Mapping nucleosome and chromatin architectures: A survey of computational methods

Kun Fang a, Junbai Wang b, Lu Liu c, Victor X. Jin a,⇑

aInstitute for Health and Equity, MCW Cancer Center, and Mellowes Center for Genome Science and Precision Medicine, Medical College of Wisconsin, Milwaukee, WI, USA
bDepartment of Clinical Molecular Biology, University of Oslo and Akershus University Hospital, Lørenskog, Norway
cDepartment of Computer Science, North Dakota State University, Fargo, ND, USA

Article info
Article history:
Received 8 March 2022
Received in revised form 22 July 2022
Accepted 22 July 2022
Available online 26 July 2022

Keywords:
Computational methods
Nucleosome
Chromatin architectures

Abstract
With ever-growing genomic sequencing data, the data variabilities and the underlying biases of the sequencing technologies pose significant computational challenges ranging from the need for accurately detecting the nucleosome positioning or chromatin interaction to the need for developing normalization methods to eliminate systematic biases. This review mainly surveys the computational methods for mapping the higher-resolution nucleosome and higher-order chromatin architectures. While a detailed discussion of the underlying algorithms is beyond the scope of our survey, we have discussed the methods and tools that can detect the nucleosomes in the genome, then demonstrated the computational methods for identifying 3D chromatin domains and interactions. We further illustrated computational approaches for integrating multi-omics data with Hi-C data and the advance of single-cell (sc)Hi-C data analysis. Our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures as well as for computational scientists who are interested in improving upon them.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents
1. Introduction ..................................................................................................... 3957
2. Computational methods in detecting nucleosomes and characterizing nucleosome organization .......................................................... 3957
3. Computational methods in mapping 3D chromatin domains and interactions .................................................................................. 3959
4. Computational approaches for integrating multi-omics data including Hi-C data. .................................................................................. 3960
5. Advances in scHi-C computational analyses ............................................. 3960
6. Conclusions and future perspectives .......................................................... 3960
Declaration of Competing Interest ................................................................. 3960
Acknowledgments ......................................................................................... 3960
References ................................................................................................... 3960

Abbreviations: 3D, three dimensional; ATAC-seq, assay for transposase-accessible chromatin using sequencing; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; ChIP-seq, chromatin immunoprecipitation sequencing; CTCF, CCCTC-binding factor; ICE, iterative correction; Mb, megabases; Micro-C, micrococcal nuclease applied chromosome conformation capture; MNase, micrococcal nuclease; MPE-seq, methidiumpropyl-EDTA sequencing; PCA, principle component analysis; PF, pioneer factor; TAD, topologically associating domain; TCC, tethered chromosome conformation capture; TF, transcription factor.

⇑Corresponding author.
E-mail address: vjin@mcw.edu (V.X. Jin).

https://doi.org/10.1016/j.csbj.2022.07.037
2001-0370/© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Fig. 1. Multiple scales of DNA folding in the nucleus (cartoons in the left panels), from the small scale organization of nucleosome array to the median scaling structure of loop, and the large scale structure of TAD as well as compartment A/B.
1. Introduction

A nucleosome is the unitary structure of the chromatin fiber, composed of two copies of each of four histone proteins, H2A, H2B, H3, and H4, and a ~147 bp DNA wrapped around the histone octamer [1]. Nucleosomes along with other proteins or complexes such as the structural insulator CCCTC-binding factor (CTCF) [2,3] and long-range interaction mediator YY1 [4] facilitate the assembly of three-dimensional (3D) chromatin structure. 3D chromatin structures, ranging from megabases (Mb)-scaled chromatin compartment, kilobases (Kb)-Mb-scaled topologically associating domain (TAD), to Kb-scaled chromatin loop, play critical roles in gene regulation [5], DNA replication [6], cell development and differentiation [7]. Recent advance in sequencing technologies allows us to map the nucleosome organization and 3D chromatin structure (Fig. 1). MNase-seq, a profiling method for mapping nucleosome landscape, is the most prevalent technique to study nucleosome organization. In addition, several other experimental methods were developed to tackle the nucleosome arrangement in the nucleus, including Array-seq [8], Methylation foot-printing [9], Chemical mapping [10], ATAC-seq [11] and MPE-seq [12]. These alternative approaches adapted the idea of cleaving DNA between the nucleosomes but provided extra information. For example, chemical mapping cleaves the DNA by introducing cysteines in histones H3/H4 and can detect nucleosome preferred positions. MPE-seq uses MPE-Fe(II) to digest the DNA and can find the presence of core histones. Chromosome conformation capture (3C) [13] and its derivatives, including but not limited to circular 3C (4C) [14], 3C carbon copy (5C) [15], Capture-C [16], Hi-C [17], TCC [18], and Micro-C [19], have been widely used to study chromatin structures.

With ever-growing genomic sequencing data, the data variabilities and the underlying biases hindered researchers from acquiring biological information accurately. For instance, MNase cleaves DNA about 30 times slower upstream of an A/G-C than 5’ of a A/T, which causes sequence bias for MNase-seq [20–22]. And Yaffe et al reported systematic bias in Hi-C dataset, including different length of restriction fragments caused by ligation efficiencies, different sequencing amplifications affected by GC contents, and differential mappability of sequences [23]. Those data variabilities and biases pose significant computational challenges ranging from the need for accurately detecting the nucleosome positioning to developing normalization methods to eliminate systematic biases.

Enormous efforts at the computational aspect have been put into processing the sequencing data and mapping the nucleosome organization and chromatin structure. These computational efforts include the development of novel computational methods, the application of sophisticated algorithms, the modeling of the distinct distribution of the sequencing data and the extraction of underlying data features. For example, NPS/INPS [24,25] applied a gaussian convolution algorithm to detect inflection points to find candidate nucleosomes, and a novel iterative correction (ICE) algorithm [26] was developed for balancing the biases in Hi-C data. The computational algorithms are often further implemented as software tools for user-friendly and interactive interfaces or downloadable executable files in public code repositories, such as NuChMM [27] for identifying nucleosome states, JuiceBox [28] for visualizing and analyzing Hi-C data, and HiSIF [29] for detecting significant interacting fragments. In addition, a series of methods and tools are sometimes combined into a workflow/pipeline to achieve its meaningful biological output. For instance, HiC-Pro [30], a popular pipeline in Hi-C analysis, combines Bowtie2 (a mapping tool) [31], ICE, ggplot2 [32] (a visualization package) and other customized scripts for Hi-C data processing.

This review mainly focuses on the computational methods available for mapping the higher-resolution nucleosome and higher-order chromatin architectures, and the summaries of the features of the methods. While a detailed discussion of the underlying algorithms is beyond the scope of our survey, we will first discuss the methods and tools that can detect the nucleosomes in the genome, then demonstrate the computational methods for identifying 3D chromatin domains and interactions. We further illustrate computational approaches for integrating multi-omics data with Hi-C data and the advance of single-cell (sc)Hi-C data analysis. Our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures and computational scientists interested in improving upon them.

2. Computational methods in detecting nucleosomes and characterizing nucleosome organization

Accumulated evidence showed that nucleosomes and nucleosome organization played essential roles in transcriptional regulation by interplaying with pioneer factors (PFs), transcription factors (TFs), and ATP-dependent chromatin remodelers [33,34]. Nucleosome organization, which is characterized by nucleosome positioning, spacing and phasing, is defined as the following (Fig. 1): 1) nucleosome positioning describes the consistency of the nucleosomes’ position in the population of cells, where nucleosome position is referred to the location of the histone-octamer relative to DNA; 2) nucleosome spacing describes the distance between the dyads of two adjacent nucleosomes; 3) nucleosome phasing describes the degree of periodicity of the nucleosome array, where nucleosome array is referred to a string of nucleosomes with no more than 350 bp spacing. We didn’t list nucleosome occupancy as a nucleosome organization feature because it is included in measuring nucleosome positioning. Since MNase-seq is the most prevalent and adaptable experimental method to detect genome-wide nucleosomes, in this section, we were mainly focused on summarizing computational methods for processing MNase-seq data. Unlike typical ChIP-seq data of TFs or histone modifications, MNase-seq data tend to have a more noised background signal. Thus classic ChIP-seq peak callers, such as MACS2 [35] and SPP [36], are not suitable for locating genome-wide nucleosomes. Although the basic idea for detecting nucleosomes is still the identification of enriched regions (peaks), the algorithmic design’s goal is instead to improve the signal-to-noise ratio and refine the peak’s width to a specific size, 147 ± 30 bp. Several methods based on the above algorithmic design have been developed, including ChIPseqg [37], iNPS [24], NORMAL [38], NPS [29], Nseq [39], NucDe [40], NucHMM [27], NucHunter [41], NucleoFinder [42], nucLet [43], PING1/2 [44], PUFFIN [45]. Further, some other methods have been focused on identifying the changes of nucleosome positioning among cell types, e.g., BINoch [46], DANPOS/DANPOS2 [47], DiNuP [48] Dimnp [49] and Nucleosome Dynamics [50]. In addition, some methods, such as NucPosSimulator [51] and cplate [52], used Monte Carlo simulations and Template-Based Bayesian to identify alternative or averaged nucleosome positions in cell populations. Few methods, e.g., NucHMM and Nucleosome Dynamic, have lately been developed to further infer nucleosome organization, including nucleosome positioning, spacing and phasing. It is worth mentioning several other methods, TemplateFilter [53], a highly-cited method for analyzing nucleosome positioning from the tiling microarray data, though outdated but usually used as the control in performance comparison, and DeNORA [54] and NucleoATAC [11], designed based on ATAC-seq data. A detailed summary of these methods is listed in Table 1. Concerning the comparisons for the performance of nucleosome
| Tools | Description | Programming Language | Input data Type | Differential analysis | Quantitative Nucleosome organization Information | Webpage |
|-------|-------------|----------------------|----------------|----------------------|-----------------------------------------------|---------|
| BiNoch | Identifying differential nucleosome occupancy regions with nucleosome stabilization–destabilization (NSD) score. | Python | Single-end (SE) Pair-end (PE) MNase-seq MNase-seq | Yes | None | https://liulab-dfci.github.io/software/ |
| ChIPseqR | A binding-event description model to locate nucleosomes, which is also flexible to handle other types of experiments. | R | SE MNase-seq ChIP-seq ChIP-seq | No | Nucleosome positioning score (binding score in the package) | https://www.bioconductor.org/packages/release/bioc/html/ChIPseqR.html |
| DANPOS2 | A toolkit for Dynamic Analysis of Nucleosome and Protein Occupancy by Sequencing, version 2. | Python | SE/PE MNase-seq ChIP-seq MNase-seq | Yes | Nucleosome positioning score | https://sites.google.com/site/danposd/ |
| deNOPA | Decoding nucleosome positions with ATAC-seq data. | Python | ATAC-seq ATAC-seq ATAC-seq | No | None | https://gitee.com/bxxu/denopa/ |
| Dimnp | Identifying regions with differential nucleosome occupancy in multiple samples using Chi-squared test. | Python/ Matlab | Not Clear | Yes | Nucleosome positioning | https://bioinfo.seu.edu.cn/ Nu_dynamics_data_public/ |
| DiNuP | A systematic approach to identify regions of differential nucleosome positioning (BDNP). | Python | SE MNase-seq SE MNase-seq | Yes | Nucleosome positioning score | https://zhanglab.tongji.edu.cn/softwares/DiNuP/download.html |
| iNPS | An improved version of NPS, which outperforms latter one and provides additional nucleosome features. | Python | SE/PE MNase-seq SE/PE MNase-seq | No | None | https://www.pich.ac.cn/hanlab/NPS.html |
| NOrMAL | Using a modified Gaussian mixture model to identify nucleosome positions. | C++ | SE/PE MNase-seq | No | Nucleosome positioning (Fuzziness) | https://github.com/antonpolishko/NOrMAL |
| NPS | A signal processing-based algorithm for identifying positioned nucleosomes from sequencing experiments at the nucleosome level. | Python | SE MNase-seq | No | None | https://liulab-dfci.github.io/software/ |
| Nseq | A multithreaded Java application for finding positioned nucleosomes from sequencing data. | Java | SE MNase-seq | No | Nucleosome positioning score | https://github.com/songlab/Nseq |
| NucDe | A Non-homogeneous hidden-State model on first order differences for automatic detection of nucleosome positions. | R | SE MNase-seq MNase-Chip MNase-seq | No | None | https://pages.stat.wisc.edu/~keles/Software/demo_Nucde.pdf |
| NucDyn | Based on nucleR and aimed at comparing the reads of two MNase-seq experiments for nucleosome positioning and detecting significant inclusions, evictions and shifts. | R | SE/PE MNase-seq MNase-seq | Yes | Nucleosome positioning score (nucleosome phasing (Termed Periocity in paper) for specific gene (from TSS to TTS)) Nucleosome spacing Nucleosome spacing | https://github.com/nucleosome-dynamics/NucDyn |
| NucHMM | A quantitative method modeling of nucleosome organization identifying functional nucleosome states, the nucleosome position is identified based on INPS. | Python/C++ | SE/PE MNase-seq ChIP-seq | No | Nucleosome positioning score Nucleosome spacing Nucleosome spacing | https://github.com/KunFang93/NucHMM |
| NucHunter | Predicting nucleosome positions with histone marks annotation from ChIP data. | Java | SE/PE ChIP-seq ChIP-seq ChIP-seq | No | Nucleosome positioning score (Fuzziness score) Nucleosome positioning score | https://eigen.molgen.mpg.de/nuchunter/ https://nucleoatac.readthedocs.io/en/latest/ |
| NucleoATAC | NucleoATAC is a python package for calling nucleosome positions and occupancy using ATAC-Seq data. | Python | PE ATAC-seq MNase-seq MNase-seq | No | None | https://sites.google.com/site/beckerjeremie/home/nucleofinder |
| NucleoFinder | A statistical approach for the detection of nucleosome positions. The authors claim it has fewer false positive detection than NPS and TemplateFilter. | R | SE MNase-seq MNase-Chip MNase-seq | No | Nucleosome spacing | https://sites.google.com/site/beckerjeremie/home/nucleofinder |
| nucleR | Using Non-parametric methods to detect nucleosome position. Its features include in situ visualization and exporting results to common genome browser formats. | R | SE/PE MNase-seq MNase-seq Tiling array data | No | Nucleosome positioning score | https://mmb.pcb.ub.es/nucleR/ |
| NucTools | NucTools accounts for the continuous distribution of nucleosome occupancy. | Perl/Matlab | SE/PE MNase-seq | Yes | Nucleosome positioning score | https://generegulation.org/nuctools/ |
peak calling tools and nucleosome characterization tools, the interested reader are referred to the review [55] and tools' original articles [24,27,44,45].

3. Computational methods in mapping 3D chromatin domains and interactions

**Identification of chromatin compartments.** Mammalian genomes are usually composed of two types of chromatin compartments, A and B compartments [17], corresponding to open and closed chromatin, respectively. Compartments were first derived and defined by the sign of the first principal component analysis (PCA) on transformed Hi-C matrices. Many tools, such as Juice-Box [56], HOMER [57] and Fan-C [58], implemented original compartment analysis in their suite. It was further extended to a finer sub-compartment on Hi-C data with a very high sequencing coverage such as Hi-C GM12878 data [59] by SNIPER developed by Xiong and Ma [60], based on a denoising autoencoder and a multilayer perceptron classifier. A recent new tool, Calder [61] was able to identify multi-scale sub-compartments at variable data resolutions, even in experiments with relatively low resolution.

**Prediction of TADs.** TAD is one of the key 3D-genome structures that control gene regulation [62]. So far, more than two dozen [59,63–66] programs or algorithms have been developed to predict TADs. Most of TAD prediction methods can be grouped into four major categories: one-dimension linear score method (e.g., TopDom [67]), two-dimension clustering method (e.g., MSTD [68]), feature method (e.g., PTADS [69]), and statistical method (e.g., HiCseg [70]). Recently, a systematic comparison of 22 computational methods for predicting TADs was carried out [66], providing a detailed description illustrating the pros and cons of each of the 22 methods. This extensive evaluation of TAD prediction tools is an excellent resource that safe-guides users in selecting a suitable tool or method for the TAD prediction, based on the desired experimental design and biological question. Although TADs are commonly conserved at a larger scale among different conditions, Sub-TADs within each large TAD are believed to be dynamically changed between the conditions [62]. Some of TAD prediction tools are specially designed to detect both the TAD and the sub-TAD, including GMAP [71], HiTAD [72], Arrowhead [73] and TADtree [74]. A more recent tool, OnTAD [75], implemented a two-level approach to generate a hierarchical TAD organization that best fits the input Hi-C interaction matrix. It is reported to performs better than other tools and reveals several interesting biological phenomena related to a TAD hierarchy. It is worth mentioning that a set of methods were particularly developed for some variant Hi-C techniques such as ChIA-PET [76], capture-Hi-C [77], HiChIP [78] and Split-Pool Recognition of Interactions by Tag Extension (SPRITE) [79] due to their sparse or asymmetric interaction matrices. GRINCH [80] is one such tool based on a non-negative matrix factorization and graph regularization, to identify TADs of genome organization from sparse chromosome interaction matrices.

**Prediction of differential TAD/TAD boundaries.** Besides the computational prediction of TADs, there is also a need for computational methods to investigate dynamical changes of TADs between the conditions. For example, researchers have recently observed differential responses at TADs, looping genes and expressed genes in endocrine-resistant breast cancer cells [81]. A new algorithmic approach DADO [82], has been proposed to identify such differentially active domains between the two conditions, providing complementary information to a general differential expression analysis. However, there are not many publicly available tools for the aforementioned task. In the future, more tools are needed to perform the differential analysis of chromatin domains under different conditions.

**Identification of chromatin interactions and loops.** In addition to mapping 3D chromatin domains, computational methods for identifying chromatin interactions or loops have also been developed quickly in the past few years. A probabilistic mixture model [83] was proposed to identify promoter-enhancer interactions from Hi-C data about five years ago. However, the sequencing depth required to obtain high-resolution interactions from Hi-C data is too big. A new tool, HIFI [84], based on a density estimation algorithm was thus designed to detect high resolution (at the restriction-fragment scale) chromatin interactions from Hi-C data. Alternately, a combination of low input “easy Hi-C” protocol for 3D genome mapping and a new analysis pipeline (HiCorr) [85] for Hi-C bias-correcton at high resolution, were able to detect enhancer-promoter loops at sub-TAD level. Besides, FithHiC/2 [86,87] implemented the statistical confidence estimation method to detect loops. And a novel computational and statistical method (HiSIF) [29] was proposed to identify genome-wide chromatin loops in Hi-C data with high resolution, by using a two-tier module (quality

### Table 1 (continued)

| Tools         | Description                                                                 | Programming Language | Input data Type | Differential analysis | Quantitative Nucleosome organization Information | Webpage                                      |
|---------------|-----------------------------------------------------------------------------|----------------------|-----------------|----------------------|-------------------------------------------------|----------------------------------------------|
| NUCwave       | A wavelet-based tool that is designed to evaluate the nucleosome-related experimental methods. | Python               | seq             | No                   | spacing (NRL)                                   | https://nucleosome.usal.es/nucwave/           |
| PING2         | A Probabilistic inference method to identify nucleosome positioning.         | R                    | PE/SE/MNase-seq  | No                   | None                                           | https://www.bioconductor.org/packages/release/bioc/html/PING.html |
| PuFFIN        | A parameter-free method to construct genome-wide nucleosome maps from paired-end sequencing data. | Python               | PE/MNase-seq    | No                   | Nucleosome positioning (Fuzziness)               | https://github.com/ucrbioinfo/PuFFIN         |
| TemplateFilter| Source code and executable files based pipeline for nucleosome positioning data processing. | Perl                 | SE/MNase-seq    | No                   | None                                           | https://compbio.cs.huji.ac.il/NucPosition/TemplateFiltering/Home.html |

**Tools Description**: NUCwave A wavelet-based tool that is designed to evaluate the nucleosome-related experimental methods. PING2 A Probabilistic inference method to identify nucleosome positioning. PuFFIN A parameter-free method to construct genome-wide nucleosome maps from paired-end sequencing data. TemplateFilter Source code and executable files based pipeline for nucleosome positioning data processing.
control and classification). HiSIF detected genes with enhanced loops showing worse survival in endocrine-treated breast cancer patients. FitzHiChIP [88] is a computational tool for calling chromatin loops from Hi-C/HiChIP/PLAC-seq data with a better performance. More recently, MUSTACHE [89] employed scale-space theory in computer vision to detect loops in Hi-C contact matrix. For readers who are interested in studying the algorithm and comparing the performance of the Hi-C related tools, recent review papers [66,90,91] are referred.

4. Computational approaches for integrating multi-omics data including Hi-C data

The availability of Hi-C data and other genome-wide omics data from 4D Nucleome [92], ENCODE [93] and other resources provide an excellent opportunity to develop computational approaches to integrating multi-omics data from the same or different sources. Some pioneering studies explored integrative computational methods in the context of 3D genome organization. Lan et al. [94] used a Mixture Poisson Regression Model and a power-law decay background to identify a highly specific set of interacting genomic regions from publicly available K562 Hi-C data, sequentially applied hierarchical clustering to ChIP-seq data of nine histone marks and an Apriori algorithm to ChIP-seq of 45 transcription factors, and collectively classified 12 different clusters of interacting loci categorized into two types of chromatin linkages. Whalen et al. [95] developed an ensemble machine learning pipeline, TargetFinder, for predicting enhancer-promoter interactions by integrating Hi-C interactions with epigenetic annotations from Segway and ChromHMM. Pancaldi et al. [96], inspired by a network topology metric used in social sciences, applied the metric to integrate epigenomic data and high-resolution promoter capture Hi-C and Hi-Cap data as well as ChIA-PET data to investigate promoter-centered chromatin interaction networks. Wang et al. [97] proposed a computational method called SPIN which integrated TSA-seq, DamID and Hi-C data from K562 cells in a unified framework based on a hidden Markov random field. They revealed a genome-wide intra-nuclear chromosome positioning and nuclear compartmentalization relative to multiple nuclear structures, such as nuclear lamina, speckles and nucleoli. Recently, a new method pTADs [69] was developed to identify TAD boundary and strength, by integrating both DNA sequence-based features (e.g., DNA shape and TF binding motif occurrence) and epigenetic profile information (e.g., CTCF, H3K36me3 and H3K20me1).

5. Advances in scHi-C computational analyses

Although tremendous progress in reconstructing the 3D chromatin structure based on population-averaged Hi-C data [98], single-cell Hi-C (scHi-C) protocols have newly been developed to identify 3D chromatin architecture at single-cell resolution [99–102], in which it has capability to delineate the 3D-regulated heterogeneity in population cells. For instance, the organization of zygote chromatin [103], the nuclear changes of stem cell differentiation [104], and single-allele chromatin interactions [105,106] have been fully examined by scHi-C technique. A crucial issue with scHi-C analysis is the inherent sparsity of the contact matrices and the technical noise mainly due to low amounts of starting material often resulting in variable capture efficiencies. Quality control of sequencing data is crucial to avoid technical artifacts. Despite of these challenges, new sets of computational methods have been developed for processing scHi-C data to reconstruct single-cell 3D chromatin structures [107–109], to impute the chromosome contact matrices [110–112], to identify TAD-like domains [113], to classify single cells [114], to identify chromatin loops [115], and to provide toolbox of scHi-C [116].

6. Conclusions and future perspectives

In summary, our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures, as well as for computational scientists who are interested in improving upon them. Despite such abundant resources, there is still an opportunity to develop more tools for integrating multi-source information and multi-level approaches in analyzing 3D transcriptional regulation. We also urgently need computational tools which could directly interpret the functionality of the nucleosome organization and 3D chromatin structure in various diseases. Another promising area worthy of attention is to incorporate molecular imaging data such as 3D super-resolution imaging (3D-SIM) [117] and genome architecture mapping (GAM) [118] to better derive the nucleosome and chromatin structures.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partly supported by National Institutes of Health (NIH R01GM114142 and U54CA217297 to V.X.J.) and NDSU Start-up Award to L.L. J.W. was supported by South-Eastern Norway Regional Health Authority [HSØ 2018107] and the Norwegian Research Council NOTUR project (nr4605k).

References

[1] Luger K et al. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997;389(6648):251–60.
[2] Splinter E et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006;20 (17):2349–54.
[3] Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell 2009;137 (7):1194–211.
[4] Weitzelub AS et al. YY1 Is a Structural Regulator of Enhancer–Promoter Loops. Cell 2017;171(7):1573–1588 e28.
[5] Miele A, Dekker J. Long-range chromosomal interactions and gene regulation. Mol Biosyst 2008;4(4):1046–57.
[6] Dekker J. Gene regulation in the third dimension. Science 2008;319 (5871):1793–4.
[7] Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol 2019;20(9):535–50.
[8] Baldisseri S et al. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. Nat Struct Mol Biol 2018;25 (9):894–901.
[9] Kilgore JA et al. Single-molecule and population probing of chromatin structure using DNA methyltransferases. Methods 2007;41(3):320–22.
[10] Brogaard K et al. A map of nucleosome positions in yeast at base-pair resolution. Nature 2012;486(7404):496–501.
[11] Schep AN et al. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res 2015;25(11):1757–70.
[12] Ishii H, Kadonaga JT, Ren B. MPE-seq, a new method for the genome-wide analysis of chromatin structure. Proc Natl Acad Sci U S A 2015;112(27): E3457–65.
[13] Dekker J et al. Capturing chromosome conformation. Science 2002;295 (5558):1306–11.
[14] Simonis M et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 2006;38(11):1348–54.
[15] Dostie J et al. Chromosome Conformation Capture Carbon Copy (3C): a massively parallel solution for mapping interactions between genomic elements. Genome Res 2006;16(10):1299–309.
[16] Hughes JR et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat Genet 2014;46 (2):205–12.
