiomics datasets can provide a holistic and multidimensional view of cancer genomes, fostering a deeper understanding that could shape future cancer research and clinical applications. This paper highlights the remarkable potential of Hi-C in revealing the dynamic 3D chromatin architecture of cancer cells.

¹Department of Life Science, University of Seoul, Seoul 02504, Republic of Korea. ²Department of Anatomy, College of Medicine, Chung-Ang University, Seoul 06974, Republic of Korea. ³These authors contributed equally: Insoo Yoon, Uijin Kim, Yousuk Song, Taesoo Park. [✉]email: dslee@uos.ac.kr

Received: 17 December 2023 Revised: 15 March 2024 Accepted: 25 March 2024

Published online: 25 April 2024

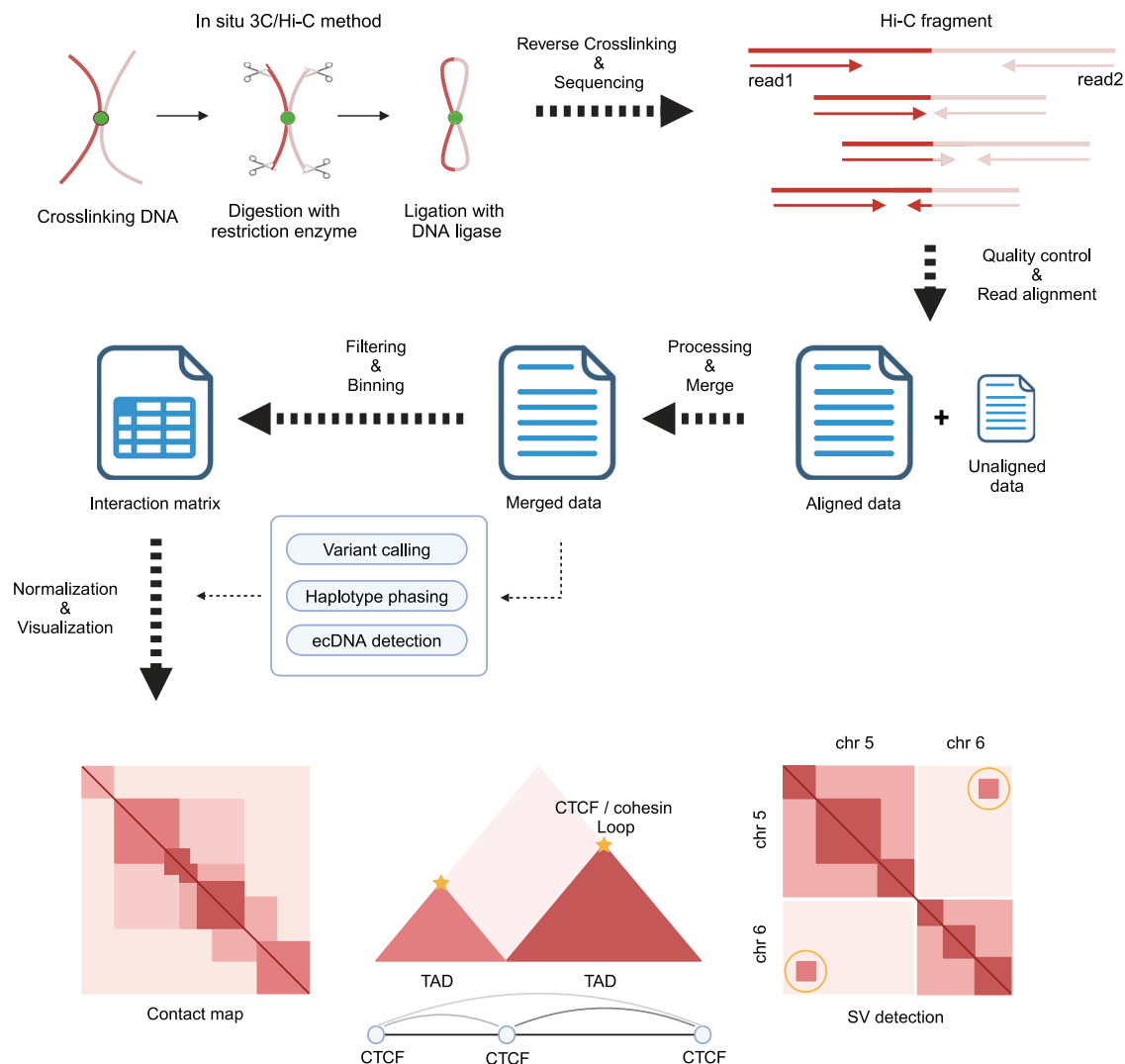


Fig. 1 Overview of the Hi-C sequencing method and analysis pipeline. In situ Hi-C is performed through DNA crosslinking, digestion, and ligation steps. A Hi-C sequencing library is constructed, and paired-end sequencing reads are generated via high-throughput sequencing platforms. The raw data are aligned to the reference genome after quality control. Unaligned reads that can remain because of the biased Hi-C sequencing library are processed and merged with the aligned data. Further processes, such as variant calling, haplotype phasing, and ecDNA detection, can be performed using the merged aligned data with various bioinformatics computational tools. This information can be integrated into an interaction matrix containing contact frequencies between genomic regions to create a contact map, topologically associated domains (TADs), and structural variant (SV) information.

PRINCIPLE AND ANALYSIS OF THE HI-C SEQUENCING METHOD

Within mammalian cells, DNA is organized into nucleosomes and then assembles into more complex chromatin structures, which play important roles in controlling the cell cycle, replication, development, and gene activity. Various interactions take place within these 3D chromatin structures, including those that boost gene expression, such as enhancer–promoter interactions, and those that repress gene expression, such as interactions with lamina-associated domains. The organization of the chromatin structure relies on key structural components, such as CTCF and cohesin^{1,7}. The 3C technique can be used to map interactions between particular genomic regions⁵. Although 3C is effective for assessing changes in the spatial proximity of two regions under different conditions, it has limitations in providing a comprehensive understanding of the entire 3D structure beyond the specific regions being examined. Research has been conducted to solve this problem, and Hi-C, a relatively new technology⁶, has enabled the profiling of the entire genome structure⁶ (Fig. 1). The procedure commences with cell fixation to maintain the spatial

arrangement of the chromatin. This step is followed by DNA cleavage using restriction enzymes. The subsequent ligation step enhances the chances of connecting DNA fragments located close to each other. Subsequently, the joined DNA sequences are separated and used to construct a sequencing library. The prepared Hi-C library is sequenced using a high-throughput sequencing platform. The raw sequencing reads are subjected to quality control to remove low-quality reads and adapter sequences. The reads are aligned to the reference genome using specialized alignment algorithms to determine their genomic locations. Chimeric reads can also be generated from the Hi-C fragments. The unaligned reads undergo further processing and are integrated into the aligned data. The aligned data are analyzed to identify pairs of genomic loci that were in close proximity to the original sample. Contact frequencies between pairs of loci are calculated, representing the frequency of interactions between genomic regions, and the information is stored in an interaction matrix. The following operations are performed only on the aligned data: variant calling, haplotype phasing, and ecDNA

Table 1. Methods for confirming chromatin conformations and the types of data they provide.

Bulk-cell methods	Data type 1	Data type 2	Data access (GEO)	Ref.
Hi-C	Chromatin conformation		GSE18199	6
ChIA-PET	Chromatin conformation		GSE18046	19,20
PLAC-seq	Chromatin conformation		GSE86150	21
HiChIP	Chromatin conformation		GSE80820	22
ChIA-drop	Chromatin conformation		GSE109355	23
Single-cell methods				
scHi-C	Chromatin conformation		GSE48262	24
snHi-C	Chromatin conformation		GSE94489	25
scMethyl-HiC	Chromatin conformation	DNA methylome	GSE119171	28
sn-m3C-seq	Chromatin conformation	DNA methylome	GSE124391 GSE130711	29
HiRES	Chromatin conformation	Transcriptome	GSE223917	30
scCARE-seq	Chromatin conformation	Transcriptome	GSE211395	31
LiMCA	Chromatin conformation	Transcriptome	GSE239969	32

ChIA-Drop chromatin interaction analysis via droplet-based and barcode-linked sequencing, *ChIA-PET* chromatin interaction analysis with paired-end tags, *Hi-C* high-throughput chromosome conformation capture, *HiChIP* Hi-C followed by chromatin immunoprecipitation, *HiRES* Hi-C and RNA-seq employed simultaneously, *LiMCA* linking mRNA to chromatin architecture, *PLAC-seq* proximity ligation-assisted ChIP-seq, *scCARE-seq* simultaneous detection of chromatin architecture and messenger RNA expression by sequencing, *scHi-C* single-cell Hi-C, *scMethyl-HiC* single-cell Methyl-HiC, *snHi-C* single-nucleus Hi-C, *sn-m3c-seq* single-nucleus methyl-3C sequencing.

detection. Contact frequencies are normalized to correct biases introduced during library preparation and sequencing, considering factors such as genomic distance or mappability. Normalized data are visualized as contact maps, where each cell represents the interaction frequency between two genomic loci. Various bioinformatic tools and algorithms are used to analyze contact maps and identify chromatin loops, TADs, compartments, SVs, and other higher-order chromatin structures.

PROGRESS IN THE DEVELOPMENT OF 3D GENOME INTERACTION RESEARCH

In recent years, considerable progress has been made in 3D genome analysis techniques, resulting in an influx of data. However, the majority of currently available Hi-C data offer resolutions between 25 kb and 1 Mb. High-resolution Hi-C data, typically ranging from 1 to 10 kb, are only accessible for a limited number of tissues or cell lines, thus affecting our ability to examine structures at the kb pair level. Notably, as data resolution increases, so does the need for more extensive sequencing, and consequently, the associated costs increase.

Recently, several computational techniques have been developed to improve the resolution of Hi-C data. For example, HiCPlus uses a superresolution convolutional neural network algorithm to replace the original matrix with a very similar interaction matrix using only 1/16 of the original sequencing reads¹⁴. HiCNN can efficiently enhance low-resolution Hi-C data using a 54-layer, very deep convolutional neural network to enhance the resolution of Hi-C data¹⁵. Li et al.¹⁶ developed SRHiC, a deep learning-based computational method, and inferred the corresponding high-resolution Hi-C interaction matrix from a low-resolution matrix with fewer layers. HiDENSEC, a new computational framework for analyzing heterogeneous solid cancers, was developed to predict copy number alterations, large-scale chromosomal rearrangements, and cancer cell fractions¹⁷.

In addition, considerable efforts have been made to improve the specificity of chromatin conformation capture techniques (Table 1). Chromatin immunoprecipitation-loop assay (ChIP-loop), a method that combines ChIP and 3C technology to identify protein-mediated chromatin interactions, has been developed¹⁸.

Fullwood et al.¹⁹ implemented an enrichment strategy using a more extended method, chromatin interaction analysis with paired-end tags (ChIA-PET), and were able to identify chromatin associations mediated by specific proteins of interest at a genome-wide scale without preference²⁰. Proximity ligation-assisted ChIP-seq (PLAC-seq)²¹ and in situ Hi-C followed by chromatin immunoprecipitation (HiChIP)²² improves the DNA contact capture efficiency and accuracy by applying in situ Hi-C, and HiChIP uses a Tn5 transposase-mediated sequencing library construction approach. ChIA-drop can isolate single chromatin complexes without ligation via droplet-based and barcode-linked sequencing²³.

DEVELOPMENT OF SINGLE-CELL TECHNIQUES FOR PROFILING CHROMATIN STRUCTURE

The initial single-cell Hi-C (scHi-C) analysis was performed on mouse CD4⁺ T cells and involved a sequential process of chromatin crosslinking, restriction enzyme digestion, biotin filling, and ligation within the permeated nuclei²⁴. Subsequently, numerous single nuclei were manually isolated after ligation, and biotinylated ligation junctions were individually pulled down from each isolated nucleus. Using single-cell Hi-C technology, Nagano et al.²⁵ reported that mouse embryonic stem cells undergo chromosomal condensation during the early G1 phase, followed by substantial reorganization during replication. This finding implies a correlation between the cell cycle and variations in chromatin interactions. Additionally, the adoption of single-nucleus Hi-C eliminated the conventional biotin-filling step instead of employing a sticky-end ligation method. This modification enabled the distinct reconstruction of the chromatin structure during the transition from oocytes to zygotes in mice²⁶. Another method, Dip-C, also excludes biotin-related steps and has been applied to single human diploid cells (GM12878) to amplify DNA through transposon-based whole-genome amplification to obtain data on haplotypes²⁷.

Recent advancements have led to the development of single-cell multiomic Hi-C techniques that allow simultaneous Hi-C and bisulfite sequencing from the same cell (Table 1). Single-cell methyl-HiC (scMethyl-HiC) was used to perform in situ Hi-C and

Table 2. Recent cancer studies related to chromatin structure.

Year	Studies	3D genome feature	Type of cancer	Refs.
2024	Ijaz, J. et al.	Chromothripsis	Esophageal adenocarcinoma	113
2024	Achinger-Kawecka, J. et al.	Alteration of genome 3D structure associated with decitabine treatment	Breast cancer	124
2023	Kim, K. et al.	Enhancer hijacking	Colorectal cancer	110
2023	Erdmann-Pham, D. D. et al.	Somatic SVs	Melanoma	17
2023	Song, T. et al.	Somatic SVs	Adenocarcinoma	58
2023	Li, P. et al.	Alteration of genome 3D structure associated with oxaliplatin resistance	Colorectal cancer	123
2022	Mallard et al.	Chromothripsis	Leukemia	55
2022	Kim, T. et al.	CTCF	Breast cancer	122
2022	Du, Y. et al.	Somatic SVs	Pancreatic cancer	59
2021	Sunkel et al.	TF(ocoprotein)-driven chromatin loop	Rhabdomyosarcoma	40
2021	Boone et al.	TF(ocoprotein)-driven chromatin loop	Ewing sarcoma	41
2021	Iyyanki et al.	Altered chromatin loops associated with FOXA1 and GATA3	Bladder cancer	49
2021	Montefiori et al.	Enhancer hijacking	Leukemia	54
2021	Sungalee et al.	Enhancer hijacking	Lymphoma	57
2021	Alpsoy et al.	CTCF	Prostate cancer	83
2021	Wong et al.	CTCF	Breast cancer	84
2021	Sivapragasam et al.	CTCF	Melanoma	85
2021	Surdez et al.	STAG2 cohesin complex subunit	Ewing sarcoma	97
2021	Carico et al.	Hinge domain of the cohesin complex	Acute myeloid leukemia	98
2021	Tothova et al.	Cohesin	Acute myeloid leukemia	100
2021	Atkin et al.	Cohesin	Acute myeloid leukemia	101
2020	Zeret	Altered chromatin loops associated with FOXA1 and GATA3	Breast cancer	48
2020	Voronina et al.	Chromothripsis	Various	56
2020	Li et al.	Somatic SVs	Various	11
2020	Akdemir et al.	Somatic SVs	Various	60
2020	Ooi et al.	TAD disruption	Gastric adenocarcinoma	70
2019	Yang et al.	TAD disruption	Leukemia	69
2019	Abdalla et al.	TAD disruption	Breast cancer	71
2019	Ma et al.	Cohesin-associated proteins (PDS5B and miRNA-223)	Pancreatic cancer	99
2018	Baxter et al.	Altered chromatin loops associated with FOXA1 and GATA3	Breast cancer	47

bisulfite conversion²⁸. Single-nucleus methyl-3C sequencing (sn-m3C-seq) was performed on a Hi-C instrument without the ligation junction enrichment step²⁹. These techniques enable the identification of subcell populations based on known DNA methylation profiles as cell type-specific epigenetic signatures. In a recent study, a multiomics profiling technique that can efficiently and accurately determine both 3D genome structure and gene expression was developed (Table 1). Liu et al.³⁰ used Hi-C and RNA-seq simultaneously (HiRES) to generate a 3D genome and transcriptome atlas of mouse embryos after implantation and identified the interaction of multiple factors that jointly shape the single-cell 3D genome during development. In addition, single-cell multimodal omics methods allowing the simultaneous detection of chromatin architecture and messenger RNA (mRNA) expression by sequencing (scCARE-seq)³¹ and linking mRNA to chromatin architecture (LiMCA)³² have been developed to coprofile chromatin architecture and transcriptomes. These studies identified changes in the processes that control the regulation of olfactory receptor expression and confirmed that periodic changes in chromatin structure and transcription are interconnected in the cell cycle. These single-cell multiomics Hi-C methods facilitate the concurrent characterization of cell type-specific chromatin organization and epigenomes in complex tissues or heterogeneous cancer genomes.

Considering the historical reliance on tumor heterogeneity signals to infer cancer traits, these innovative methodologies present invaluable resources for mapping the development of tumors. Gaining insights into the source of mutations in the course of tumorigenesis can reveal the mutations that promote tumor initiation, proliferation, infiltration, and metastasis. This knowledge is pivotal in advancing our understanding of cancer development and progression.

RECENT CANCER STUDIES FROM THE PERSPECTIVE OF CHROMATIN STRUCTURAL LEVELS

Recent advancements in cancer research have revealed an intricate relationship between 3D chromatin structure and transcriptional regulation in cancer. Here, we summarize recent cancer studies related to chromatin structure by referring to previous reviews by Deng et al.³ and Wang et al.⁴ (Table 2 and Fig. 2).

Chromatin loops

Chromatin loops, identified in Hi-C data as concentrated regions with increased interactions^{33,34}, play a crucial role in gene regulation. Whereas traditional cis-chromatin sequencing methods, such as Hi-C, can detect long-range chromatin loops, newer

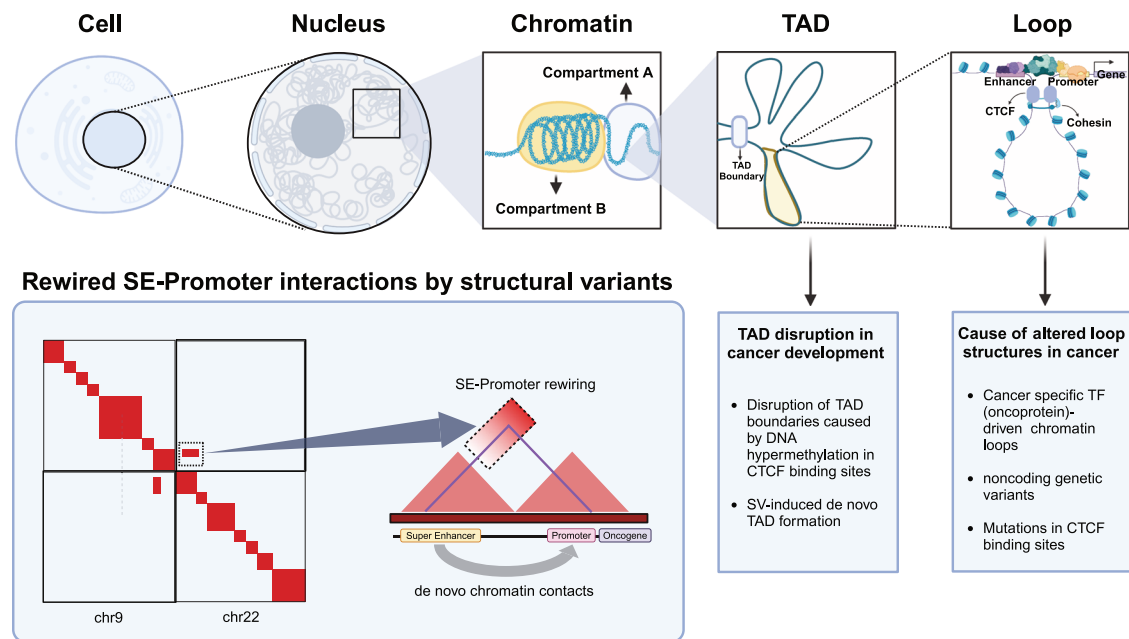


Fig. 2 Deciphering the 3D genomic landscape by investigating chromatin loops and SV-induced tumorigenesis. Chromatin is a complex of DNA and proteins that is organized in the nucleus. It is divided into compartments A and B, which represent different structural and functional states, respectively. TADs are self-interacting genomic regions that play crucial roles in organizing chromatin into discrete functional units. Enhancers are regulatory DNA sequences that enhance the transcription of specific genes. These enhancers often interact with gene promoters over long distances, forming DNA loops that bring distant regulatory elements in close proximity to their target genes. Chromosome conformation capture (3C) maps are used to study these 3D interactions and provide insights into the spatial organization of chromosomes and the formation of TADs and DNA loops. SVs are alterations in genomic DNA sequences, including insertions, deletions, duplications, inversions, and translocations. These variants can be detected through 3C maps, revealing instances of enhancer hijacking, where SVs alter these interactions and dysregulate gene expression, thereby contributing to the development of diseases, such as cancer.

techniques, such as HiChIP^{22,35}, TrAC looping³⁶, and Capture-HiC³⁷, focus on the high-resolution sequencing of shorter-range loops. These methods provide PET images that aid in defining chromatin interactions, including enhancer–promoter interactions.

Incorporating the micrococcal nuclease enzyme into sequencing increases the resolution, which has led to recent methods, such as Micro-C³⁸ and Microcapture-C³⁹. These approaches offer high-resolution chromatin interaction data, enabling the examination of short-range chromatin interactions in human cancers.

These high-resolution interactions also provide insights into transcription factor (TF) binding within chromatin loops, offering new ways to study TF-driven childhood malignancies, such as rhabdomyosarcoma⁴⁰ and Ewing sarcoma⁴¹. In rhabdomyosarcoma, fusion-positive subtypes show unique chromatin looping driven by the PAX3–FOXO1 oncoprotein^{40,42}, offering a potential target for clinical interventions^{35,43}. Fusion-negative subtypes exhibit different chromatin looping patterns, which can be influenced by clinical agents, such as trametinib^{42,44,45}.

TF motifs play a role in shaping regulatory chromatin loops in human cancers. Recent studies have highlighted the role of noncoding genetic variants in altering chromatin interactions⁴⁶ and emphasized the importance of pioneering factors, such as FOXA1 and GATA3, which are associated with altered chromatin loops in breast^{47,48} and bladder cancer⁴⁹. Mutations in the CTCF-binding sites in various cancers also affect chromatin loop structures⁵⁰.

In summary, the study of chromatin loops using advanced techniques offers insights into gene regulation and high-resolution interactions and their relevance to various diseases, including pediatric and adult cancers. These findings revealed potential targets for clinical intervention, highlighted the impact of genetic variants on chromatin interactions, and emphasized the role of pioneering factors and CTCF mutations in shaping chromatin loops in cancer.

Structural variations

Recent studies have highlighted the power of Hi-C, particularly when combined with whole-genome sequencing, to identify SVs in the human genome, particularly in the context of cancer^{10,51}. Hi-C, a chromatin sequencing and imaging method, has proven to be efficient for the de novo detection of SVs in cancer genomes⁵². These studies have several implications.

One key aspect is the newfound understanding of how SVs, such as deletions, inversions, and translocations, systematically impact enhancers and insulators that play crucial roles in gene regulation. Previous studies provide valuable insights into the relationship between gene regulation and SVs, shedding light on how SVs alter the regulatory functions of these elements.

The topological context of copy number variations (CNVs) and gene fusion events in cancer can also be revealed using Hi-C^{49,53}. This revealed how these genetic alterations fit within chromatin domains. Furthermore, elucidating the effects of SVs on both noncoding and coding regions of the genome continue to provide insights into the epigenetic mechanisms that drive tumor development.

A notable study of leukemia genomes illustrated the potential of Hi-C for understanding SVs. In this case, SVs were found to modify the proximity of the *BCL11B* gene locus and its enhancer, thereby affecting gene expression in progenitor cells⁵⁴. This study employed HiChIP data mapped onto patient-specific reference genomes to account for SVs, highlighting the importance of accounting for these variations in genome analysis.

The potential of Hi-C to reveal SV patterns, such as chromothripsis (multiple catastrophic chromosomal rearrangements), makes it a powerful tool, particularly in cases where other methods, such as SNP arrays or RNA-seq, may not capture these patterns effectively^{55,56}.

Hi-C has been used to elucidate how SV-induced enhancer transpositions regulate the expression of oncogenic

drivers^{49,54,57–59} in various cancer types, including leukemia, lymphoma, bladder cancer, adenocarcinoma, and pancreatic cancer^{49,54,57–59}. These findings emphasize the potential of using low-depth Hi-C or Hi-ChIP to investigate patterns of SV and altered enhancer function in both clinical and basic research contexts. Ultimately, the integration of Hi-C with other sequencing methods offers a comprehensive view of the SVs in cancer genomes, providing valuable insights into the mechanisms driving cancer development and progression.

Complex SVs in human cancers often differ from those in normal tissues, and their association with chromatin interactions is becoming increasingly evident. There are common “SV pathways” leading to recurrent fusion-oncogene events, along with unique SV patterns associated with specific cancer types^{11,60}. Specific types of cancer exhibit recurrent alterations in chromatin domain boundaries, connecting SVs with gene dysregulation. The integration of SVs into the context of transcriptional regulatory elements and domains is expected to define the structural drivers of cancer.

Overall, the combination of Hi-C and whole-genome sequencing is a powerful tool for understanding the role of SVs in cancer development and revealing their complex effects on the regulatory and structural architecture of the genome.

TADs

A TAD is a region in the genome characterized by increased self-interactions^{61,62}. These are crucial for maintaining proper genomic organization and function. TAD boundaries are primarily regulated by CTCF.

TADs are established during embryonic development^{63,64} and play a role in various cellular processes, including cell cycle regulation and DNA replication^{25,26,65–67}. The dysregulation of TADs has been linked to diseases, particularly developmental malformations and tumorigenesis. De novo TAD formation, often referred to as neo-TAD, is a phenomenon in which SVs disrupt existing TADs in the genome⁵². This disruption creates new TAD boundaries, leading to changes in gene regulation. In the context of cancer research, neo-TAD formation is particularly relevant because it can result in the activation of oncogenes or the inactivation of tumor suppressor genes, contributing to tumorigenesis. Researchers have identified neo-TADs using Hi-C, and understanding their functional consequences is crucial for deciphering disease mechanisms⁶⁸. Various studies have shown that TAD disruption can contribute to the development of cancers, such as leukemia⁶⁹, gastric adenocarcinoma⁷⁰, breast cancer⁷¹, and multiple myeloma⁷².

The strength of CTCF-mediated transcriptional insulation depends on factors such as the number of CTCF protein-binding sites, surrounding DNA sequences, and the location of these sites within or outside the TAD boundaries⁷. Targeting CTCF and TAD regulation may have broad applications in the treatment of cancer and other inherited diseases.

Recent studies have revealed subdomains within TADs known as chromatin nanodomains (CNDs)⁷³, which persist in cells even when CTCF or cohesin, another important protein, is depleted, indicating their structural importance⁷⁴. Understanding the role of CNDs, particularly in the context of enhancer–promoter interactions, can provide new insights into cancer development and progression.

Exploring the characteristics of self-renewing cancer stem cells, as well as the dynamic chromatin changes that occur during cancer progression using high-resolution techniques will be vital for advancing our understanding of cancer biology and treatment.

CTCF

CTCF, a highly conserved protein, plays a central role in regulating 3D chromatin architecture and gene expression^{61,75}. It is a versatile factor involved in processes such as transcriptional

regulation, insulation, genomic imprinting, DNA repair and alternative splicing^{76–78}. Its binding sites can be influenced by epigenetic modifications, mutations, and environmental factors, highlighting its dynamic nature⁷⁶.

CTCF-mediated looping is highly conserved and critical during early embryonic development⁷⁹. It regulates cell differentiation, fate, and development and is indispensable for these processes⁸⁰. CTCF binding also responds to environmental stimuli⁸¹, further highlighting its dynamic role.

Recently, extensive research has focused on the association between CTCF and cancer. Studies have shown that CTCF aberrations can lead to tumorigenesis. For instance, in gliomas, mutations can lead to DNA hypermethylation, disruption of TAD boundaries, and oncogene activation⁸². In prostate cancer, CTCF interacts with other proteins to regulate gene expression⁸³. Breast cancer metastasis can be controlled by CTCF and early growth response 1 (EGR1)⁸⁴. Additionally, CTCF binding has been linked to UV damage and mutation hotspots in melanoma⁸⁵.

The role of CTCF in human developmental diseases has also been discussed, with particular emphasis on its impact on cis-regulatory interactions and 3D chromatin architecture. Furthermore, the involvement of CTCF in cell aging and senescence has been highlighted, suggesting its potential role in restoring the function of senescent cells and its relevance to tumorigenesis^{86,87}.

In summary, CTCF is a multifaceted protein with a wide range of functions, and its dysregulation can contribute to various diseases, particularly cancer and developmental disorders. Understanding the dynamic nature of CTCF and its role in these processes is essential for advancing our knowledge of disease mechanisms and potential therapeutic interventions.

Cohesin

Cohesin is a well-known structural maintenance protein in the chromosome complex that plays a pivotal role in sister chromatid cohesion during the cell cycle, ensuring proper chromosome segregation⁸⁸. It also contributes to maintaining the 3D architecture of interphase chromosomes^{8,33,89}. After the establishment of CTCF anchors, cohesin assists in the loop extrusion process, which can occur through various mechanisms, including diffusion⁹⁰, ATP hydrolysis⁹¹, and translocating factors, such as RNA polymerase II⁹².

This protein complex plays a regulatory role in gene expression by modulating chromatin structure. It collaborates with CTCF and other proteins to regulate gene expression and genome organization. Cohesin and CTCF work together to separate chromatin into distinct spatial domains, facilitating long-range interactions and coregulating genes within those domains⁹.

In the context of cancer, dysregulation of cohesin-related proteins is associated with tumor initiation and development^{93,94}. Mutations in genes encoding cohesin proteins have been identified in various cancers, particularly myeloid malignancies⁹⁵ and solid tumors^{96,97}. The hinge domain of the cohesin complex plays a critical role in these mutations⁹⁸. Additionally, cohesin-associated proteins, such as PDS5B and miRNA-223, have been linked to cancer, suggesting that these proteins may be potential therapeutic targets⁹⁹.

Targeting dysfunctional or mutated cohesin complexes may be new strategies for cancer therapy, and patients with cohesin mutations can be considered distinct subgroups in clinical trials. Abnormal expression of cohesin complex genes can also serve as a prognostic tool in cancer.

In hematological malignancies, cohesin mutations are key drivers of DNA damage repair and alterations in chromatin architecture, thereby providing therapeutic opportunities¹⁰⁰. DNA breakage at the CTCF/cohesin binding sites may become a robust early detection tool for individuals susceptible to blood cancer¹⁰¹.

Interactions of the cohesin complex with other architectural proteins such as CTCF and polycomb group proteins emphasize its

fundamental role in 3D chromatin structure and gene regulation¹⁰². Understanding the differences between cohesin subunits such as STAG1 and STAG2 could lead to more targeted therapeutic approaches.

In conclusion, the cohesin protein complex is a versatile player in various biological processes and has the potential to open new therapeutic avenues for cancer and other diseases, especially as advanced technologies provide deeper insights into its functions.

DETECTION OF STRUCTURAL VARIANTS THAT DRIVE ONCOGENE ACTIVATION

Hi-C has emerged as a powerful tool for identifying cancer-specific alterations within TADs, shedding light on their pivotal role in cancer development^{60,75}. These alterations originate primarily from SVs and include duplications, deletions, inversions, and translocations. These genomic rearrangements have the potential to disrupt the 3D organization of the genome⁵². Such disruption, often referred to as enhancer hijacking, is known to activate oncogenes, promoting uncontrolled cell proliferation and tumor growth, and has been documented across a spectrum of cancers, including acute myeloid leukemia^{2,103}, neuroblastoma¹⁰⁴, medulloblastoma¹⁰⁵ and colorectal cancer¹⁰⁶.

Superenhancers (SEs) can also be redirected by enhancer hijacking to promote the expression of oncogenes or other genes that drive tumorigenesis. This aberrant activation of SEs substantially contributes to cancer development and progression^{107,108} (Fig. 3a). Understanding how SEs are hijacked in cancer cells has become a crucial focus for both basic research and potential therapeutic strategies in the field of oncology.

Furthermore, Xu et al.¹⁰⁹ harnessed Hi-C data derived from cancer cell lines and patient samples. Their work precisely identified specific loci where recurrent changes manifested within the 3D structure of the genome. Notably, these structural alterations are strongly associated with well-known oncogenes, such as *TERT*, *MYC*, and *CCND1*. These findings were substantiated by CRISPR-Cas9 genome engineering. This cutting-edge technology enables researchers to anticipate the activation of oncogenes by scrutinizing the interactions between distinct chromatin regions and SE activity.

Kim et al.¹¹⁰ revealed a pivotal insight: recurrent enhancer hijacking, orchestrated by SVs, can instigate a surge in oncogenic expression, all without a pronounced surge in CNVs. This discovery further underscores the indispensable role of Hi-C data in revealing the intricate web of oncogenic regulation in cancer, thereby shedding light on the multifaceted mechanisms driving cancer development and progression. This study adds a vital piece to the puzzle of cancer research by showing the complexity of cancer genomes and how seemingly subtle changes in their 3D structures can have profound consequences in the context of oncogenic expression.

RECONSTRUCTION OF COMPLEX GENOMIC REARRANGEMENTS AT THE HAPLOTYPE LEVEL

Hi-C data play a pivotal role in haplotype phasing, a fundamental process that is crucial for identifying the parental origin of genetic variants located along a chromosome¹². Haplotype phasing can be performed either statistically or experimentally. The statistical method relies on population genetic data and linkage disequilibrium but may introduce errors at recombination hotspots¹¹⁰. In contrast, experimental phasing uses multiple variants within sequencing reads but faces challenges owing to the limited prevalence of heterozygous polymorphisms¹¹¹.

Using the Hi-C data, primary contigs are assembled and extended into chromosome-level allele-specific contigs. Primary contigs are partitioned into allele blocks, and linkages are established, ultimately leading to the creation of extended

allele-specific contigs (Fig. 3b). Several computational algorithms, such as HapCUT¹¹², leverage Hi-C data for this purpose.

Recent efforts have integrated Hi-C data with other sequencing techniques, especially long-read sequencing, leading to substantial improvements in the fractionation of phased variants^{12,113}. The synergy between Hi-C and long-read sequencing, linked-read sequencing, and genome assembly methods can achieve more comprehensive and precise haplotype phasing. This method plays a crucial role in understanding complex genomic rearrangements at the haplotype level and offers valuable insights into allele-specific chromatin interactions, particularly in cancer research. It not only aids in the identification of allele-specific structural variants and somatic mutations but also provides a nuanced perspective on somatic clonal evolution¹¹⁴. Analysis of allele-specific somatic mutations and variants enables deeper exploration of the dynamics of genomic changes in the context of disease progression, considerably enhancing our understanding of these processes, particularly in the study of diseases, such as cancer.

Notably, Garg et al.¹¹⁵ developed an efficient approach for chromosome-level haplotype reconstruction using Hi-C and long-read sequencing in cancer cell lines, surpassing the existing methods in terms of accuracy and speed. This approach revealed a detailed structural variant landscape, which not only aids in personalized diagnosis but can also predict therapeutic responses in patients with cancer.

In summary, integrating Hi-C data with advanced sequencing techniques is crucial for haplotype phasing and offers insights into complex genomic rearrangements, allele-specific interactions, and somatic evolution.

ECDNA DETECTION

EcDNA refers to circular DNA segments that exist outside linear chromosomes and vary in size from kb to Mb. EcDNA plays a crucial role in generating tumor heterogeneity, contributing to drug resistance and poor prognosis¹³. It is considered to be a major factor in the genomic diversity of cancers, particularly glioblastoma^{116,117}.

The identification of ecDNA has been substantially facilitated by the emergence of Hi-C technology. This is primarily due to unique characteristics that distinguish ecDNA from conventional chromosomal DNA⁵³. These characteristics encompass a range of distinctive features, including specific spatial proximities, the formation of chimeric DNA molecules, unconventional DNA interactions, SVs, and enhanced signals detectable within the Hi-C data. This technology effectively leverages these traits to accurately locate and differentiate ecDNA within the genomic landscape.

The principal objective of Hi-C is to reveal the intricate spatial organization of various genomic regions within the cell nucleus⁷⁵. Notably, ecDNA exhibits distinct spatial relationships compared with chromosomal DNA^{13,118}, making Hi-C exceptionally adept at capturing these spatial proximities. During Hi-C, DNA fragments originate from diverse genomic amalgamation areas, resulting in the formation of chimeric DNA molecules⁷⁵. The amalgamation of ecDNA into chimeric structures results in recognizable patterns that serve as unequivocal indicators of the presence of ecDNA.

Specifically, regions containing ecDNA may exhibit higher concentrations of Hi-C interactions^{13,118} and coamplify enhancer elements^{119,120}, signifying areas of elevated spatial interactions (Fig. 3c). These intensified signals in the Hi-C data further enhanced the detectability of ecDNA within the genome. By harnessing these unique attributes, Hi-C offers a robust and comprehensive approach for detecting and characterizing ecDNA, making a substantial contribution to our understanding of the genomic landscape. This is particularly vital in the context of diseases, such as cancer, where ecDNA functions as a mobile enhancer that amplifies chromosomal transcription.

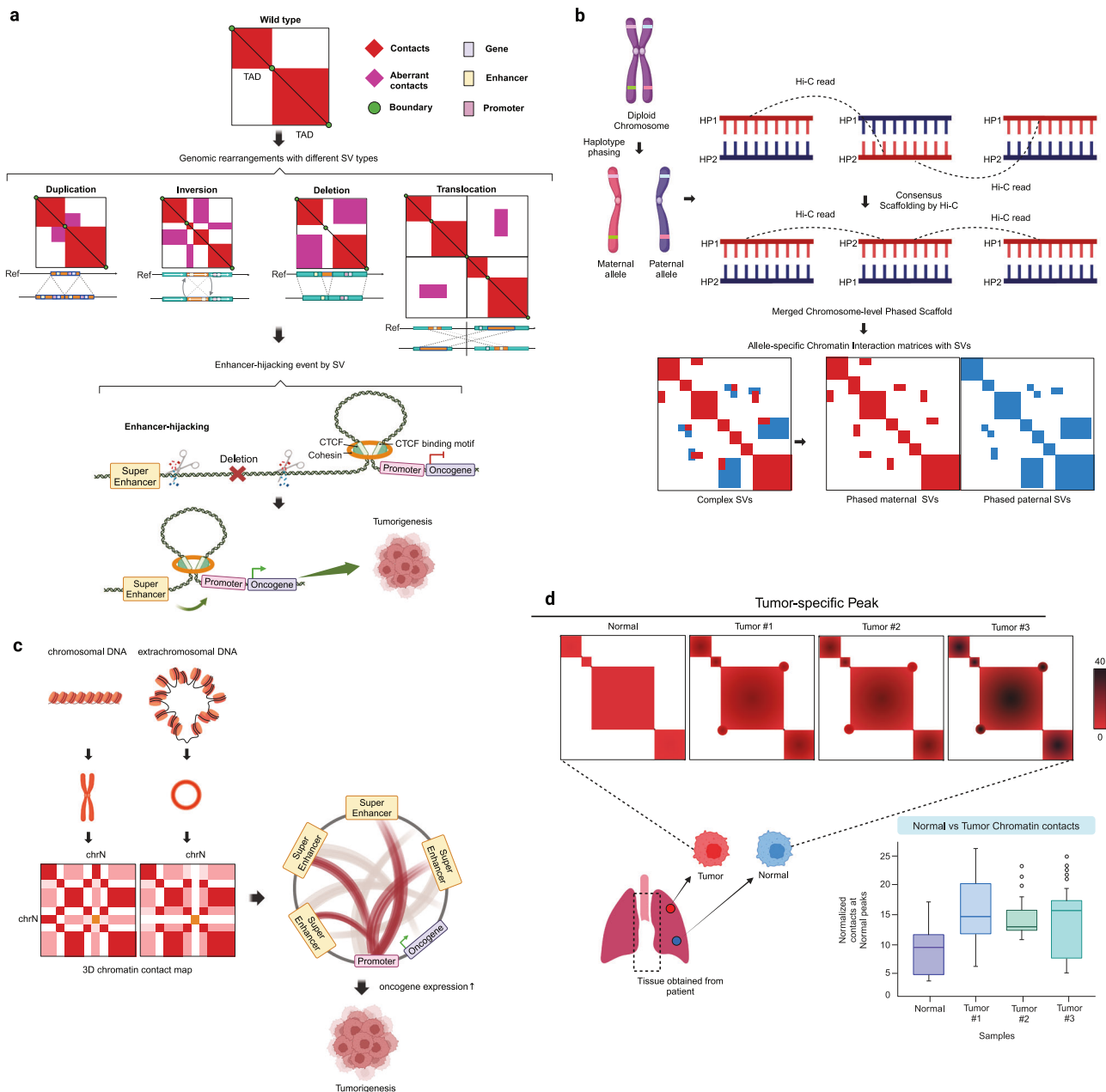


Fig. 3 Applications of Hi-C in human cancer genome research. **a** Hi-C contact matrices can be used to effectively identify enhancer hijacking events resulting from SVs. These matrices can reveal alterations in enhancer–promoter interactions resulting from genomic rearrangements. Enhancer hijacking occurs when regulatory elements such as super-enhancers are misdirected and interact with the promoters of genes, including oncogenes, with which they normally do not interact. This aberrant interaction can lead to dysregulated gene expression, contributing to cancer oncogenesis and development. **b** Chromosome-level haplotype blocks are constructed using Hi-C reads to link alleles in contigs, thus revealing specific combinations of genetic variants present on each chromosome. The precise characterization of individual genomes, facilitated by chromosome-level haplotype blocks, allows for a better understanding of genetic diversity and inheritance patterns within populations. It provides information on the specific combinations of genetic variants present on each chromosome, aiding in personalized medicine and genetic studies. **c** In the context of ecDNA detection, Hi-C contact matrices leverage the characteristic pattern of ecDNA interactions, which typically involve local self-interactions and lack interchromosomal interactions. This distinctive behavior allows Hi-C to identify and characterize ecDNAs within the genomic landscape, particularly in the context of diseases, such as cancer. **d** Comparative analysis of cancer cell lines and normal mammary cells revealed CTCF-dependent changes in chromatin organization, affecting chromatin accessibility, transcription, and epigenetic features. These disruptions provide critical insights into the structural and functional aspects of 3D chromatin in cancer cells.

CHARACTERIZATION OF 3D CHROMATIN ORGANIZATION IN CANCER THROUGH COMPARATIVE ANALYSIS

Chromatin exhibits a 3D architecture that plays essential roles in fundamental biological processes, such as the cell cycle, DNA replication, cellular differentiation, and transcription regulation¹²¹. Conversely, disruptions in these 3D structures are closely

associated with the emergence of anomalies and various diseases, among which cancer is particularly notable³. The association between disruptions in 3D chromatin organization and cancer is an area of active research with substantial implications for our understanding of cancer development and progression.

Kim et al.¹²² comprehensively characterized 3D chromatin organization in various breast cancer cell lines and compared it to that in normal mammary epithelial cells and tissues (Fig. 3d). This study revealed that CTCF-dependent cancer-specific changes in 3D chromatin organization are associated with perturbed chromatin accessibility and transcriptional dysregulation, affecting compartment domains, TADs, and chromatin loops. These changes, which are strongly linked to epigenetic modifications and gene expression, are partially conserved in cancer tissues, highlighting the disruptive nature of 3D chromatin organization in breast cancer.

In summary, disruptions in 3D chromatin organization in cancer revealed a connection between these structural changes, epigenetic features, and gene expression, offering critical insights into the 3D chromatin structural features of cancer. In addition to previous findings, studies have shown that drug resistance or drug efficacy causes differences in 3D chromatin organization^{123,124}. Understanding how drugs affect chromatin organization can provide insight into the mechanisms underlying drug resistance and efficacy, potentially leading to the development of more effective treatments. This knowledge serves as a foundation for diagnostic and therapeutic investigations and advances our understanding of cancer etiology.

CONCLUDING REMARKS

Here, we review how 3C technology has revolutionized our understanding of cancer biology by revealing the complexities of 3D chromatin interactions, including chromatin loops, TADs, and interactions between regulatory elements and oncogenes. This knowledge offers promising avenues for identifying critical factors involved in tumorigenesis and potential therapeutic interventions.

3C-based techniques, such as HiChIP and Micro-C, can provide higher-resolution data, which are particularly valuable for studying enhancer–promoter interactions and understanding the impact of oncogenic drivers on gene expression in cancer. 3C technology has also emerged as an effective tool for detecting SVs within cancer genomes, shedding light on how these alterations disrupt enhancers and insulators and affect both coding and noncoding regions.

The process of haplotype phasing, which is vital for studying oncogene activation, has been streamlined using Hi-C technology, particularly when integrated with long-read sequencing. This integration enables precise identification of allele-specific SVs and somatic mutations. The unique ability of Hi-C-based methods to detect ecDNA is essential for understanding its role in cancer development, as it often acts as a mobile enhancer and amplifies gene transcription. Comparative analyses of 3D chromatin organization in cancer have revealed disruptions associated with epigenetic and gene expression alterations, providing insights into the importance of 3D chromatin structures in cancer development.

Ongoing advancements, including the integration of multiple omics data and the development of single-cell multiomics 3C techniques, offer opportunities to characterize cell type-specific chromatin organization in complex tissues and heterogeneous cancer genomes. Despite this progress, challenges remain, such as handling repetitive sequences and refining methods for mapping structural variants.

In summary, 3C technology has substantially advanced our understanding of cancer biology, offering exciting possibilities for more targeted therapies and shaping our knowledge of cancer biology and its clinical applications.

REFERENCES

- Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
- Krijger, P. H. L. & De Laat, W. Regulation of disease-associated gene expression in the 3D genome. *Nat. Rev. Mol. Cell Biol.* **17**, 771–782 (2016).
- Deng, S., Feng, Y. & Pauklin, S. 3D chromatin architecture and transcription regulation in cancer. *J. Hematol. Oncol. Pharm.* **15**, 1–23 (2022).
- Wang, M., Sunkel, B. D., Ray, W. C. & Stanton, B. Z. Chromatin structure in cancer. *BMC Mol. Cell Biol.* **23**, 1–10 (2022).
- Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
- Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Huang, H. et al. CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. *Nat. Genet.* **53**, 1064–1074 (2021).
- Wutz, G. et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).
- Merkenschlager, M. & Nora, E. P. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genom. Hum.* **17**, 17–43 (2016).
- Dixon, J. R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).
- Li, Y. et al. Patterns of somatic structural variation in human cancer genomes. *Nature* **578**, 112–121 (2020).
- Kronenberg, Z. N. et al. Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nat. Commun.* **12**, 1935 (2021).
- Wu, S. et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699–703 (2019).
- Zhang, Y. et al. Enhancing Hi-C data resolution with deep convolutional neural network HiCPlus. *Nat. Commun.* **9**, 750 (2018).
- Liu, T. & Wang, Z. HiCNN: a very deep convolutional neural network to better enhance the resolution of Hi-C data. *Bioinformatics* **35**, 4222–4228 (2019).
- Li, Z. & Dai, Z. SRHiC: a deep learning model to enhance the resolution of Hi-C data. *Front. Genet.* **11**, 353 (2020).
- Erdmann-Pham, D. D. et al. Tracing cancer evolution and heterogeneity using Hi-C. *Nat. Commun.* **14**, 7111 (2023).
- Fullwood, M. J. & Ruan, Y. ChIP-based methods for the identification of long-range chromatin interactions. *J. Cell. Biochem.* **107**, 30–39 (2009).
- Fullwood, M. J., Wei, C.-L., Liu, E. T. & Ruan, Y. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* **19**, 521–532 (2009).
- Li, G. et al. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. *Genome Biol.* **11**, 1–13 (2010).
- Fang, R. et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**, 1345–1348 (2016).
- Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
- Zheng, M. et al. Multiplex chromatin interactions with single-molecule precision. *Nature* **566**, 558–562 (2019).
- Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
- Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
- Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
- Tan, L., Xing, D., Chang, C.-H., Li, H. & Xie, X. S. Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018).
- Li, G. et al. Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat. Methods* **16**, 991–993 (2019).
- Lee, D.-S. et al. Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nat. Methods* **16**, 999–1006 (2019).
- Liu, Z. et al. Linking genome structures to functions by simultaneous single-cell Hi-C and RNA-seq. *Science* **380**, 1070–1076 (2023).
- Qu, J. et al. Simultaneous profiling of chromatin architecture and transcription in single cells. *Nat. Struct. Mol. Biol.* **30**, 1393–1402 (2023).
- Xie, X. et al. Simultaneous single-cell three-dimensional genome and gene expression profiling uncovers dynamic enhancer connectivity underlying olfactory receptor choice. Preprint at <https://doi.org/10.21203/rs.3.rs-3210240/v1> (2023).
- Rao, S. S. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320. e324 (2017).
- Canela, A. et al. Genome organization drives chromosome fragility. *Cell* **170**, 507–521. e518 (2017).
- Gryder, B. E., Khan, J. & Stanton, B. Z. Measurement of differential chromatin interactions with absolute quantification of architecture (AQUA-HiChIP). *Nat. Protoc.* **15**, 1209–1236 (2020).
- Lai, B. et al. Trac-looping measures genome structure and chromatin accessibility. *Nat. Methods* **15**, 741–747 (2018).

37. Hughes, J. R. et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
38. Hsieh, T.-H. S. et al. Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162**, 108–119 (2015).
39. Hua, P. et al. Defining genome architecture at base-pair resolution. *Nature* **595**, 125–129 (2021).
40. Sunkel, B. D. et al. Evidence of pioneer factor activity of an oncogenic fusion transcription factor. *iScience* **24**, 102867 (2021).
41. Boone, M. A. et al. The FLI portion of EWS/FLI contributes a transcriptional regulatory function that is distinct and separable from its DNA-binding function in Ewing sarcoma. *Oncogene* **40**, 4759–4769 (2021).
42. Gryder, B. E. et al. Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nat. Genet.* **51**, 1714–1722 (2019).
43. Laubscher, D. et al. BAF complexes drive proliferation and block myogenic differentiation in fusion-positive rhabdomyosarcoma. *Nat. Commun.* **12**, 6924 (2021).
44. Pomella, S. et al. Interaction between SNAI2 and MYOD enhances oncogenesis and suppresses differentiation in fusion negative rhabdomyosarcoma. *Nat. Commun.* **12**, 192 (2021).
45. Yohe, M. E. et al. MEK inhibition induces MYOG and remodels super-enhancers in RAS-driven rhabdomyosarcoma. *Sci. Transl. Med.* **10**, eaan4470 (2018).
46. Yan, J. et al. Systematic analysis of binding of transcription factors to noncoding variants. *Nature* **591**, 147–151 (2021).
47. Baxter, J. S. et al. Capture Hi-C identifies putative target genes at 33 breast cancer risk loci. *Nat. Commun.* **9**, 1028 (2018).
48. Zaret, K. S. Pioneer transcription factors initiating gene network changes. *Annu. Rev. Genet.* **54**, 367–385 (2020).
49. Iyyanki, T. et al. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Genome Biol.* **22**, 1–20 (2021).
50. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
51. Wang, X., Luan, Y. & Yue, F. EagleC: a deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. *Sci. Adv.* **8**, eabn9215 (2022).
52. Spielmann, M., Lupiáñez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).
53. Wang, J. et al. Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Sci. Adv.* **7**, eabg4126 (2021).
54. Montefiori, L. E. et al. Enhancer hijacking drives oncogenic BCL11B expression in lineage-ambiguous stem cell leukemia. *Cancer Discov.* **11**, 2846–2867 (2021).
55. Mallard, C. et al. Hi-C detects genomic structural variants in peripheral blood of pediatric leukemia patients. *Mol. Case Stud.* **8**, a006157 (2022).
56. Voronina, N. et al. The landscape of chromothripsis across adult cancer types. *Nat. Commun.* **11**, 2320 (2020).
57. Sungalee, S. et al. Histone acetylation dynamics modulates chromatin conformation and allele-specific interactions at oncogenic loci. *Nat. Genet.* **53**, 650–662 (2021).
58. Song, T. et al. Integrative identification by Hi-C revealed distinct advanced structural variations in Lung Adenocarcinoma tissue. *Phenomics* **3**, 390–407 (2023).
59. Du, Y. et al. Dynamic Interplay between structural variations and 3D genome organization in pancreatic cancer. *Adv. Sci.* **9**, 2200818 (2022).
60. Akdemir, K. C. et al. Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nat. Genet.* **52**, 294–305 (2020).
61. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
62. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
63. Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381. e320 (2017).
64. Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).
65. Sima, J. et al. Identifying cis elements for spatiotemporal control of mammalian DNA replication. *Cell* **176**, 816–830. e818 (2019).
66. Tanizawa, H., Kim, K.-D., Iwasaki, O. & Noma, K.-I. Architectural alterations of the fission yeast genome during the cell cycle. *Nat. Struct. Mol. Biol.* **24**, 965–976 (2017).
67. Li, Y. et al. Transcription-coupled structural dynamics of topologically associating domains regulate replication origin efficiency. *Genome Biol.* **22**, 1–29 (2021).
68. Wang, X. et al. Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. *Nat. Methods* **18**, 661–668 (2021).
69. Yang, M. et al. Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nat. Commun.* **10**, 1519 (2019).
70. Ooi, W. F. et al. Integrated paired-end enhancer profiling and whole-genome sequencing reveals recurrent CCNE1 and IGF2 enhancer hijacking in primary gastric adenocarcinoma. *Gut* **69**, 1039–1052 (2020).
71. Abdalla, M. O. A. et al. The Eleanor ncRNAs activate the topological domain of the ESR1 locus to balance against apoptosis. *Nat. Commun.* **10**, 3778 (2019).
72. Wu, P. et al. 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. *Nat. Commun.* **8**, 1937 (2017).
73. Szabo, Q. et al. Regulation of single-cell genome organization into TADs and chromatin nanodomains. *Nat. Genet.* **52**, 1151–1157 (2020).
74. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
75. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
76. Liu, F., Wu, D. & Wang, X. Roles of CTCF in conformation and functions of chromosome. *Semin. Cell Dev. Biol.* **90**, 168–173 (2019).
77. Tanwar, V. S., Jose, C. C. & Cuddapah, S. Role of CTCF in DNA damage response. *Mutat. Res.* **780**, 61–68 (2019).
78. Alharbi, A. B., Schmitz, U., Bailey, C. G. & Rasko, J. E. CTCF as a regulator of alternative splicing: new tricks for an old player. *Nucleic Acids Res.* **49**, 7825–7838 (2021).
79. Nakamura, R. et al. CTCF looping is established during gastrulation in medaka embryos. *Genome Res.* **31**, 968–980 (2021).
80. Agrawal, P. & Rao, S. Super-enhancers and CTCF in early embryonic cell fate decisions. *Front. Cell Dev. Biol.* **9**, 653669 (2021).
81. Lyu, X., Rowley, M. J. & Corces, V. G. Architectural proteins and pluripotency factors cooperate to orchestrate the transcriptional response of hESCs to temperature stress. *Mol. Cell* **71**, 940–955. e947 (2018).
82. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110–114 (2016).
83. Alpsoy, A. et al. BRD9 is a critical regulator of androgen receptor signaling and prostate cancer progression. *Cancer Res.* **81**, 820–833 (2021).
84. Wong, K. M., Song, J. & Wong, Y. H. CTCF and EGR1 suppress breast cancer cell migration through transcriptional control of Nm23-H1. *Sci. Rep.* **11**, 491 (2021).
85. Sivapragasam, S. et al. CTCF binding modulates UV damage formation to promote mutation hot spots in melanoma. *EMBO J.* **40**, e107795 (2021).
86. Hou, Y. et al. CTCF mediates replicative senescence through POLD1. *Front. Cell Dev. Biol.* **9**, 618586 (2021).
87. Miyata, K. et al. Pericentromeric noncoding RNA changes DNA binding of CTCF and inflammatory gene expression in senescence and cancer. *Proc. Natl Acad. Sci.* **118**, e2025647118 (2021).
88. Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35–45 (1997).
89. Gassler, J. et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
90. Brackley, C. A. et al. Nonequilibrium chromosome looping via molecular slip links. *Phys. Rev. Lett.* **119**, 138101 (2017).
91. Terakawa, T. et al. The condensin complex is a mechanochemical motor that translocates along DNA. *Science* **358**, 672–676 (2017).
92. Ocampo-Hafalla, M., Muñoz, S., Samora, C. P. & Uhlmann, F. Evidence for cohesin sliding along budding yeast chromosomes. *Open Biol.* **6**, 150178 (2016).
93. De Koninck, M. & Losada, A. Cohesin mutations in cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026476 (2016).
94. Waldman, T. Emerging themes in cohesin cancer biology. *Nat. Rev. Cancer* **20**, 504–515 (2020).
95. Fisher, J. B., McNulty, M., Burke, M. J., Crispino, J. D. & Rao, S. Cohesin mutations in myeloid malignancies. *Trends Cancer* **3**, 282–293 (2017).
96. Bailey, M. L. et al. Paralogous synthetic lethality underlies genetic dependencies of the cancer-mutated gene STAG2. *Life Sci. Alliance* **4**, e202101083 (2021).
97. Surdez, D. et al. STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma. *Cancer Cell* **39**, 810–826. e819 (2021).
98. Carico, Z. M., Stefan, H. C., Justice, M., Yimit, A. & Downen, J. M. A cohesin cancer mutation reveals a role for the hinge domain in genome organization and gene expression. *PLoS Genet.* **17**, e1009435 (2021).
99. Ma, J. et al. miR-223 regulates cell proliferation and invasion via targeting PDS5B in pancreatic cancer cells. *Mol. Ther. Nucleic Acids* **14**, 583–592 (2019).
100. Tothova, Z. et al. Cohesin mutations alter DNA damage repair and chromatin structure and create therapeutic vulnerabilities in MDS/AML. *JCI Insight* **6**, e142149 (2021).
101. Atkin, N. D., Raimier, H. M., Wang, Z., Zang, C. & Wang, Y. H. Assessing acute myeloid leukemia susceptibility in rearrangement-driven patients by DNA breakage at topoisomerase II and CCCTC-binding factor/cohesin binding sites. *Genes Chromosomes Cancer* **60**, 808–821 (2021).

102. Kriz, A. J., Colognori, D., Sunwoo, H., Nabet, B. & Lee, J. T. Balancing cohesin eviction and retention prevents aberrant chromosomal interactions, Polycomb-mediated repression, and X-inactivation. *Mol. Cell* **81**, 1970–1987. e1979 (2021).
103. Gröschel, S. et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* **157**, 369–381 (2014).
104. Valentijn, L. J. et al. TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nat. Genet.* **47**, 1411–1414 (2015).
105. Northcott, P. A. et al. Enhancer hijacking activates GF11 family oncogenes in medulloblastoma. *Nature* **511**, 428–434 (2014).
106. Weischenfeldt, J. et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2017).
107. Jia, Q., Chen, S., Tan, Y., Li, Y. & Tang, F. Oncogenic super-enhancer formation in tumorigenesis and its molecular mechanisms. *Exp. Mol. Med.* **52**, 713–723 (2020).
108. Zhuang, H.-H., Qu, Q., Teng, X.-Q., Dai, Y.-H. & Qu, J. Superenhancers as master gene regulators and novel therapeutic targets in brain tumors. *Exp. Mol. Med.* **55**, 290–303 (2023).
109. Xu, Z. et al. Structural variants drive context-dependent oncogene activation in cancer. *Nature* **612**, 564–572 (2022).
110. Kim, K. et al. Spatial and clonality-resolved 3D cancer genome alterations reveal enhancer-hijacking as a potential prognostic marker for colorectal cancer. *Cell Rep.* **42**, 112778 (2023).
111. Wakita, S. et al. Experimental method for haplotype phasing across the entire length of chromosome 21 in trisomy 21 cells using a chromosome elimination technique. *J. Hum. Genet.* **67**, 565–572 (2022).
112. Edge, P., Bafna, V. & Bansal, V. HapCUT2: robust and accurate haplotype assembly for diverse sequencing technologies. *Genome Res.* **27**, 801–812 (2017).
113. Ijaz, J. et al. Haplotype-specific assembly of shattered chromosomes in esophageal adenocarcinomas. *Cell Genom.* **4**, 100484 (2024).
114. Sakamoto, Y. et al. Phasing analysis of lung cancer genomes using a long read sequencer. *Nat. Commun.* **13**, 3464 (2022).
115. Garg, S. Towards routine chromosome-scale haplotype-resolved reconstruction in cancer genomics. *Nat. Commun.* **14**, 1358 (2023).
116. Snuderl, M. et al. Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* **20**, 810–817 (2011).
117. Szerlip, N. J. et al. Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc. Natl. Acad. Sci.* **109**, 3041–3046 (2012).
118. Kumar, P. et al. ATAC-seq identifies thousands of extrachromosomal circular DNA in cancer and cell lines. *Sci. Adv.* **6**, eaba2489 (2020).
119. Morton, A. R. et al. Functional enhancers shape extrachromosomal oncogene amplifications. *Cell* **179**, 1330–1341. e1313 (2019).
120. Helmsauer, K. et al. Enhancer hijacking determines extrachromosomal circular MYCN amplicon architecture in neuroblastoma. *Nat. Commun.* **11**, 5823 (2020).
121. Misteli, T. The self-organizing genome: principles of genome architecture and function. *Cell* **183**, 28–45 (2020).
122. Kim, T. et al. Comparative characterization of 3D chromatin organization in triple-negative breast cancers. *Exp. Mol. Med.* **54**, 585–600 (2022).
123. Li, P. et al. Alteration of chromatin high-order conformation associated with oxaliplatin resistance acquisition in colorectal cancer cells. *Exploration* **3**, 20220136 (2023).
124. Achinger-Kawecka, J. et al. The potential of epigenetic therapy to target the 3D epigenome in endocrine-resistant breast cancer. *Nat. Struct. Mol. Biol.* **31**, 498–512 (2024).

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation (NRF) of Korea Grants funded by the Korean Government (2021R1C1C1006798 and 2022M3A9D3016848). This work was supported by the 2022 Advanced Facility Fund of the University of Seoul.

FUNDING

National Research Foundation (NRF) of Korea Grants funded by the Korean Government (2021R1C1C1006798 and 2022M3A9D3016848).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Dong-Sung Lee.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024, corrected publication 2024