BMB
Reports

BMB Rep. 2025; 58(5): 203-208
www.bmbreports.org

Contributed Mini Review

Unraveling the three-dimensional genome structure using machine learning

Jiho Lee[1], Hye-Lim Mo[2], Yoon Ha[2], Dong Yeon Nam[2], Geumnim Lim[2], Jeong-Woon Park[2], Seoyoung Park[2], Woo-Young Choi[2], Hyun Ji Lee[2] & Je-Keun Rhee[1,2,]*

[1]School of Systems Biomedical Science, Soongsil University, Seoul 06978, [2]Department of Bioinformatics & Life Science, Soongsil University, Seoul 06978, Korea

The study of chromatin interactions has advanced considerably with technologies such as high-throughput chromosome conformation capture (Hi-C) sequencing, providing a genome-wide view of physical interactions within the nucleus. These techniques have revealed the existence of hierarchical chromatin structures such as compartments, topologically associating domains (TADs), and chromatin loops, which are crucial in genome organization and regulation. However, identifying and analyzing these structural features require advanced computational methods. In recent years, machine learning approaches, particularly deep learning, have emerged as powerful tools for detecting and analyzing structural information. In this review, we present an overview of various machine learning-based techniques for determining chromosomal organization. Starting with the progress in predicting interactions from DNA sequences, we describe methods for identifying various hierarchical structures from Hi-C data. Additionally, we present advances in enhancing the chromosome contact frequency map resolution to overcome the limitations of Hi-C data. Finally, we identify the remaining challenges and propose potential solutions and future directions. [BMB Reports 2025; 58(5): 203-208]

INTRODUCTION

The three-dimensional (3D) organization of the genome is pivotal to orchestrating various biological processes, such as DNA replication, gene regulation, and cellular differentiation (1). Advances in high-throughput chromosome conformation capture (Hi-C) sequencing have revolutionized our understanding of genome organization and structural features at different scales, such as compartments, topologically associating domains (TADs), and chromatin loops, by providing genome-wide contact frequency maps (Fig. 1) (2-5).

At the megabase scale, the genome is partitioned into two types of compartments, A and B (Fig. 1B) (3, 6). A compartments are transcriptionally active regions associated with open chromatin (euchromatin). In contrast, B compartments are transcriptionally inactive regions associated with closed chromatin (heterochromatin). These compartments are represented as characteristic checkerboard or plaid patterns in the contact frequency maps.

TADs, self-interacting chromatin regions, are considered fundamental structural and functional genome units because they constrain the interaction between regulatory elements and their target genes (Fig. 1C) (4, 7). The TAD boundaries are enriched with insulator proteins, such as CTCF, and are characterized by distinct squares along the diagonal of Hi-C contact frequency maps.

At the finest scale, chromatin loops represent long-range spatial interactions between two loci, such as enhancers and promoters (Fig. 1D, E) (8, 9). These loops are formed by binding

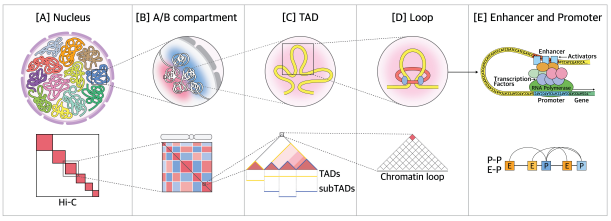

Fig. 1. Multiscale 3D genomic structures identified from Hi-C data. (A) DNA is packed into chromosomal territories in the nucleus. (B) Chromatin consists of two compartments, A and B. Compartments are represented as characteristic checkerboard or plaid patterns in contact frequency maps. (C) TADs and their nested subTADs are self-interacting functional units on a scale ranging from 100 kilobases to a few megabases. These can be identified along the diagonal of the Hi-C contact map. (D) Loops represent long-range spatial interactions on a scale ranging from 10 kilobases to 100 kilobases. The loops exhibit intense signals on the contact frequency map. (E) Enhancer–promoter interactions and promoter-promoter interactions are global long-range contacts for transcription regulation. These interactions are represented as peaks in the probability distribution.

*Corresponding author. Tel: +82-2-828-7038; Fax: +82-2-820-0816; E-mail: jkrhee@ssu.ac.kr

https://doi.org/10.5483/BMBRep.2024-0020

Received 24 January 2024, Revised 7 March 2024, Accepted 6 September 2024, Published online 21 April 2025

Keywords: 3D genome, Chromatin interaction, Deep learning, Hi-C sequencing, Machine learning

ISSN: 1976-670X (electronic edition)

some architectural proteins, such as CTCF, and cohesin appears as intense focal signals on the contact frequency maps.

Interpreting Hi-C data and identifying biologically meaningful regions require advanced computational methods. In particular, machine learning has emerged as a powerful tool for deciphering the intricacies of 3D genome organization (6, 10, 11). Unsupervised learning methods without labeled target values, such as dimensionality reduction and clustering, have been used to identify topological features from contact frequency maps. Supervised learning methods, trained on known chromatin interactions, can predict new potential interactions in unexplored genomic regions.

Furthermore, deep learning (deep neural networks, DNNs) approaches have gained significant attention due to their ability to capture complex nonlinear relationships (Fig. 2). For instance, convolutional neural networks (CNNs) can capture local spatial dependencies and identify intricate patterns in genomic structures (6). Recurrent neural networks (RNNs), and long short-term memory (LSTM) networks, have been employed to model the sequential nature of the genome (6). Moreover, generative models, such as generative adversarial networks (GANs), learn the underlying distribution of the Hi-C data (12).

In this review, we provide an overview of the current machine learning-based techniques for analyzing 3D genome structures (Fig. 3). Furthermore, we describe approaches to overcome some limitations in fully understanding genome organization, including the challenges posed by low-resolution Hi-C data.

CHROMATIN INTERACTION PREDICTION FROM DNA SEQUENCES

The advent of various machine learning methods has enabled chromatin interaction prediction from DNA sequences, providing valuable insights into the relationships between genome sequences and 3D genome organization (Fig. 3A, Supplementary Table 1). The models predict chromatin interactions at various scales, ranging from enhancer–promoter interactions (EPIs) to TADs and compartments.

SPEID employs a combination of CNN and LSTM structures to predict EPIs based on sequence features (13). Zhuang *et al.* proposed a simpler CNN-based model that achieved comparable performance with reduced computational costs (14). Moreover, the model substantially improved the predictive performance by applying transfer learning approaches. DeepTACT predicts both EPIs and promoter-promoter interactions by incorporating the regulatory sequences and their chromatin accessibility scores by combining CNN and attention-based bidirectional LSTM (BiLSTM) structures (15). SEPT employs a similar architecture by combining CNNs and LSTM layers but introduces a transfer learning strategy and addresses the challenge of model generalization across cell types (16). The chromatin interaction neural network (ChINN), another CNN-based architecture, can predict interactions between open chromatin regions (17). Furthermore, DeepMILO and DeepCTCFLoop were developed to predict CTCF-mediated chromatin loop formation from DNA sequences by integrating a CNN and a BiLSTM (18, 19).

In addition to predicting the specific interactions between sequence pairs, several computational models have attempted to predict the contact frequency map from raw DNA sequences. Akita utilizes high-resolution Hi-C and Micro-C data as training targets to predict the contact frequency maps for a sequence region up to 1 Mb in length based on CNN structures (20). DeepC predicts genome folding from Mb-scale DNA sequences by employing CNNs with transfer learning (21). Orca, consisting of hierarchical encoder and decoder components with convolutional blocks, represents a significant advancement in multiscale chromatin interaction prediction (22). Orca can predict EPIs and TADs along with compartments A and B observed at the Mb scale by expanding the sequence context.

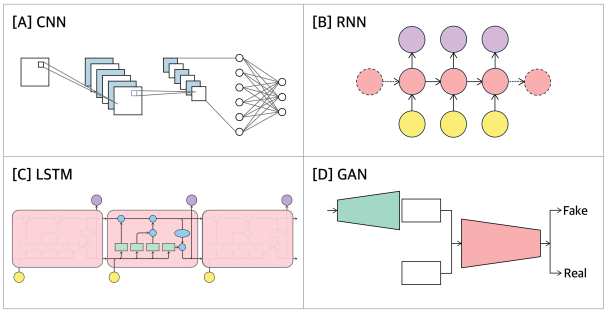

Fig. 2. Schematic representation of key deep learning architectures used in Hi-C data analysis. (A) CNN captures local patterns and spatial relationships through convolutional operations. (B) RNN processes sequential data by retaining information from previous inputs, enabling the learning of temporal patterns. (C) LSTM network is similar to RNN, but the internal structure of the LSTM cell allows for long-term dependency learning. (D) GAN implements a competitive learning process between a generator and discriminator.

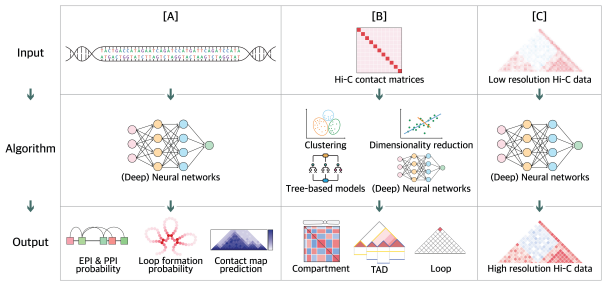

Fig. 3. Unlocking 3D genome structures using machine learning. (A) 3D organization prediction from DNA sequence data. (B) Prediction of chromatin organization, such as compartments, TADs, and loops, from Hi-C contact frequency matrices. (C) Enhancing the Hi-C data resolution.

COMPARTMENTS A AND B DETECTION FROM Hi-C DATA

Various computational methods have been developed to detect distinct chromatin structures from contact frequency maps generated by Hi-C (Fig. 3B). Understanding the arrangement of compartments A and B typically involves eigenvector analysis of normalized chromatin contact frequency matrices. The change in the first eigenvector sign indicates the boundary between the two compartments (23, 24). For example, several tools, such as HOMER, Juicer, and Pentad, utilize principal component analysis (PCA) to annotate and visualize the compartments in Hi-C contact frequency matrices (25-27). Calder employs Fisher's z-transformed correlation and hierarchical clustering to infer the chromatin compartments (28). dcHiC identifies significant changes in compartmentalization across multiple contact frequency maps via Fisher's z-transformed correlation and parallelized partial singular value decomposition (SVD) (29). CscoreTool defines C-scores derived from log-likelihood function, providing the probability of assigning a genomic window to compartment A (30).

SNIPER combines a denoising autoencoder and a multilayer perceptron classifier for compartment annotation (31). It directly imputes interchromosomal contact frequency maps and identifies subcompartments even with low-coverage Hi-C data by addressing the noise and bias in the data. However, the performance can be parameter-dependent, limiting its generalizability to diverse datasets. Supplementary Table 2 shows the machine learning-based tools used to detect compartments A and B.

DETECTING TADS AND SUBTADS FROM Hi-C DATA

TADs and subTADs can be detected based on the distribution of chromosomal contact frequency matrices (Fig. 3B, Supplementary Table 3). Earlier models, such as Armatus, insulation score, TopDom, and HiCDB, detected TAD boundaries by identifying abrupt changes in contact frequencies at the boundaries based on the assumption that intra-TAD interactions are more frequent than inter-TAD interactions (32-35). However, more recently, unsupervised learning techniques have been adopted predominantly to capture distinctive contact patterns. IC-Finder uses hierarchical clustering to infer TAD boundaries (36). SpectralTAD represents the contact frequencies between genomic loci as nodes and edges, and TAD boundaries are identified by a graph-based spectral clustering framework (37).

Some models discern the hierarchical organization of TADs and subTADs. ClusterTAD employs various clustering methods, such as hierarchical clustering, expectation-maximization, and k-means clustering. Additionally, ClusterTAD automatically determines the number of clusters and distinguishes between larger TADs and smaller subTADs by iteratively splitting larger clusters into smaller clusters (38). TADtree assumes that contact enrichment increases proportionally with TAD length and detects the nested TAD hierarchy (39). OnTAD employs a two-phase approach for identifying nested TADs (40). In the first phase, potential TAD boundaries are identified via an adaptive local minimum search. These candidates are subsequently combined to form hierarchical TAD structures via a recursive dynamic programming algorithm.

Other advanced machine learning techniques have also been leveraged to predict TAD boundaries. preciseTAD uses transfer learning and a random forest model (41). Despite first being trained on low-resolution Hi-C data, the model effectively enables TAD detection at the base-pair level via transfer learning. TADL transforms the TAD identification problem into an image classification task by utilizing a CNN-based model that integrates a residual neural network (ResNet) and a squeeze-and-excitation network (SENet) (42). RefHiC adopts an attention-based encoder architecture to learn topological patterns (43). RefHiC enhances its generalization performance and reliability in detecting TADs by using a wide variety of Hi-C datasets across various cell types, species, and sequencing depths. But the performance varies in terms of computational efficiency and the characteristics of identified TADs. For example, RefHiC identifies larger TADs in general, while preciseTAD requires more computational time.

CHROMATIN LOOP DETECTION VIA Hi-C

Chromatin loop identification generally involves comparing contact-enriched regions against their background expectations based on random interactions and genomic distances (Fig. 3B, Supplementary Table 4). Representative early methods, such as Fit-Hi-C and HiCCUPS, employ statistical significance tests with different background estimation approaches; Fit-Hi-C utilizes a genome-wide model, whereas HiCCUPS applies a local context (44, 45).

Some approaches have incorporated computer vision techniques and various image analysis methods by treating Hi-C data as images to detect chromatin loops. For instance, Mustache employs a scale-space representation that applies Gaussian kernels of increasing width to the contact frequency maps to detect loops of various sizes (46). Chromosight segments the contact frequency maps into smaller squares and computes the correlation between each square and a loop template, filtering squares with high correlation values to determine the final loop calls (47). SIP enhances the loop signal and reduces background noise by applying general image processing steps (48).

Supervised learning approaches and DNNs have also been applied to chromatin loop detection. Peakachu utilizes a random forest classification framework to predict chromatin loops in genome-wide contact maps (49). The model is trained on a set of known loops using features derived from the contact map, such as the interaction intensity and the distance between the interacting loci. DLoopCaller is a CNN-based model that takes Hi-C contact frequency maps as input and predicts the probability of each pixel being a chromatin loop (50). GILoop utilizes a dual-branch neural model consisting of

a U-Net branch that learns pixelwise image features and a graph convolutional network (GCN) branch that learns edgewise features, where the edge represents the interaction of contact frequency maps at a graph representation (51). DeepLoop employs a denoising autoencoder and U-Net structures, reducing bias and noise in Hi-C data and enabling robust loop signal detection even in low-depth and single-cell Hi-C data (52). These learning approaches can potentially improve the accuracy of the loop detection, effectively reflecting the nonlinear relationships and complex interaction patterns.

Hi-C DATA RESOLUTION ENHANCEMENT

As interest in the regulation of biological processes within cell nuclei grows, demand for high-quality 3D genome structure data is increasing. Although methods for identifying genomic interactions have become more accurate and refined, their prediction power is often limited by the quality of the original data (53, 54). To overcome these limitations, numerous methods have been developed to enhance the resolution of Hi-C data by amplifying positive signals and removing technical biases (Fig. 3C, Table 1).

Deep learning approaches, which are commonly used for image manipulation, can be utilized to enhance Hi-C data resolution. HiCPlus employs CNN structures that are trained using low-resolution data sampled from high-resolution data (55). A deeper convolutional model, HiCNN, incorporates residual layers to address the vanishing gradient caused by the increase in computational network depth (56). SRHiC employs an architecture that combines fewer convolution layers with residual blocks and skip connections to reduce computational resources (57). HiCARN uses cascading residual networks to effectively capture the complex features and representations of data (58).

Some models have applied GANs to generate refined data from low-resolution contact frequency maps. hicGAN generates high-resolution data that closely resemble the original data at the pixel level through adversarial training (59). DeepHiC adopts a loss function that combines adversarial loss, pixelwise error, total variation, and perceptual loss to generate high-resolution data (60). EnHiC leverages the nonnegative and symmetric properties of Hi-C data to extract multiscale features that capture the hierarchical structures of Hi-C contact frequency maps (61). EnHiC can predict high-resolution Hi-C contact frequency maps, by employing a series of subpixel convolutional layers and adversarial training, combined with these features from different scales, demonstrating its applicability to various cell types and sequencing depths.

Some models that aim to reduce noise from data have also been developed. HiCSR integrates a denoising autoencoder into a GAN to be robust to noise-corrupted Hi-C data (62). VEHiCLE employs a variational autoencoder (VAE) and adversarial training strategy and introduces chromosome topology-inspired insulation loss (63). iEnhance consists of a dense encoder and decoder structure designed to extract multiscale global and local features, enabling the capture of hierarchical structures even with sparse information, such as single-cell Hi-C and Micro-C data (64).

These deep learning models differ in their architectural design and computational demands, which affects their training and application characteristics. For instance, DeepHiC requires more training iterations to achieve convergence, while iEnhance has higher memory requirements for both its training process and predictive applications.

DISCUSSION

We reviewed the latest computational approaches for analyzing 3D genome organization, with a particular focus on the application of machine learning. Machine learning techniques have greatly enhanced our ability to detect and characterize complex patterns in Hi-C images and similar contact frequency maps.

Despite the remarkable results achieved by applying machine learning in 3D genome research, several major challenges remain to be addressed. One issue is the lack of high-resolution data from Hi-C experiments. Developing more precise Hi-C data resolution enhancement methods will become increasingly important to overcome this limitation. Moreover, the robustness of the computational models should be improved across varying sequencing depths, resolutions, and cell types. The recent emergence of single-cell Hi-C sequencing will make the development of ad-

Table 1. Methods for Hi-C data resolution enhancement

Tool	Approach	Github
HiCPlus	CNN	https://github.com/wangjuan001/hicplus
HiCNN2	CNN	http://dna.cs.miami.edu/HiCNN2/
hicGAN	GAN, CNN	https://github.com/kimmo1019/hicGAN
DeepHiC	GAN, CNN	https://github.com/omegahh/DeepHiC
HiCSR	GAN, Denoising autoencoder	https://github.com/PSI-Lab/HiCSR
SRHiC	CNN	https://github.com/hzlzldr/SRHiC
VEHiCLE	GAN, CNN, VAE	https://github.com/Max-Highsmith/VEHiCLE
HiCARN	GAN, CNN	https://github.com/OluwadareLab/HiCARN
EnHiC	GAN, CNN	https://github.com/wmalab/EnHiC
iEnhance	Encoder-decoder, CNN	https://github.com/onlybugs/iEnhance

vanced computational techniques even more critical (65, 66). In addition, the use of machine learning algorithms to investigate the dynamic interplay of chromatin in 3D space over time and analyze temporal datasets will be essential for obtaining a more comprehensive understanding of genome organization (67).

Another promising issue is the integration of multimodal data, such as imaging, genomic, epigenomic, and transcriptomic information (12). Developing machine learning frameworks that can effectively incorporate these multimodal datasets can provide new insights into unraveling the complex interplay between 3D genome organization and gene regulation.

The interpretability of machine learning models also remains a critical challenge. The development of explainable AI techniques tailored to 3D genome data is necessary to better understand the biological significance of the features extracted by these models.

In conclusion, further advances in cutting-edge machine learning techniques will lead to a comprehensive understanding of genome organization and provide deeper insights into the biological functions and underlying basic mechanisms associated with various diseases.

CONFLICTS OF INTEREST

The authors have no conflicting interests.

REFERENCES

1. Li Y, Hu M and Shen Y (2018) Gene regulation in the 3D genome. Hum Mol Genet 27, R228-R233
2. Lee BH and Rhie SK (2021) Molecular and computational approaches to map regulatory elements in 3D chromatin structure. Epigenetics Chromatin 14, 14
3. Lieberman-Aiden E, van Berkum NL, Williams L et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289-293
4. Dixon JR, Selvaraj S, Yue F et al (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376-380
5. Pombo A and Dillon N (2015) Three-dimensional genome architecture: players and mechanisms. Nat Rev Mol Cell Biol 16, 245-257
6. Zhang Y, Boninsegna L, Yang M, Misteli T, Alber F and Ma J (2024) Computational methods for analysing multiscale 3D genome organization. Nat Rev Genet 25, 123-141
7. Austenaa LM, Barozzi I, Simonatto M et al (2015) Transcription of mammalian cis-regulatory elements is restrained by actively enforced early termination. Mol Cell 60, 460-474
8. Grubert F, Srivas R, Spacek DV et al (2020) Landscape of cohesin-mediated chromatin loops in the human genome. Nature 583, 737-743
9. Kadauke S and Blobel GA (2009) Chromatin loops in gene regulation. Biochim Biophys Acta 1789, 17-25
10. Pal K, Forcato M and Ferrari F (2019) Hi-C analysis: from data generation to integration. Biophys Rev 11, 67-78
11. Raffo A and Paulsen J (2023) The shape of chromatin: insights from computational recognition of geometric patterns in Hi-C data. Brief Bioinform 24, bbad302
12. Gong H, Yang Y, Zhang S, Li M and Zhang X (2021) Application of Hi-C and other omics data analysis in human cancer and cell differentiation research. Comput Struct Biotechnol J 19, 2070-2083
13. Ganji M, Shaltiel IA, Bisht S et al (2018) Real-time imaging of DNA loop extrusion by condensin. Science 360, 102-105
14. Zhuang Z, Shen X and Pan W (2019) A simple convolutional neural network for prediction of enhancer-promoter interactions with DNA sequence data. Bioinformatics 35, 2899-2906
15. Li W, Wong WH and Jiang R (2019) DeepTACT: predicting 3D chromatin contacts via bootstrapping deep learning. Nucleic Acids Res 47, e60
16. Jing F, Zhang SW and Zhang S (2020) Prediction of enhancer-promoter interactions using the cross-cell type information and domain adversarial neural network. BMC Bioinformatics 21, 507
17. Cao F, Zhang Y, Cai Y et al (2021) Chromatin interaction neural network (ChINN): a machine learning-based method for predicting chromatin interactions from DNA sequences. Genome Biol 22, 226
18. Trieu T, Martinez-Fundichely A and Khurana E (2020) DeepMILO: a deep learning approach to predict the impact of non-coding sequence variants on 3D chromatin structure. Genome Biol 21, 79
19. Kuang S and Wang L (2021) Deep learning of sequence patterns for CCCTC-binding factor-mediated chromatin loop formation. J Comput Biol 28, 133-145
20. Fudenberg G, Kelley DR and Pollard KS (2020) Predicting 3D genome folding from DNA sequence with Akita. Nat Methods 17, 1111-1117
21. Schwessinger R, Gosden M, Downes D et al (2020) DeepC: predicting 3D genome folding using megabase-scale transfer learning. Nat Methods 17, 1118-1124
22. Zhou J (2022) Sequence-based modeling of three-dimensional genome architecture from kilobase to chromosome scale. Nat Genet 54, 725-734
23. Mohanta TK, Mishra AK and Al-Harrasi A (2021) The 3D genome: from structure to function. Int J Mol Sci 22, 11585
24. Fortin JP and Hansen KD (2015) Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. Genome Biol 16, 180
25. Durand NC, Shamim MS, Machol I et al (2016) Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst 3, 95-98
26. Magnitov MD, Garaev AK, Tyakht AV, Ulianov SV and Razin SV (2022) Pentad: a tool for distance-dependent analysis of Hi-C interactions within and between chromatin compartments. BMC Bioinformatics 23, 116
27. Heinz S, Texari L, Hayes MGB et al (2018) Transcription elongation can affect genome 3D structure. Cell 174, 1522-1536.e22
28. Liu Y, Nanni L, Sungalee S et al (2021) Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. Nat Commun 12, 2439
29. Chakraborty A, Wang JG and Ay F (2022) dcHiC detects differential compartments across multiple Hi-C datasets.

Nat Commun 13, 6827
30. Zheng X and Zheng Y (2018) CscoreTool: fast Hi-C compartment analysis at high resolution. Bioinformatics 34, 1568-1570
31. Xiong K and Ma J (2019) Revealing Hi-C subcompartments by imputing inter-chromosomal chromatin interactions. Nat Commun 10, 5069
32. Filippova D, Patro R, Duggal G and Kingsford C (2014) Identification of alternative topological domains in chromatin. Algorithms Mol Biol 9, 14
33. Crane E, Bian Q, McCord RP et al (2015) Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature 523, 240-244
34. Shin H, Shi Y, Dai C et al (2016) TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res 44, e70
35. Chen F, Li G, Zhang MQ and Chen Y (2018) HiCDB: a sensitive and robust method for detecting contact domain boundaries. Nucleic Acids Res 46, 11239-11250
36. Haddad N, Vaillant C and Jost D (2017) IC-Finder: inferring robustly the hierarchical organization of chromatin folding. Nucleic Acids Res 45, e81
37. Cresswell KG, Stansfield JC and Dozmorov MG (2020) SpectralTAD: an R package for defining a hierarchy of topologically associated domains using spectral clustering. BMC Bioinformatics 21, 319
38. Oluwadare O and Cheng J (2017) ClusterTAD: an unsupervised machine learning approach to detecting topologically associated domains of chromosomes from Hi-C data. BMC Bioinformatics 18, 480
39. Weinreb C and Raphael BJ (2016) Identification of hierarchical chromatin domains. Bioinformatics 32, 1601-1609
40. An L, Yang T, Yang J et al (2019) OnTAD: hierarchical domain structure reveals the divergence of activity among TADs and boundaries. Genome Biol 20, 282
41. Stilianoudakis SC, Marshall MA and Dozmorov MG (2022) preciseTAD: a transfer learning framework for 3D domain boundary prediction at base-pair resolution. Bioinformatics 38, 621-630
42. Yang JY and Chang JM (2022) Pattern recognition of topologically associating domains using deep learning. BMC Bioinformatics 22, 634
43. Zhang Y and Blanchette M (2022) Reference panel guided topological structure annotation of Hi-C data. Nat Commun 13, 7426
44. Ay F, Bailey TL and Noble WS (2014) Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res 24, 999-1011
45. Rao SS, Huntley MH, Durand NC et al (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665-1680
46. Roayaei Ardakany A, Gezer HT, Lonardi S and Ay F (2020) Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. Genome Biol 21, 256
47. Matthey-Doret C, Baudry L, Breuer A et al (2020) Computer vision for pattern detection in chromosome contact maps. Nat Commun 11, 5795
48. Rowley MJ, Poulet A, Nichols MH et al (2020) Analysis of Hi-C data using SIP effectively identifies loops in organisms from C. Elegans to mammals. Genome Res 30, 447-458
49. Salameh TJ, Wang X, Song F et al (2020) A supervised learning framework for chromatin loop detection in genome-wide contact maps. Nat Commun 11, 3428
50. Wang S, Zhang Q, He Y et al (2022) DLoopCaller: a deep learning approach for predicting genome-wide chromatin loops by integrating accessible chromatin landscapes. PLoS Comput Biol 18, e1010572
51. Wang F, Gao T, Lin J et al (2022) GILoop: robust chromatin loop calling across multiple sequencing depths on Hi-C data. iScience 25, 105535
52. Zhang S, Plummer D, Lu L et al (2022) DeepLoop robustly maps chromatin interactions from sparse allele-resolved or single-cell Hi-C data at kilobase resolution. Nat Genet 54, 1013-1025
53. Lee BH, Wu Z and Rhie SK (2022) Characterizing chromatin interactions of regulatory elements and nucleosome positions, using Hi-C, Micro-C, and promoter capture Micro-C. Epigenetics Chromatin 15, 41
54. Lazaris C, Kelly S, Ntziachristos P, Aifantis I and Tsirigos A (2017) HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. BMC Genomics 18, 22
55. Zhang Y, An L, Xu J et al (2018) Enhancing Hi-C data resolution with deep convolutional neural network HiCPlus. Nat Commun 9, 750
56. Liu T and Wang Z (2019) HiCNN: a very deep convolutional neural network to better enhance the resolution of Hi-C data. Bioinformatics 35, 4222-4228
57. Li Z and Dai Z (2020) SRHiC: a deep learning model to enhance the resolution of Hi-C data. Front Genet 11, 353
58. Hicks P and Oluwadare O (2022) HiCARN: resolution enhancement of Hi-C data using cascading residual networks. Bioinformatics 38, 2414-2421
59. Liu Q, Lv H and Jiang R (2019) hicGAN infers super resolution Hi-C data with generative adversarial networks. Bioinformatics 35, i99-i107
60. Hong H, Jiang S, Li H et al (2020) DeepHiC: a generative adversarial network for enhancing Hi-C data resolution. PLoS Comput Biol 16, e1007287
61. Hu Y and Ma W (2021) EnHiC: learning fine-resolution Hi-C contact maps using a generative adversarial framework. Bioinformatics 37, i272-i279
62. Dimmick MC, Lee LJ and Frey BJ (2020) HiCSR: a Hi-C super-resolution framework for producing highly realistic contact maps. bioRxiv 2020.02.24.961714
63. Highsmith M and Cheng J (2021) VEHiCLE: a variationally encoded Hi-C loss enhancement algorithm for improving and generating Hi-C data. Sci Rep 11, 8880
64. Li K, Zhang P, Wang Z et al (2023) iEnhance: a multi-scale spatial projection encoding network for enhancing chromatin interaction data resolution. Brief Bioinform 24, bbad245
65. Zhou T, Zhang R and Ma J (2021) The 3D genome structure of single cells. Annu Rev Biomed Data Sci 4, 21-41
66. Galitsyna AA and Gelfand MS (2021) Single-cell Hi-C data analysis: safety in numbers. Brief Bioinform 22, bbab316
67. Dekker J, Alber F, Aufmkolk S et al (2023) Spatial and temporal organization of the genome: current state and future aims of the 4D nucleome project. Mol Cell 83, 2624-2640