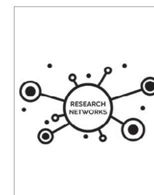
Computational and Structural Biotechnology Journal 21 (2023) 312–325

journal homepage: www.elsevier.com/locate/csbj



Review

Novel biological insights revealed from the investigation of multiscale genome architecture

Tianyi Ding, He Zhang *

Institute for Regenerative Medicine of Shanghai East Hospital, Frontier Science Research Center for Stem Cells, School of Life Science and Technology, Tongji University, Shanghai, PR China

ARTICLE INFO

Article history:
Received 12 August 2022
Received in revised form 6 December 2022
Accepted 6 December 2022
Available online 9 December 2022

Keywords:
Three-dimensional genome
Chromosome architecture
RNA-chromatin interaction
Phase separation

ABSTRACT

Gene expression and cell fate determination require precise and coordinated epigenetic regulation. The complex three-dimensional (3D) genome organization plays a critical role in transcription in myriad biological processes. A wide range of architectural features of the 3D genome, including chromatin loops, topologically associated domains (TADs), chromatin compartments, and phase separation, together regulate the chromatin state and transcriptional activity at multiple levels. With the help of 3D genome informatics, recent biochemistry and imaging approaches based on different strategies have revealed functional interactions among biomacromolecules, even at the single-cell level. Here, we review the occurrence, mechanistic basis, and functional implications of dynamic genome organization, and outline recent experimental and computational approaches for profiling multiscale genome architecture to provide robust tools for studying the 3D genome.

© 2022 The Authors. 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 . . . 313
2. Elementary structure of eukaryotic chromatin. . . . 313
3. Dynamic higher-order genome organization . . . 314
4. Topological insulators and chromatin mediators as a functional and architectural bridge. . . . 315
5. Phase separation helps create a stable nuclear body . . . 315
6. Detecting gene expression changes at the base level . . . 316
7. Detecting chromatin accessibility and biomacromolecule interactions at the nucleosome level . . . 316
8. Mapping genome-wide DNA-chromatin contacts . . . 317
9. RNA-chromatin interactions reveal transcriptional regulation in the genome. . . . 319
10. Sequencing and imaging at the single-cell level . . . 319
11. Data-driven 3D genome informatics. . . . 320
12. Conclusions and perspectives . . . 320
CRediT authorship contribution statement . . . 322
Declaration of Competing Interest . . . 322
Acknowledgments . . . 322
Author contributions. . . . 322
References . . . 322

* Corresponding author at: School of Life Science and Technology, Tongji University, Shanghai 200092, PR China.
E-mail address: zhanghe@tongji.edu.cn (H. Zhang).

https://doi.org/10.1016/j.csbj.2022.12.009
2001-0370/© 2022 The Authors. 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/).

1. Introduction

From a static perspective, the eukaryotic genome exhibits a multiscale architecture. Elements such as DNA, nucleosomes, histone modifications and interactions among biomacromolecules together form the elementary structure of the genome. Following the elementary linear structure described above, the dynamic folding genome is condensed into a three-dimensional organization including chromatin loops, TADs, A/B compartments, and chromatin territories [1,2]. Genes are selectively activated or silenced by reconstructing the chromatin architecture in the nucleus, thereby controlling the self-sustainment or directional differentiation of cells and determining tissue specificity and cell destiny, leading to the formation of complex tissues, organs and individual organisms [2,3]. Therefore, studying the higher-order structure of chromosomes and its regulatory mechanism is of great significance for understanding the expression patterns of genes and the mechanism of epigenetic regulation in the process of cell proliferation and differentiation.

Recent studies have revealed that the 3D organization of chromatin modulates biological processes such as DNA replication, transcription and meiosis, which are critical for cell differentiation and development [4,5]. Aberrant 3D genome organization is likely responsible for finger deformity, sex reversal, rheumatoid arthritis and diabetes [6–9]. Moreover, aberrant genome folding is relevant to common cancers, such as leukemias, prostate and gastric cancers [10–12], and orphan cancers including adenoid cystic carcinoma, Ewing sarcoma, multiple myeloma and ocular tumors [13–18]. Given that numerous variables and low-frequency chromatin interactions may shape genome structure and consequently regulate gene expression, it is crucial to elucidate genome architecture at different scales.

In recent decades, diverse approaches for exploring 3D genome folding have emerged and been continuously improved. Here, we first profile multiscale eukaryotic genome architecture at elementary and higher-order levels. We then review biochemistry and imaging technologies and data-driven 3D genome informatics approaches that are mainly applied to identify functional contacts and illustrate genome organization.

2. Elementary structure of eukaryotic chromatin

The elementary structure of chromatin serves as the cornerstone of DNA regulation. DNA, the basis for conveying hereditary information, contains both protein-coding sequences and noncoding sequences (Fig. 1). The methylation modification of DNA molecules can affect the binding of DNA-binding proteins, such as transcription factors, thereby regulating gene expression. Changes in nucleosome density at different locations in the genome can activate or repress gene expression [19]. Nucleosome positions are nonrandom in the genome and can determine the accessibility of DNA to regulatory proteins [20]. There are three major components controlling nucleosome positions: the DNA sequence, the binding of other protein factors and ATP-dependent remodeling complexes [21]. Recent studies have focused on the architecture and function of nucleosomes, with some researches analyzing the structures of the chromatin remodeling proteins Snf2 and ISWI in different states and elucidating the mechanism of chromatin remodeling by the RSC complex [22–25].

Additionally, methylation and acetylation on histones, which are important indicators of gene activation and inhibition, influence the recognition and aggregation of specific protein factors, thereby altering and regulating other structures or functions (Fig. 1). Recently, it has been shown that histone demethylase enzymes, as a subclass of molecular dioxygenases that depend on oxygen for their activity, can be affected by hypoxia to induce rapid changes in histone methylation and reprogram chromatin [26]. Notably, whether a given gene is expressed or not and its expres-

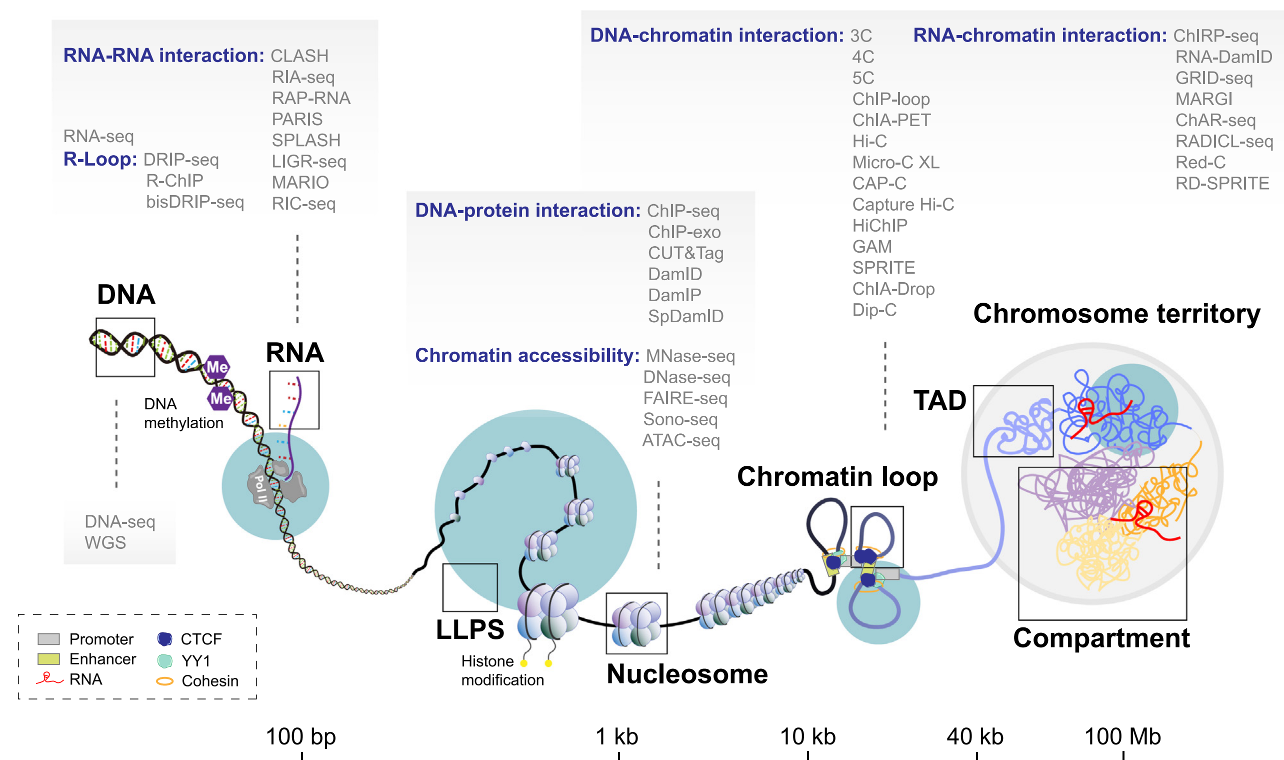


Fig. 1. Multiscale genome architecture and analysis methods related to different genomic layers. A schematic representation of multiscale genome architecture within the nucleus. From left to right: DNA, nucleosome, chromatin loop, TAD, compartment and chromosome territory. LLPS, shown in green circles, appears at different genomic scales. The methods are listed in gray boxes to show the respective detection locations.

sion level is determined by mutual antagonism through multiple histone modifications. On the one hand, the activation of genes is generally closely related to the modification of histones H3K4me3 and H3K27ac at the transcriptional start sites (TSSs) of genes. In detail, active enhancers can be identified based on H3K4me1 and H3K27ac enrichment, H3K4me3 exists at gene promoters, and H3K36me3 is associated with the transcriptional region of the gene body [26,27]. On the other hand, gene suppression can be mediated by different mechanisms, including modification with H3K9me3 and H3K27me3, which are both inhibitory signals but show different characteristics. H3K27me3 acts as a temporary signal that controls developmental regulators and is mainly found in promoters in gene-enriched regions; PCR2-accociated H3K27me3 initially accumulates in large intergenic domains and can then spread into genes only in the context of histone deacetylation and gene silencing [28]. In contrast, H3K9me3 is a signal of permanent heterochromatin, associated with chromosomal regions with tandem repeat structures. H3K9me3 is commonly found in regions lacking genes, such as satellite repeats, telomeres and regions near centromeres, and marks certain families of retrotransposons and zinc fingers (KRAB-ZFP) [3,29].

R-loops are another common transcription-associated structure recently identified at the elementary level of chromatin. R-loops are special triplex structures composed of a DNA–RNA hybrid and single-stranded DNA that are enriched at CpG islands (CGIs) and can regulate chromatin states [30]. The R-loop structure is one of the main causes of genome instability in cells, which has driven cells to evolve a set of mechanisms (e.g., RNase H1) for recognizing and degrading R-loops [31]. In recent years, a large number of R-loop structures have been discovered in a variety of organisms, and their functions have been associated with many crucial biological functions, including chromatin accessibility determination, DNA methylation modification, histone posttranslational modification, transcriptional regulation, and DNA damage repair [31–34]. Disordered R-loops often lead to the dysregulation of transcription and replication and a decline in genome stability, thus promoting the occurrence and development of many diseases. Research on R-loops relies heavily on detection techniques that can accurately locate and quantify R-loop structures. DNA:RNA hybrid immunoprecipitation and sequencing (DRIP-seq) [35], first reported in 2012, brought the study of R-loops into the age of genomics. In 2017, R-ChIP [36] first appeared; this new sequencing technology can accurately locate and quantify R-loops in vivo. A new method based on DRIP-seq and bisulfite conversion (for identifying non-template single-chain regions in R-loops), single-strand DNA ligation-based library construction of DNA:RNA hybrid immunoprecipitation and sequencing (bisDRIP-seq) [37,38], was developed to improve the resolution of DRIP-Seq and can produce a signal distribution similar to that of R-ChIP.

3. Dynamic higher-order genome organization

The nuclei of human cells harbor 22 sets of autosomes and 2 sex chromosomes containing genetic information in the form of genes. A model in which chromosomes are densely packed and folded into hierarchical domains at wide scales provides a profile of chromosomes in the eukaryotic nucleus (Fig. 1). The exploration of higher-order genome organization has revealed several avenues for future research.

Chromatin is organized at multiple scales that show variability and initially forms domains composed of individual loops, each of which occurs at relatively low frequency (1–3 %) in a cell population [39,40]. Because chromatin loops present high variability between distinct cell types, they affect gene expression under different conditions by enriching chromatin-chromatin interactions, especially by promoting interactions between distal individual genes or between genes and regulators [41,42]. Most loops are dynamic, not static, and undergo repeated formation and reformation throughout a typical 14–24 hr mammalian cell cycle [43]. Large local chromatin interaction domains referred to as TADs are a pervasive structural features and fundamental units of genome organization that facilitate further chromatin folding among populations of cells [44]. TADs are defined by boundary elements occupied by architectural chromatin proteins, and the size of TADs generally ranges from 200 kb to 1 Mb [45]. The boundaries of TADs are enriched in CCCTC-binding factor (CTCF) proteins, cohesin, housekeeping genes, transfer RNAs and short interspersed element (SINE) retrotransposons [44]. Depleting cohesin may abolish TADs on average at the population level without altering the prevalence of TAD-like structures in single cells [46]. TADs act as the basic regulatory unit for functions and are characterized by extensive internal interactions, such as those between enhancers and promoters. Individual TADs are relatively independent, and their internal interactions are much stronger than those with other TADs [44]. Nevertheless, TADs can further form higher-order domains, known as metaTADs, and the metaTAD hierarchy extends across genomic scales to the size range of entire chromosomes [47]. Within metaTADs, complex inter-TAD interactions can be understood as relatively simple tree-like hierarchical structures irrespective of the cell type [2,47]. TADs are smaller organizational units in the face of their further folding structure—chromatin compartments. As large-scale genomic structures, chromatin compartments can be divided into two types, A and B compartments, based on heterogeneity. Compartment A regions, which are relatively transcriptionally active have the tendency to spatially cluster around nuclear speckles and represent open chromatin, can be defined by DNase hypersensitivity. Compartment B regions, which are transcriptionally inactive, tend to be located close to the nucleolus and the nuclear lamina and associated with closed chromatin; these regions can be defined by repressive epigenetic marks of heterochromatin [48,49]. However, it is notable that some heterochromatin domains contain active transcription sites [50]. A/B compartment switches may indicate that topological remodeling is required for the transcriptional regulation of genes during physiological and biochemical processes, as observed in the epithelial-mesenchymal transition (EMT) spectrum [51]. More globally, chromosomes reside in separate territories that preferentially show transcriptional activity in a nonrandom manner [52]. According to the genome-wide view of the relationship between transcriptional activity and interactions with nuclear architecture, nuclear lamina-associated loci and nuclear speckle-associated loci are correlated with lower and higher transcriptional activities, respectively [53]. Furthermore, highly expressed genes seem to be prepositioned to active nuclear territories, and proximity to nuclear bodies may contribute extensively to regulating the chromatin states of DNA loci [54].

Major features of the genome are evolutionarily conserved, but the genome is highly plastic in its spatial arrangement; for instance, the position of domains within compartments is variable, and chromosomes are preferentially located, although they can be found anywhere in the nucleus [45,55]. Structural heterogeneity in genome organization is stochastic and intrinsic to the DNA polymer and may be a functional and regulated feature of a cell; chromatin topology may also affect cell function [1]. Both the structure and activity of individual genes are dynamic and stochastic, showing cyclic bursts of activity and periods of inactivity. Genome architectures are highly variable, in line with the observed stochasticity of gene expression. Genome organization events have been linked to functional changes, such as changes related to immune cell fate and function [56], and aberrant organization may also be a sign of diseases such as uveal melanoma [18].

4. Topological insulators and chromatin mediators as a functional and architectural bridge

Topological insulators and chromatin mediators are critical for 3D genome organization. Active transcription generally occurs in short bursts at irregular intervals [57]. This pattern of punctuated gene transcription by RNA polymerase pausing is common in all organisms and is particularly widespread in mammals [58]. Chromatin accessibility has been shown to be a determinant of gene expression, regulating the intermittent access of promoters to transcription factors [59]. Most prominently, the regulation of transcription occurs via the dynamic formation of promoter-enhancer loops. The interaction of chromatin proteins with their substrates is dynamic [60], even among architectural proteins that form stable loops, such as CTCF and cohesion[43]. CTCF functions in an orientation-dependent manner to bind to CTCF-binding sites in mammalian genomes[61]. CTCF mediates genome-wide long-range chromatin interactions together with associated cohesin proteins. Generally, the boundaries of TADs are enriched with insulators and cohesin proteins, which play a key role in forming TADs and maintaining their stability[44]. CTCF binding sites arranged in tandem in the genome can balance the spatial contact of the genome and the topological selection of enhancers and promoters[62]. The abnormal absence of CTCF-guided chromatin loops is a crucial factor in the loss of IGF2 imprinting and tumorigenesis [63]. CTCF plays a key role in the establishment of the 3D chromatin structure during human embryogenesis [64]. A recent study proposed a static model of TADs in which TADs exist in a fully unlooped or looped state, stably bridged by one cohesin molecule [65]. It was also suggested that CTCF-guided insulation may be regulated by individual extrusion-blocking CTCF boundaries to a greater extent than the rare fully looped state [65]. Therefore, within a TAD, frequent interactions mediated by cohesin may be more important for regulatory contacts than rare CTCF-CTCF loops. In addition, Yin Yang 1 (YY1), another zinc finger protein except CTCF, has been reported to facilitate enhancer-promoter interactions in a manner analogous to CTCF [66]. Constitutive CTCF-mediated interactions may form a preexisting topological framework and subsequently demarcate the locations of YY1-mediated interactions [67]. Long noncoding RNAs (lncRNAs) can control the structure and gene expression of chromatin and directly interact with histones and DNA-modifying enzymes through covalent modification while acting as cofactors of chromatin remodeling factors through noncovalent interactions (ATP-dependent) [68]. Taken together, the available evidence indicates that CTCF functions in an orientation-dependent manner to organize multiscale genome folding and is the basis of most loop formation events, which are also enriched at TAD boundaries. Changing the status of CTCFs can perturb domain boundaries and disrupt the expression of nearby genes. YY1 is a transcription factor and architectural protein that plays a key role in gene expression regulation, cell fate change and disease occurrence through the formation of chromatin loops. However, the mechanism of YY1-mediated chromatin loop formation has not been elucidated and needs to be further studied.

In the loop extrusion model, CTCF and cohesin are essential for spatial and temporal chromosome organization, although many recent papers have suggested that they may not be necessary in TAD formation or gene promoter and enhancer interaction. In the epiblast and in neuronal tissues, CTCF-mediated loops are not required for *Sox2* promoter and distal enhancer contact or for *Sox2* expression, but high-affinity enhancer-promoter interactions themselves can bypass CTCF/cohesin-mediated insulation and function in maintaining phenotypic robustness [69]. By generating contact matrices at a 20 bp resolution with an extension of the MNase-based 3C approach, also known as Tiled-Micro-Capture-C (Tiled-MCC), it has been shown that within *cis*-regulatory elements with distinct internal nanoscale structures, local insulation is dependent on CTCF but independent of cohesion [70]. Specifically, the depletion of cohesin subtly reduces enhancer–promoter contacts and leads to a small decrease in gene expression, but CTCF depletion can induce rewiring of interactions and ectopic gene activation in some cases, instead of generally reducing interactions between certain enhancers and promoters [70]. Furthermore, a recent study showed that a TAD physically separates *Zfp42* and *Fat1*, which have distinct local enhancers that drive their independent expression in ESCs (embryonic stem cells), but this separation is driven by chromatin activity independent of CTCF/cohesion [71].

5. Phase separation helps create a stable nuclear body

Liquid-liquid phase separation (LLPS) is a phenomenon in which different macromolecules in a cell collide and fuse with each other, driven by physical force, to form droplets such that some components are enclosed within the droplet, while some are restricted to the outside [72]. In early research, P particles were found to be a kind of protein with characteristics unlike solids but similar to droplets that collide and merge with each other and disperse into small droplets after vigorous shaking and then quickly fuse to form large droplets [73]. Many membraneless organelles in cells, such as nucleoli, Cajal bodies, stress granules, miRISC (microRNA-induced silencing complex) and the synaptic cytoskeleton, represent specific protein or RNA phase transitions [74–76]. Whether a solution of macromolecules undergoes phase separation depends on two main factors. First, the concentrations and identities of the macromolecules in the solution play a key role, and second, environmental conditions are also important, including temperature, salt types and concentrations, cosolutes, pH, and the volume excluded by other macromolecules [77,78].

Due to the ubiquitous occurrence of phase separation in cells, this phenomenon is associated with and performs functions in various biological processes, including gene transcription regulation, epigenetics, posttranslational modification of proteins, classical signaling pathways, neurobiology, tumorigenesis, lipid transport and immunity. Phase separation disorders may cause degenerative neurological or muscular diseases [79]. Genetic abnormalities associated with cancer generally appear in proteins with low-complexity regions (LCRs), which are also called intrinsically disordered regions (IDRs) [80]. The relationship between cancers and IDRs has long been obscure; however, a research team recently revealed a key role of IDRs in tumorigenesis, in which IDRs contained within NUP98–HOXA9 recurrently detected in leukemias were shown to be essential for establishing LLPS puncta of chimeras and for inducing leukemic transformation [81].

Phase separation provides a new perspective for understanding genome assembly, creating stability out of organizational variability (Fig. 1). HP1 is a well-known key silencing protein, and its homolog Swi6 can significantly increase the accessibility and dynamic changes of hidden histone residues in nucleosomes [82]. It is unknown how HP1 compacts chromatin into phase-separated condensates. A recent study showed that Swi6 contributes to nucleosome remodeling and can increase the chance of multivalent interactions between nucleosomes, thereby promoting phase separation [83]. To reveal whether and how physical forces can directly restructure chromatin, the CasDrop system, a novel CRISPR-Cas9-based optogenetic technology developed in 2018, can quantify and locate the phase separation of multiple proteins [84]. With CasDrop, it has been shown that the LLPS phenomenon in the nucleus can sense and reshape the chromatin structure [84]. Phase separation leads to the segregation of molecules or DNA loci and can result in the formation of a stable nuclear body resulting from a complex combination of variable associa-

tions [39]. Reconstituted chromatin undergoes histone tail-driven LLPS in a physiological salt solution and produces dense, dynamic droplets when microinjected into cell nuclei, revealing a framework based on the intrinsic phase separation of the chromatin polymer for better understanding eukaryotic genome organization and regulation [85]. Additionally, RNA-driven phase separation can regulate RNA binding protein (RBP) activity and reveal an essential role of this process in genome maintenance. The repetitive sequence architectures of *NORAD* and *NEAT1* lncRNAs suggest that phase separation may be a widely used mechanism of lncRNA-mediated regulation [86,87].

The process of phase separation appears to promote the formation of heterochromatin domains and super enhancer (SE) clusters [88–90]. An SE is a cluster of enhancers formed by serially arranged enhancers in series, the emergence of which generally accompanies most key genes that determine cell characteristics and functions [91,92]. The transcriptional coactivators BRD4 and MED1 can undergo phase separation at an SE to form droplets and achieve compartmentalization of the transcription process [88]. Therefore, SEs regulate gene expression through phase separation, which provides a new perspective on the regulation of the expression of key genes during cell fate determination and disease development. The phase separation of proteins can segregate loci by excluding DNA from dense protein droplets and tethering them together to form a single droplet [84], suggesting the probability of phase-separating proteins stabilizing the dynamic organization. LLPS of NUP98-HOXA9 promotes chromatin occupancy of chimeric transcription factors and is required for the formation of a broad 'super enhancer'-like binding pattern typically observed in leukemogenic gene sequences, which potentiates transcriptional activation [81]. Phase-separated NUP98-HOXA9 induces the enrichment of CTCF-independent chromatin loops in proto-oncogenes, which are induced by LLPS transcription factor proteins and are completely different from CTCF-dependent chromatin loops.

In addition to LLPS, there is another type of phase separation known as polymer–polymer phase separation (PPPS), or bridging-induced phase separation (BIPS), which is a type of biomolecular condensation [93–96]. PPPS can generate an ordered collapsed globule, which is induced by chromatin-associated proteins that bridge nucleosomes residing in close spatial proximity to each other [94]. PPPS can clarify chromatin architecture variability across single cells. Strings and Binders (SBS) polymer physics model predictions for single-molecule structures provide evidence that chromosomal architecture is controlled by a thermodynamic PPPS mechanism [93]. The thermodynamic degeneracy of single-molecule conformations leads to extensive architectural and temporal variability of TAD-like interaction patterns, and PPPS appears to be a robust mechanism of chromatin organization that is reversible [95]. In addition, bare DNA is bridged by a single structural maintenance of chromosome (SMC) protein complex to increase the local DNA concentration and subsequently induce the binding of more SMC complexes in this region, leading to the formation of a large DNA/SMC-holocomplex droplet in an ATP-independent manner [96]. With PRISMR, a previously developed machine-learning approach, contact patterns from Hi-C data can be explained based on the molecular mechanisms envisaged by the SBS polymer model [95]. Although some advances have been achieved recently, the elucidation of the intricacies of the associations among phase separation, chromatin changes, and genome function awaits further research.

6. Detecting gene expression changes at the base level

At the base level, DNA-seq is a key library construction technology that provides a foundation for other next-generation sequencing (NGS) library construction technologies. The similar technology of RNA-seq analyzes a collection of RNAs that can be transcribed in a specific cell in a certain functional state, thereby providing an overview of gene functions in a specific context. Many advanced methods have been based on these technologies in recent years. In mammalian cells, mRNAs are responsible for protein coding, and noncoding RNAs (ncRNAs) that do not encode proteins play critical roles in essential biological activities, including RNA splicing, by interacting with other RNAs [97].

Various types of RNA-RNA interaction detection technologies are continuously being developed (Fig. 1). High-throughput methods based on sequencing, such as the crosslinking, ligation and sequencing of hybrids (CLASH) [98], RNA interactome analysis and sequencing (RIA-seq) [99] and RNA antisense purification followed by RNA sequencing (RAP-RNA) [100] methods, mainly involve the initial crosslinking of interacting RNAs, followed by the enrichment of protein-specific RNA interactions by using antibodies. This kind of method represents the earliest application of high-throughput technology for large-scale RNA-RNA identification, resulting in the realization one-to-many interaction identification and providing specific interaction site information. With the development of sequencing methods, novel approaches have been developed at the whole-transcriptome level, including the psoralen analysis of RNA interaction and structures (PARIS) [101], sequencing of psoralen crosslinked, ligated, and selected hybrids (SPLASH) [102], ligation of interacting RNA followed by high-throughput sequencing (LIGR-seq) [103] and mapping of the RNA interactome in vivo (MARIO) [104] methods. Recently, a new method, RNA *in situ* conformation sequencing (RIC-seq) [105], was shown to capture the advanced structures of RNA and the targets of various types of noncoding RNAs and to depict RNA-RNA spatial interactions *in situ*. These technologies have established RNA-RNA interaction profiles in different species and multiple cell lines, making it possible to construct and analyze RNA-RNA interaction networks.

7. Detecting chromatin accessibility and biomacromolecule interactions at the nucleosome level

Histones and 147 bp DNA sequences form nucleosomes, which act as the core structure of chromatin. The composition of nucleosomes and posttranscriptional modifications reflect different functional states and regulate chromatin accessibility [2]. Open euchromatin tends to bind transcription factors (TFs), RNA polymerase, or other proteins, allowing active gene transcription. In contrast, DNA tightly wrapped in nucleosomes or higher-order heterochromatin loses its binding capability, leading to gene silencing. At this level, interactions emerge among genes, *cis*-acting elements such as promoters and transcription factor-binding sites (TFBSs), and distal regulatory elements such as enhancers, repressors, and insulators.

Therefore, studying chromatin accessibility and proteins in open regions of chromatin represents major work to be conducted at the nucleosome level. There are several main techniques for assessing genome-wide chromatin accessibility (Fig. 1). In MNase-seq, a micrococcal nuclease (MNase) derived from *Staphylococcus aureus* is used to digest exposed genomic regions, chromatin-bound sequences are protected from MNase digestion, and then DNA bound to nucleosomes is recovered and sequenced [106]. MNase-seq can mainly be used to identify nucleosome localization and occupancy and the regulatory factors that bind to nucleosomes [106]. Similarly, DNase-seq employs deoxyribonuclease I (DNase I), a nonspecific endonuclease, to digest exposed regions of the genome [107]. DNase-seq identifies DNase I digestion sites with a base-pair resolution, providing a method that is

the inverse of MNase-seq [108]. DNase-seq is a mature, practical experimental method that can clearly explain the deviation of DNase I splicing and can allow the reverse deduction of unsplit genome regions, known as the DNase footprinting method, with the aim of identifying transcription factors and nucleosome binding sites. In such experiments, it is important to set a control because the shearing bias of DNase I may lead to false conclusions. Formaldehyde-assisted isolation of regulatory elements sequencing (FAIRE-seq) [109] and Sono-seq [110] are often used to study open chromatin regions. The two methods are slightly different experimentally but operate on the same principle. FAIRE-seq is widely used to identify activated regulatory elements in different human cell lines and to compare differences in chromosomal accessibility between normal cells and cells in disease states. The independence of FAIRE-seq on proteins such as antibodies and enzymes avoid the sequence preference of MNase and DNase I DNA cutting. However, successful FAIRE-seq shows high dependence on the formaldehyde fixation efficiency. Compared with other chromatin accessibility analysis methods, the sequencing signal–noise ratio of FAIRE-seq is relatively low, and excessive background signals may interrupt the data analysis. Tn5 transposase is widely applied in sequencing library generation because it can ligate synthetic oligonucleotides at both ends of fragmented double-stranded DNA and easily binds to open chromatin [111]. Therefore, the Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) [112] utilizes a hyperactive Tn5 enzyme preloaded with DNA adapters for sequencing to fragment and tag the genome, which can allow chromatin accessibility to be efficiently assayed. ATAC-seq requires a small number of cells and is characterized by a high signal–noise ratio. The whole process is simple and fast. It can detect the open state of chromatin in the whole genome. By adapting the scheme of flow cytometry or microfluidic technology, single-cell analysis can be performed.

To further study open regions of the genome with high accessibility, especially the interactions between nucleic acids and proteins, chromatin immunoprecipitation sequencing (ChIP-seq) and derivatives thereof are utilized in most situations (Fig. 1). The combination of ChIP-seq, DNase-seq, and FAIRE-seq can reveal TFBSs, nucleosome distribution locations, open regions of chromatin, and the relationships among the three. Based on the principle of DNA-protein interaction, ChIP-seq isolates and enriches DNA fragments interacting with proteins of interest to study TFBSs. However, in ChIP-seq experiments, the use of nonspecific antibodies can dilute the DNA-protein complex library of interest, and the target protein must be known and capable of inducing antibodies. Another method based on ChIP-seq, ChIP-exo, is a base-pair-resolution method for evaluating the protein binding distribution with a relatively high signal-to-noise ratio, to map the bindings of a protein of interest (POI) in the genome [113]. ChIP-exo can better identify low-distribution binding sites and allow in-depth analysis of the relationship between DNA sequences and the distribution of transcription factors. The Cleavage Under Targets and Tagmentation (CUT&Tag) [114] approach is a new method for studying protein-DNA interactions that can analyze very small numbers cells (60 cells) or even single cells. The antibody-guided ChiTag enzyme used in CUT&Tag is a hyperactive Tn5 transposase-Protein A (pA-Tn5) fusion protein loaded with sequencing adapters that only fragments target DNA in regions containing target histone modification markers, transcription factors or chromatin regulatory proteins that bind to chromatin [114]. CUT&Tag and ChIP-seq differ in two main aspects. First, during sample preparation, CUT&Tag involves a cell permeabilization treatment. Second, to obtain the target DNA fragment, instead of adding a primary antibody and Protein A magnetic beads, which may be included in ChIP-seq, the CUT&Tag procedure allows the use of a secondary antibody to bind to a primary antibody that has already bound the protein of interest after entering the cell; thereafter, the pA-Tn5 transposome binds to the secondary antibody, and Mg^{2+} is finally added to activate the transposome and fragment and extract DNA. DNA adenine methyltransferase identification (DamID) [115–117] can identify whole-genome protein-DNA interactions without antibodies, fixation or pull-down as long as cells are introduced into the Dam system in advance, which means that DamID is limited to some cell models. ChIP can only provide information on the interaction between chromatin and DNA at a particular point in time (when formaldehyde cross-linking is performed), while DamID is able to reveal the dynamic changes in chromatin interactions. DamID has a relatively low resolution because it only recognizes the GATC motif. Another method based on DamID, Dam immunoprecipitation (DamIP) [118,119] facilitates DNA adenine methyltransferase (DamK9A) targeting of the ATC motif, which is more common than the GATC motif within genes. Additionally, Split DamID (SpDamID) [120] can label DNA in living cells only if two labeled proteins are located close enough to interact with each other on the same DNA strand.

To allow gene expression alterations and cell reprogramming, chromatin remains in a dynamic state that is constantly being reshaped. The exploration of chromatin structure and nucleosome localization reveals epigenetic mechanisms related to specific cellular processes and disease states. Corresponding transcription factors or other regulatory proteins binding to open regions of a chromosome directly affect intracellular gene replication and transcription and play an important role in transcriptional regulation. The accurate identification of these specific open DNA regions in the genome is critical for discovering regulatory elements.

8. Mapping genome-wide DNA-chromatin contacts

DNA-chromatin interactions can be detected and mapped by different strategies, from one-to-one to genome-wide scales. The chromatin conformational capture (3C) approach focuses on interactions between two known loci (especially promoters and enhancer; “one versus one”), by using site-specific PCR [121]. At present, 3C has become the basis for a range of related technologies for larger-scale, higher-throughput, or higher-specificity analyses. By self-circularizing cut DNA fragments, the circular chromosome conformational capture (4C) [122] strategy recognizes an unknown region of DNA that interacts with the target site (“one versus all”). Primers for a region of interest, such as a certain gene promoter, are designed to amplify all possible ligation partners of the target locus, referred to as the ‘viewpoint’ [123]. 4C is the preferred strategy for assessing DNA contact profiles of individual genomic loci with highly reproducible data. Chromosome conformation capture carbon copy (5C) [124] allows the simultaneous determination of interactions between multiple sequences and is a high-throughput version of 3C. 5C provides a sequencing depth of approximately 60 million reads per library to achieve a resolution of 15–20 kb for a 1-Mb region and is an ideal strategy when the aim is to understand all interactions in the target area, for example, when a detailed interaction matrix must be drawn for a specific chromosome (“many versus many”) [125].

The ChIP-loop method, combining features of 3C and ChIP-seq, relies on the interaction between known promoters and enhancers to detect the interaction between two target loci mediated by a protein of interest and reveal the functions of specific transcription factors. ChIP-loop analysis allows characterization with good specificity but without sufficient information, so the results should be cross-referenced with ChIP data [126]. ChIP-loop is suitable for identifying suspected interactions when used to confirm whether two known DNA regions interact with a protein of interest. However, ChIP-loop presents the problem of unnatural loop formation

during the process of DNA concentration before ligation occurs. To avoid this, precipitation should be performed after ligation. Chromatin Interaction Analysis with Paired-End Tag Sequencing (ChIA-PET) [127] is a high-throughput version of ChIP-loop that integrates ChIP, chromatin hinges, pair-end tags, and high-throughput sequencing to study genome-wide remote chromatin interactions; this method can detect long-range chromatin interactions through a target protein in the whole genome and allows the construction of a chromatin interaction network mediated by known transcription factors. ChIA-PET is best suited for discovering the interactions of target proteins with unknown DNA, such as revealing the binding sites of transcription factors. ChIA-PET requires large amounts of cells but generates small fragments of reads, and few reads of genes of interest are produced.

The emergence of high-throughput chromosome conformation capture (Hi-C) marked the start of a new era of 3D genome research [48] (Fig. 2). If extensive genome-wide coverage is needed, Hi-C appears to be a great choice without concern about resolution. Hi-C library amplification must generate enough products for analysis while avoiding false PCR products. Additionally, the read length should be appropriate, and the proper sequencing unit size should be selected [128]. In general, Hi-C can resolve all chromatin conformations, but deep sequencing requires the full analysis of chromatin features; therefore, the specificity of Hi-C is not good, and the signal-to-noise ratio is relatively low. As one of the most commonly used 3D genomic technologies, the advent of Hi-C will certainly give rise to various new advanced techniques in the next decade. *In situ* Hi-C, developed in 2014, can reveal the 3D structure of the human genome at a superresolution of 1 kb because it uses intact nuclei for later reactions, which reduces the background noise caused by random ligation, and the use of a four-cutter enzyme may also improve the resolution. The experimental conditions were further optimized in Hi-C 2.0, in which heat stimulation is used instead of SDS to inactivate the enzyme, and end-labeled biotin is removed from unligated fragments to decrease the ratio of dangling ends [129]. In Micro-C XL, MNase is employed to further improve the resolution of interaction maps, and double cross-linkers are used to obtain more interactions at the source [130] (Fig. 2). During the updating of these technologies, two key parameters, cross-linking and chromatin fragmentation, determine the ability to generate Hi-C maps. On this basis, Hi-C 3.0, in which formaldehyde with disuccinimidyl glutarate (DSG) is used for cross-linking and *Dde*I with *Dpn*II is used for DNA fragmentation, can be applied to identify both compartments and chromatin loops [131]. In addition, chemical-crosslinking assisted proximity capture (CAP-C) relies on multifunctional chemical

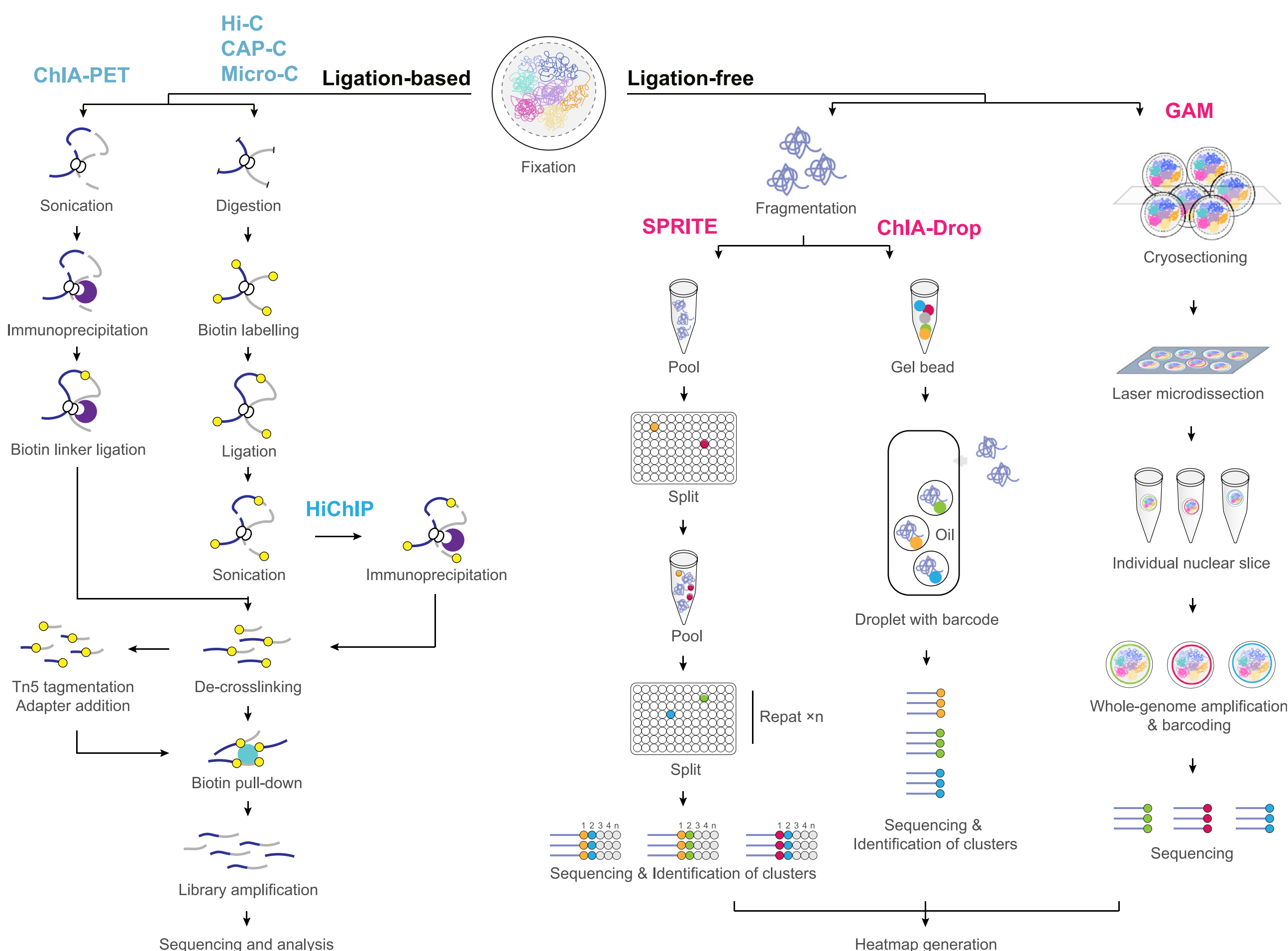


Fig. 2. Main ligation-based and ligation-free biochemistry methods for genome-wide chromatin interaction mapping. The main ligation-based approaches are chromatin interaction analysis with paired-end tag sequencing (ChIA-PET); high-throughput chromosome conformation capture (Hi-C); chemical-crosslinking assisted proximity capture (CAP-C); Micro-C; and *in situ* Hi-C followed by chromatin immunoprecipitation (HiChIP). They share some steps and have individual specific steps. Hi-C, CAP-C and Micro-C are different crosslinkers and enzymes used for digestion. The most common ligation-free methods are split-pool recognition of interactions by tag extension (SPRITE), chromatin-interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop) and genome architecture mapping (GAM). They are based on different strategies of barcoding and identification.

crosslinkers that capture proximal DNA loci through direct crosslinking upon ultraviolet (UV) irradiation [132] (Fig. 2). Therefore, CAP-C enables the mapping of chromatin contact at a sub-kilobase resolution without formaldehyde prefixation and shows low background noise.

Based on a Hi-C library combined with 3C, oligonucleotide capture technology (OCT) and high-throughput sequencing, in the Capture Hi-C (CHi-C) method, probes are designed for target region (below 6 Mb) for capture and sequencing; this method is suitable for analyzing the genome-wide interactions corresponding to multiple target regions [133]. In contrast to 4C, one CHi-C library can capture the interaction information of multiple sites. Compared with Hi-C, the CHi-C target sequence shows greater enrichment and a higher resolution. However, 4C can verify the results of Hi-C and CHi-C. The difference in the amount of sequencing data among the three techniques is that 4C generally generates 10 G per sample, but CHi-C and Hi-C depend separately on the capture area and the resolution requirements. In addition, *in situ* Hi-C followed by chromatin immunoprecipitation (HiChIP) is a method for analyzing chromatin conformation using the *in situ* Hi-C principle and the transposase-mediated construction of a library [134] (Fig. 2). It combines Hi-C and ChIA-PET to obtain higher-resolution chromatin three-dimensional structure information with a smaller amount of data but can only generate high-dimensional chromatin structures when bound by proteins of interest. Compared with ChIA-PET and Hi-C, HiChIP requires a small number of cells and produces many reads of the gene of interest with a high signal-to-noise ratio and good specificity. It is fast, efficient, and simple but only generates chromatin conformations bound by the protein of interest.

The above approaches that rely on pairwise proximity ligation generally cannot reveal the detailed nature of chromatin interactions and present limitations such as formaldehyde crosslinking bias, insufficient ligation efficiency and the omission of multipartite interactions in sequencing. Ligation-free methods can address these technical challenges and provide more information for capturing in vivo chromatin interactions (Fig. 2). The genome-wide method of genome architecture mapping (GAM) was the first genome-wide approach developed for capturing chromatin contacts between any number of genomic loci without ligation. GAM involves ultrathin cryosectioning with laser microdissection in fixed cells embedded in sucrose and achieves the isolation of single nuclear profiles [135]. In a recent study, GAM was modified by the addition of immunolabeling, and the new approach was referred to as immunoGAM, which requires low cell numbers (approximately 1,000 cells) within a complex tissue and avoids tissue dissociation [136]. The split-pool recognition of interactions by tag extension (SPRITE) method is not dependent on proximity ligation and identifies genome-wide higher-order interactions occurring simultaneously between multiple DNA sites [137]. SPRITE enables the mapping of long-range interactions than can be observed by Hi-C. Chromatin-interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop) with ChIA-DropBox is the predominant method for revealing promoter-centered multivalent interactions with single-molecule precision [138].

A chromatin contact matrix can be used to evaluate the 3D structure of chromatin, thereby revealing the physical interactions between distant genomic regions. The advent of 3C technology and its derivatives have made this type of mapping possible. Each method presents distinct advantages for specific applications. Based on the diversity of the available methodologies, it is challenging to choose a method according to a specific purpose. A recent study reported one way to benchmark these various techniques, including Hi-C, GAM and SPRITE, in a simplified and controlled computational framework against known higher-order architectures of polymer models of DNA loci [139].

9. RNA-chromatin interactions reveal transcriptional regulation in the genome

Recently, as the study of DNA-chromatin interactions has progressed, researchers have come to pay attention to RNA-chromatin interactions. These chromatin-associated RNAs (caRNAs) function as epigenomic modifiers and play important roles in gene regulation and genome organization in the context of both human biological activities and plant development [140,141]. Beyond the techniques for studying proteins on DNA, chromatin isolation by RNA purification sequencing (ChIRP-seq) specifically focuses on RNA and detects DNA and proteins bound to RNA across the genome [142]. ChIRP-seq can simultaneously analyze the interactions of lncRNAs, binding proteins and DNA, which is an effective way to obtain lncRNAs regulating target genes. A biotin probe inversely complementary to a target RNA sequence is first designed and bound to streptavidin, and this probe then specifically binds to the target RNA to capture the chromosome fragments regulated by RBPs involved in transcriptional regulation. RNA-DNA adenine methylase identification (RNA-DamID) [143] detects lncRNA-genome interactions in vivo in a cell type-specific manner using two systems, Gal4-UAS and MCP-MS2. It shows higher sensitivity and accuracy than ChART, RAP, and ChIRP and requires fewer cells. Genome-wide methods such as the global RNA interactions with DNA by deep sequencing (GRID-seq) [144], mapping RNA-genome interactions (MARGI) [145], chromatin-associated RNA sequencing (ChAR-seq) [146], RNA and DNA interacting complexes ligated and sequenced (RADICL-seq) [147] and Red-C [148] techniques focus on capturing the *in situ* global RNA-chromatin interaction landscape, including coding mRNAs and ncRNAs that bind to active promoters and enhancers, to identify all potential regulatory RNAs and assign their action sites in an unbiased manner. These technologies based on proximity ligation mostly use a biotinylated bivalent linker to facilitate the ligation of RNA to DNA *in situ*. In addition, RNA & DNA SPRITE (RD-SPRITE), developed by improving the efficiency of the RNA-tagging steps of the SPRITE approach, comprehensively maps higher-order RNA-chromatin structures with split-and-pool barcoding [149].

Compared with the Hi-C method, RNA-chromatin interaction methods can generate three-dimensional interaction maps of RNA-chromatin and reveal cell-type-specific gene regulation, in which RNAs might also play a role in coordinating functional DNA elements in the context of chromatin structures, to study the regulatory roles of RNA. Recent technologies still cannot provide any information on the specific proteins involved and present limitations in detecting RNAs of low abundance.

10. Sequencing and imaging at the single-cell level

The techniques mentioned above generally illustrate the folding situation at some specific moment in the dynamic process, basically implying the average status of millions of cells. Despite the ubiquitous presence of all the main structural characteristics of genomes, there is considerable variability and heterogeneity in genome organization at the single-cell level. If a certain region presents a dynamic changing tendency or cell differences occur within the population, data derived from diverse methods may be inconsistent. Gene expression occurs in a stochastic fashion at the single-cell level, and stable phenotypes at a population level are derived from variable single-cell gene expression patterns [150]. Within individual cells in a population, genomes can assume many distinct, albeit related, spatial conformations mediated by rare, short-lived chromatin-chromatin interactions rather than by persistent and pervasive associations [151,152]. Individual chromatin

interactions identified in bulk biochemical assays, such as Hi-C and ChIA-PET, typically occur in a relatively small fraction of cells in the population at any given time. To address this problem, several single-cell biochemistry and imaging approaches for mapping the 3D genome are under development. In a single-cell chromatin conformation capture method known as diploid chromatin conformation capture (Dip-C), 3D information of chromatin structure is encoded by the proximity ligation of chromatin fragments, and the ligation product is then amplified by multiplex end-tagging amplification (META)and subsequently sequenced [90].

In addition to biochemistry approaches, imaging approaches such as fluorescence *in situ* hybridization (FISH), which is widely used to identify the locations of particular genomic regions and to calculate the distances between them, can be combined with superresolution imaging to investigate chromatin organization at an unprecedented resolution. In 2016, following modifications of the multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) approach, multiplexed DNA was successfully used to evaluate chromatin conformation using a dual-oligonucleotide version of an array-derived oligonucleotide (oligo) probe (Oligopaint), allowing the mapping of spatial genomics and tracing of TAD positions [49]. Additionally, in 2018, a high-throughput Oligopaint labeling and imaging method was applied to observe the dynamic genome structure inside the nuclei of different single mammalian cells [46]. In 2019, a method that could trace genomic regions in single cells with a nanoscale precision and a two-kilobase resolution emerged and was named the optical reconstruction of chromatin architecture (ORCA) [153] method. In 2020, MERFISH was upgraded to a high-throughput genome-wide imaging version to characterize chromosome architectures and *trans*-chromosomal interactions in single cells [53]. In addition, based on sequential fluorescence *in situ* hybridization (seqFISH+) [154], a DNA seqFISH+ method was newly developed in 2021; this new method uses a combination of multimolecule detection and multiple axles, and along with multiplexed immunofluorescence and RNA seqFISH, which is a robust spatial multimodal approach for illustrating chromosome architectures, nuclear bodies, chromatin states and gene expression within the same single cell [54]. Using DNA seqFISH+, thousands of DNA loci along with RNAs and epigenetic markers can be revealed simultaneously in single mouse ESCs and even in the mouse brain [54,155]. In addition to multiplexed DNA FISH approaches, other imaging methods, such as *in situ* genome sequencing (IGS), can simultaneously sequence and profile genomes and thereby directly connect sequences and structures within intact biological samples [156]. Single-cell genomics and imaging advances have contributed to the generation of static snapshots of higher-order genome structures in single cells [157], but seldom do methods enable the dynamics of chromatin looping to be elucidated. To address this shortcoming, a recent study reported the direct visualization of chromatin looping at the *Fbn2* TAD in mouse embryonic stem cells with superresolution live-cell imaging and achieved the quantification of looping dynamics by Bayesian inference [65].

11. Data-driven 3D genome informatics

Biochemical and imaging techniques laid the foundation for capturing higher-order genome architecture and revealing critical biomacromolecules involved in the chromatin folding process, while data-driven 3D genome informatics are a robust tool that has paved the way for mining massive amounts of in-depth information to reconstruct the dynamic chromosome structure and clarify the gene regulation mechanism. A two-step computational algorithm reported in 2014 (shortest-path reconstruction in 3D; ShRec3D) [158] and a method developed in 2016 for population-based structural modeling of whole diploid genomes [159] can accommodate sparse and noisy contact maps and provide important tools for revealing spatial genome organization. In 2017, the TADbit [160] computational framework was developed, which can realize the automatic analysis and 3D modeling of Hi-C data. In addition to well-known computational pipelines such as HOMER [161], HiC-Pro [162], HICUP [163] and JuiceBox [164], novel pipelines such as HiC-Reg [165], reported in 2019, can predict interaction counts from one-dimensional regulatory signals and identify TADs and significant interactions enriched for CTCF bidirectional motifs and contacts. The Comparison of Hi-C Experiments using Structural Similarity (CHESS) [166] approach, developed in 2020, enables quantitative comparisons of chromatin contact data and automatic feature extraction. NeoLoopFinder [167], reported in 2021, is a computational framework for identifying chromatin interactions induced by structural variations (SVs) that can automatically resolve complex SVs, reconstruct local Hi-C maps surrounding breakpoints, normalize copy number variation and allele effects and predict chromatin loops induced by SVs. To overcome the limited 1–10 kb resolution of conventional Hi-C data profiling, the high-throughput chromosome conformation capture with nucleosome orientation (Hi-CO) pipeline allows proximity analysis at every nucleosome locus to derive 3D nucleosome positions with their orientations [168]. In addition, a novel computational method called Spatial Position Inference of the Nuclear genome (SPIN) allows the integration of TSA-seq, DamID, and Hi-C libraries and the effective inference of global nuclear compartmentalization [169]. With the implementation of CHi-C, a data processing and interaction calling toolkit (CHiCANE) [170] was designed in 2021 for analyzing CHi-C and promoter capture Hi-C (pCHi-C) data. This comprehensive workflow can annotate candidate interaction peaks and enable biological interpretation [170]. For the HiChIP library, a recently developed software tool named HiC-DC + enables Hi-C and HiChIP interaction calling and more accurate identification of enhancer-promoter contacts in H3K27ac HiChIP [171]. Single-cell Hi-C identifies higher-order genome organization in individual cells; two novel algorithms (Decode the domains of chromosomes (DeDoc) [172], reported in 2018, and deTOKI [173], reported in 2021) for predicting TADs with a high resolution based on single-cell Hi-C data with a low sequencing depth can classify single cells based on TADs. A recently developed algorithm based on hypergraph representation learning called Higashi [174] can characterize multiscale 3D genome architecture and incorporate epigenomic signals in single cells to delineate gene regulation patterns. For the wide utilization and in-depth understanding of 3D genome data, two platforms, 3DGenBench [175] and VIVID [176], have been designed recently. 3DGenBench is a web server that allows benchmarking computational models of 3D genome organization and aims to allow the standardization of methods and metrics to compare predictions and experiments [175]. VIVID provides an interactive and user-friendly platform that enables the automatic mapping of genotypic information and population genetic analysis.

The investigation of how variability in spatial genome organization contributes to functions will be greatly aided by single-cell biochemical and imaging techniques that allow the probing of many loci per cell in many individual cells. With the help of quantitative imaging and single-cell sequencing technologies and advanced informatics analysis methods, clearer regulatory patterns in single cells might be clarified in the future.

12. Conclusions and perspectives

Recent observations of higher-order chromosome architectures, such as loops, TADs, A/B compartments, and even liquid–liquid

Table 1
Methods used to reveal genome organization at multiple levels.

Application	Methods	Features	Single-cell level
Chromatin accessibility	MNase-seq	MNase	
	DNase-seq	DNase I	
	FAIRE-seq	Sonication	
	Sono-seq	Sonication	
	ATAC-seq	Tn5	
DNA-protein interaction	ChIP-seq	Antibodies	
	ChIP-exo	Low-distribution binding sites	
	CUT&Tag	ChiTag (pAG-Tn5)	Yes
	DamID	Ligation-free	
	DamIP	DamK9A	
	SpDamID	In living cells	
DNA-chromatin interaction	3C	One versus one	
	4C	One versus many	
	5C	Many versus many	
	ChIP-loop	One versus one	
	ChIA-PET	All versus all	
	Hi-C	All versus all	Yes
	Micro-C XL	All versus all; MNase	
	CAP-C	All versus all; chemical crosslinkers	
	CHi-C	Many versus all	
	HiChIP	One versus all; protein of interest	
	GAM	All versus all; Ligation-free	
	SPRITE	All versus all; Ligation-free	
	ChIA-Drop	All versus all; Ligation-free	
	Dip-C	All versus all; META	Yes
RNA-chromatin interaction	ChIRP-seq	Simultaneous analyzation of RNA, protein and DNA	
	RNA-DamID	Ligation-free	
	GRID-seq	Biotinylated bivalent linker	
	MARGI	Biotinylated bivalent linker	
	ChAR-seq	Biotinylated bivalent linker	
	RADICL-seq	Biotinylated bivalent linker	
	Red-C	Biotinylated bivalent linker	
	RD-SPRITE	RNA-tagging	
Imaging	MERFISH	Oligopaints	Yes
	ORCA	Oligopaints	Yes
	DNA SeqFISH+	Multiple axles	Yes
	IGS	Spatial barcodes	Yes
Algorithm	ShRec3D	Two-step	
	TADbit	Automatic analysis	
	HiC-Reg	TADs identification	
	CHESS	Quantitative comparison	
	NeoLoopFinder	Structural variations	
	Hi-CO	3D nucleosome positions	
	SPIN	Global nuclear compartmentalization	
	CHiCANE	CHi-C and pCHi-C analysis	
	HiC-DC+	HiChIP and Hi-C analysis	
	DeDoc	Based on single-cell Hi-C data	Yes
	deTOKI	Based on single-cell Hi-C data	Yes
	Higashi	In single cells	Yes
	3DGenBench	A web-server; benchmark	
	VIVID	An interactive and user-friendly platform	

[a]*Abbreviations:* MNase-seq, micrococcal nuclease sequencing; DNase-seq, DNase I sequencing; FAIRE-seq, Formaldehyde Assisted Isolation of Regulatory Elements sequencing; ATAC-seq, Assay for Transposase-Accessible Chromatin with high throughput sequencing; ChIP-seq, chromatin immunoprecipitation sequencing; CUT&Tag, Cleavage Under Targets and Tagmentation; 3C, chromosome conformation capture; DamID, DNA adenine methyltransferase identification; DamIP, DNA adenine methyltransferase immunoprecipitation; SpDamID, DNA adenine methyltransferase identification; 4C, circular chromosome conformation capture; 5C, chromosome conformation capture carbon copy; ChIA-PET, Chromatin Interaction Analysis with Paired-End Tag Sequencing; Hi-C, high-throughput chromosome conformation capture; CAP-C, chemical-crosslinking assisted proximity capture; CHi-C, Capture Hi-C; HiChIP, *In situ* Hi-C followed by chromatin immunoprecipitation; GAM, genome architecture mapping; SPRITE, split-pool recognition of interactions by tag extension; ChIA-Drop, chromatin- interaction analysis via droplet-based and barcode-linked sequencing; Dip-C, diploid chromatin conformation capture; ChIRP-seq, chromatin isolation by RNA purification sequencing; RNA-DamID, RNA-DNA adenine methylase identification; GRID-seq, global RNA interactions with DNA by deep sequencing; MARGI, mapping RNA-genome interactions; ChAR-seq, chromatin-associated RNA sequencing; RADICL-seq, RNA and DNA interacting complexes ligated and sequenced; RD-SPRITE, RNA & DNA split-pool recognition of interactions by tag extension; MERFISH, multiplexed error-robust fluorescence *in situ* hybridization; ORCA, optical reconstruction of chromatin architecture; DNA seqFISH+, DNA sequential fluorescence *in situ* hybridization; IGS, *in situ* genome sequencing; ShRec3D, shortest-path reconstruction in 3D; CHESS, Comparison of Hi-C Experiments using Structural Similarity; DeDoc, decode the domains of chromosomes; Hi-CO, high-throughput chromosome conformation capture with nucleosome orientation; SPIN, Spatial Position Inference of the Nuclear genome.

phase separation, have revealed that eukaryotic genome folding (especially long-distance loops between gene promoters and enhancers) plays a key role in the physical structure of chromatin interactions and the regulation of gene expression. During 3D genome folding, insulators and mediators, including CTCF, cohesin and even lncRNA, build a bridge for chromatin looping, and studies of this phenomenon contribute to the understanding of gene expression regulation [67,177]. In addition to these familiar factors, there are other potential factors and biomacromolecules waiting to be discovered in the context of different folding processes. Methods based on different strategies, including both proximity ligation and recent ligation-free approaches, have provided a convenient way to illustrate various aspects of genome organization (Table 1). Recently, technologies for studying RNA-chromatin interactions

have helped better elucidate transcriptional regulation, with a focus on transcription-activity-linked genomic interactions in the nucleus. Single-cell-level studies shown dynamic gene expression changes in individual cells, which indicates the average status of bulks of cells. Additionally, with the development and application of 3D genome informatics, data obtained from the above methods can be systematically and efficiently analyzed and further used to obtain more information for better understanding the eukaryotic genome.

Modern systems biology research includes two innovative technologies: genomics and single-cell biology. The former has the ability to monitor all genes and proteins in an organism at the same time, and the latter can track specific genes of a single cell in the natural microenvironment. Both technologies are powerful but have complementary limitations: genomic methods average the heterogeneity and spatial complexity of a cell population, while single-cell technology can only analyze a few genes at a time. Although existing technologies have greatly explored the complex eukaryotic genome architecture, there is still room for improvement. First, deep high-throughput sequencing requires many cells, and the relatively deficient resolution of this technology is still a bottleneck in revealing further details. Second, most methods are based on formaldehyde fixation and proximity-ligation strategies, which might influence ligation efficiency and the comprehensive capture of chromosome conformation. Ligation-free methods are relatively difficult to perform and expensive when applied widely. Third, there are few technologies that can visualize chromatin loops, let alone their dynamic formation, and determine the proteins involved in this process. Finally, informatic analysis methods are not always widely disseminated and are difficult to employ among researchers in non-bioinformatics fields. Thus, the development and application of innovative sequencing and imaging methods as well as easy-to-use computational platforms for 3D genome folding seem to be urgent goals.

The stochasticity of gene expression is indeed paralleled by high variability in genome organization. In this context, 3D genome technologies build bridges for capturing the whole dynamic chromatin conformation. Recent studies have shown the important function of 3D genome folding in developmental processes and disease occurrence. Depending on a more profound understanding of the eukaryotic genome, novel mechanisms of gene regulation will be revealed, and new therapeutics will consequently be developed for disease treatment.

CRediT authorship contribution statement

Tianyi Ding: Conceptualization, Investigation, Writing – original draft. **He Zhang:** Conceptualization, Supervision, Writing – review & editing.

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 supported by the National Key Research and Development Plan (2017YFE0196300), the National Natural Science Foundation of China (31870748), the Shanghai Natural Science Foundation (22ZR1466100), the ShuGuang Project of Shanghai Municipal Education Commission and Shanghai Education Development Foundation (17SG19), the Outstanding Yong Medical Scholar of Shanghai Municipal Commission of Health and Family Planning (2017YQ067), and the Outstanding Yong Scholar Grant of Tongji University (PA2019000239), the Startup Funding of Frontier Science Research Center for Stem Cells & Shanghai East Hospital of Tongji University (DFRC2019003).

Author contributions

T.Y.D designed and drafted the manuscript. H.Z discussed, revised, and approved the manuscript. All authors contributed to the article and approved the submitted version.

References

[1] Finn EH, Misteli T. Molecular basis and biological function of variability in spatial genome organization. Science 2019;365. https://doi.org/10.1126/science.aaw9498.
[2] Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet 2019;20:207–20. https://doi.org/10.1038/s41576-018-0089-8.
[3] Holla S et al. Positioning heterochromatin at the nuclear periphery suppresses histone turnover to promote epigenetic inheritance. Cell 2020;180:150–164 e115. https://doi.org/10.1016/j.cell.2019.12.004.
[4] Takayama N et al. The transition from quiescent to activated states in human hematopoietic stem cells is governed by dynamic 3D genome reorganization. Cell Stem Cell 2021;28:488. https://doi.org/10.1016/j.stem.2020.11.001.
[5] Zhang X et al. The loss of heterochromatin is associated with multiscale three-dimensional genome reorganization and aberrant transcription during cellular senescence. Genome Res 2021. https://doi.org/10.1101/gr.275235.121.
[6] Lupianez DG et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 2015;161:1012–25. https://doi.org/10.1016/j.cell.2015.04.004.
[7] Gonen N et al. Sex reversal following deletion of a single distal enhancer of Sox9. Science 2018;360:1469–73. https://doi.org/10.1126/science.aas9408.
[8] Tsuchiya H et al. Parsing multiomics landscape of activated synovial fibroblasts highlights drug targets linked to genetic risk of rheumatoid arthritis. Ann Rheum Dis 2020. https://doi.org/10.1136/annrheumdis-2020-218189.
[9] Miguel-Escalada I et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. Nat Genet 2019;51:1137–48. https://doi.org/10.1038/s41588-019-0457-0.
[10] Kloetgen A et al. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. Nat Genet 2020;52:388–400. https://doi.org/10.1038/s41588-020-0602-9.
[11] Rhie SK et al. A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. Nat Commun 2019;10. https://doi.org/10.1038/s41467-019-12079-8.
[12] Okabe A et al. Cross-species chromatin interactions drive transcriptional rewiring in Epstein-Barr virus-positive gastric adenocarcinoma. Nat Genet 2020;52:919–30. https://doi.org/10.1038/s41588-020-0665-7.
[13] Drier Y et al. An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. Nat Genet 2016;48:265–72. https://doi.org/10.1038/ng.3502.
[14] Nacev BA et al. The epigenomics of sarcoma. Nat Rev Cancer 2020;20:608–23. https://doi.org/10.1038/s41568-020-0288-4.
[15] Wu P et al. 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. Nat Commun 2017;8:1937. https://doi.org/10.1038/s41467-017-01793-w.
[16] Fan J et al. A cohesin-mediated intrachromosomal loop drives oncogenic ROR lncRNA to accelerate tumorigenesis. Mol Ther 2019;27:2182–94. https://doi.org/10.1016/j.ymthe.2019.07.020.
[17] Fan J et al. Long non-coding RNA ROR decoys gene-specific histone methylation to promote tumorigenesis. Genome Biol 2015;16:139. https://doi.org/10.1186/s13059-015-0705-2.
[18] Chai P et al. Generation of onco-enhancer enhances chromosomal remodeling and accelerates tumorigenesis. Nucleic Acids Res 2020;48:12135–50. https://doi.org/10.1093/nar/gkaa1051.
[19] Escobar TM et al. Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. Cell 2019;179:953–963 e911. https://doi.org/10.1016/j.cell.2019.10.009.
[20] Buitrago D et al. Nucleosome Dynamics: a new tool for the dynamic analysis of nucleosome positioning. Nucleic Acids Res 2019;47:9511–23. https://doi.org/10.1093/nar/gkz759.
[21] Cutter AR, Hayes JJ. A brief review of nucleosome structure. FEBS Lett 2015;589:2914–22. https://doi.org/10.1016/j.febslet.2015.05.016.
[22] Ye Y et al. Structure of the RSC complex bound to the nucleosome. Science 2019;366:838–43. https://doi.org/10.1126/science.aay0033.
[23] Li M et al. Mechanism of DNA translocation underlying chromatin remodelling by Snf2. Nature 2019;567:409–13. https://doi.org/10.1038/s41586-019-1029-2.

[24] Liu X, Li M, Xia X, Li X, Chen Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. Nature 2017;544:440–5. https://doi.org/10.1038/nature22036.
[25] Yan L, Wang L, Tian Y, Xia X, Chen Z. Structure and regulation of the chromatin remodeller ISWI. Nature 2016;540:466–9. https://doi.org/10.1038/nature20590.
[26] Batie M et al. Hypoxia induces rapid changes to histone methylation and reprograms chromatin. Science 2019;363:1222–6. https://doi.org/10.1126/science.aau5870.
[27] Xia W et al. Resetting histone modifications during human parental-to-zygotic transition. Science 2019;365:353–60. https://doi.org/10.1126/science.aaw5118.
[28] Zylicz JJ et al. The implication of early chromatin changes in X chromosome inactivation. Cell 2019;176:182–197 e123. https://doi.org/10.1016/j.cell.2018.11.041.
[29] Nicetto D et al. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. Science 2019;363:294–7. https://doi.org/10.1126/science.aau0583.
[30] Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet 2015;16:583–97. https://doi.org/10.1038/nrg3961.
[31] Kabeche L, Nguyen HD, Buisson R, Zou L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 2018;359:108–13. https://doi.org/10.1126/science.aan6490.
[32] Sun QW, Csorba T, Skourti-Stathaki K, Proudfoot NJ, Dean C. R-Loop Stabilization Represses Antisense Transcription at the Arabidopsis FLC Locus. Science 2013;340:619–21. https://doi.org/10.1126/science.1234848.
[33] Yasuhara T et al. Human Rad52 promotes XPG-mediated R-loop processing to initiate transcription-associated homologous recombination repair. Cell 2018;175:558-+. https://doi.org/10.1016/j.cell.2018.08.056.
[34] Crossley MP, Bocek M, Cimprich KA. R-loops as cellular regulators and genomic threats. Mol Cell 2019;73:398–411. https://doi.org/10.1016/j.molcel.2019.01.024.
[35] Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell 2012;45:814–25. https://doi.org/10.1016/j.molcel.2012.01.017.
[36] Chen JY, Zhang X, Fu XD, Chen L. R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. Nat Protoc 2019;14:1661–85. https://doi.org/10.1038/s41596-019-0154-6.
[37] Dumelie JG, Jaffrey SR. Defining the location of promoter associated R-loops at near-nucleotide resolution using bisDRIP-seq. Elife 2017;6. https://doi.org/10.7554/eLife.28306.
[38] Xu W et al. The R-loop is a common chromatin feature of the Arabidopsis genome. Nat Plants 2017;3:704–14. https://doi.org/10.1038/s41477-017-0004-x.
[39] Finn EH et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. Cell 2019;176:1502–1515 e1510. https://doi.org/10.1016/j.cell.2019.01.020.
[40] Cattoni DI et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. Nat Commun 2017;8:1753. https://doi.org/10.1038/s41467-017-01962-x.
[41] Rao SSP et al. Cohesin loss eliminates all loop domains. Cell 2017;171:305–320 e324. https://doi.org/10.1016/j.cell.2017.09.026.
[42] Tang Z et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. Cell 2015;163:1611–27. https://doi.org/10.1016/j.cell.2015.11.024.
[43] Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. Elife 2017;6. https://doi.org/10.7554/eLife.25776.
[44] Dixon JR et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485:376–80. https://doi.org/10.1038/nature11082.
[45] Rao SS et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665–80. https://doi.org/10.1016/j.cell.2014.11.021.
[46] Bintu B et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science 2018;362. https://doi.org/10.1126/science.aau1783.
[47] Fraser J et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Mol Syst Biol 2015;11:852. https://doi.org/10.15252/msb.20156492.
[48] Lieberman-Aiden E et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009;326:289–93. https://doi.org/10.1126/science.1181369.
[49] Wang S et al. Spatial organization of chromatin domains and compartments in single chromosomes. Science 2016;353:598–602. https://doi.org/10.1126/science.aaf8084.
[50] Grewal SI, Elgin SC. Transcription and RNA interference in the formation of heterochromatin. Nature 2007;447:399–406. https://doi.org/10.1038/nature05914.
[51] Pang QY et al. 3D genome organization in the epithelial-mesenchymal transition spectrum. Genome Biol 2022;23:121. https://doi.org/10.1186/s13059-022-02687-x.
[52] Bickmore WA, van Steensel B. Genome architecture: domain organization of interphase chromosomes. Cell 2013;152:1270–84. https://doi.org/10.1016/j.cell.2013.02.001.
[53] Su JH, Zheng P, Kinrot SS, Bintu B, Zhuang X. Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. Cell 2020;182:1641–1659 e1626. https://doi.org/10.1016/j.cell.2020.07.032.
[54] Takei Y et al. Integrated spatial genomics reveals global architecture of single nuclei. Nature 2021;590:344–50. https://doi.org/10.1038/s41586-020-03126-2.
[55] Gibcus JH, Dekker J. The hierarchy of the 3D genome. Mol Cell 2013;49:773–82. https://doi.org/10.1016/j.molcel.2013.02.011.
[56] Cuartero S, Stik G, Stadhouders R. Three-dimensional genome organization in immune cell fate and function. Nat Rev Immunol 2022. https://doi.org/10.1038/s41577-022-00774-5.
[57] Zoller B, Nicolas D, Molina N, Naef F. Structure of silent transcription intervals and noise characteristics of mammalian genes. Mol Syst Biol 2015;11:823. https://doi.org/10.15252/msb.20156257.
[58] Yakhnin AV et al. NusG controls transcription pausing and RNA polymerase translocation throughout the Bacillus subtilis genome. Proc Natl Acad Sci U S A 2020;117:21628–36. https://doi.org/10.1073/pnas.2006873117.
[59] Fraser LCR, Dikdan RJ, Dey S, Singh A, Tyagi S. Reduction in gene expression noise by targeted increase in accessibility at gene loci. Proc Natl Acad Sci U S A 2021;118. https://doi.org/10.1073/pnas.2018640118.
[60] Vivante A, Brozgol E, Bronshtein I, Garini Y. Genome organization in the nucleus: From dynamic measurements to a functional model. Methods 2017;123:128–37. https://doi.org/10.1016/j.ymeth.2017.01.008.
[61] Yin M et al. Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. Cell Res 2017;27:1365–77. https://doi.org/10.1038/cr.2017.131.
[62] Jia Z et al. Tandem CTCF sites function as insulators to balance spatial chromatin contacts and topological enhancer-promoter selection. Genome Biol 2020;21:75. https://doi.org/10.1186/s13059-020-01984-7.
[63] Zhang H et al. Interruption of intrachromosomal looping by CCCTC binding factor decoy proteins abrogates genomic imprinting of human insulin-like growth factor II. J Cell Biol 2011;193:475–87. https://doi.org/10.1083/jcb.201101021.
[64] Chen X et al. Key role for CTCF in establishing chromatin structure in human embryos. Nature 2019;576:306–10. https://doi.org/10.1038/s41586-019-1812-0.
[65] Gabriele M et al. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. Science 2022;eabn6583. https://doi.org/10.1126/science.abn6583.
[66] Weintraub AS et al. YY1 is a structural regulator of enhancer-promoter loops. Cell 2017;171:1573–1588 e1528. https://doi.org/10.1016/j.cell.2017.11.008.
[67] Beagan JA et al. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. Genome Res 2017;27:1139–52. https://doi.org/10.1101/gr.215160.116.
[68] Han P, Chang CP. Long non-coding RNA and chromatin remodeling. RNA Biol 2015;12:1094–8. https://doi.org/10.1080/15476286.2015.1063770.
[69] Chakraborty S et al. High affinity enhancer-promoter interactions can bypass CTCF/cohesin-mediated insulation and contribute to phenotypic robustness. bioRxiv 2022.
[70] Aljahani A et al. Analysis of sub-kilobase chromatin topology reveals nano-scale regulatory interactions with variable dependence on cohesin and CTCF. Nat Commun 2022;13:2139. https://doi.org/10.1038/s41467-022-29696-5.
[71] Ringel AR et al. Repression and 3D-restructuring resolves regulatory conflicts in evolutionarily rearranged genomes. Cell 2022;185:3689–3704 e3621. https://doi.org/10.1016/j.cell.2022.09.006.
[72] Alberti S, Gladfelter A, Mittag T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. Cell 2019;176:419–34. https://doi.org/10.1016/j.cell.2018.12.035.
[73] Brangwynne CP et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 2009;324:1729–32. https://doi.org/10.1126/science.1172046.
[74] Boeynaems S et al. Protein phase separation: A new phase in cell biology. Trends Cell Biol 2018;28:420–35. https://doi.org/10.1016/j.tcb.2018.02.004.
[75] Alberti S et al. A user's guide for phase separation assays with purified proteins. J Mol Biol 2018;430:4806–20. https://doi.org/10.1016/j.jmb.2018.06.038.
[76] Courchaine EM, Lu A, Neugebauer KM. Droplet organelles? EMBO J 2016;35:1603–12. https://doi.org/10.15252/embj.201593517.
[77] Nott TJ et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell 2015;57:936–47. https://doi.org/10.1016/j.molcel.2015.01.013.
[78] Franzmann TM, Alberti S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. J Biol Chem 2019;294:7128–36. https://doi.org/10.1074/jbc.TM118.001190.
[79] Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 2017;18:285–98. https://doi.org/10.1038/nrm.2017.7.
[80] Ahn JH et al. Phase separation drives aberrant chromatin looping and cancer development. Nature 2021;595:591–5. https://doi.org/10.1038/s41586-021-03662-5.
[81] Ahn JH et al. Phase separation drives aberrant chromatin looping and cancer development. Nature 2021. https://doi.org/10.1038/s41586-021-03662-5.

[82] Verdel A. Destabilizing heterochromatin: Does Swi6/HP1 make the choice? Mol Cell 2006;22:709–10. https://doi.org/10.1016/j.molcel.2006.06.004.
[83] Sanulli S et al. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. Nature 2019;575:390–4. https://doi.org/10.1038/s41586-019-1669-2.
[84] Shin Y et al. Liquid nuclear condensates mechanically sense and restructure the genome. Cell 2018;175:1481–1491 e1413. https://doi.org/10.1016/j.cell.2018.10.057.
[85] Gibson BA et al. Organization of chromatin by intrinsic and regulated phase separation. Cell 2019;179:470–484 e421. https://doi.org/10.1016/j.cell.2019.08.037.
[86] Yamazaki T et al. Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. Mol Cell 2018;70:1038-+. https://doi.org/10.1016/j.molcel.2018.05.019.
[87] Elguindy MM, Mendell JT. NORAD-induced Pumilio phase separation is required for genome stability. Nature 2021. https://doi.org/10.1038/s41586-021-03633-w.
[88] Sabari BR et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science 2018;361. https://doi.org/10.1126/science.aar3958.
[89] Larson AG et al. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature 2017;547:236–40. https://doi.org/10.1038/nature22822.
[90] Tan L, Xing D, Chang CH, Li H, Xie XS. Three-dimensional genome structures of single diploid human cells. Science 2018;361:924–8. https://doi.org/10.1126/science.aat5641.
[91] Blobel GA, Higgs DR, Mitchell JA, Notani D, Young RA. Testing the super-enhancer concept. Nat Rev Genet 2021;22:749–55. https://doi.org/10.1038/s41576-021-00398-w.
[92] Pott S, Lieb JD. What are super-enhancers? Nat Genet 2015;47:8–12. https://doi.org/10.1038/ng.3167.
[93] Conte M et al. Polymer physics indicates chromatin folding variability across single-cells results from state degeneracy in phase separation. Nat Commun 2020;11:3289. https://doi.org/10.1038/s41467-020-17141-4.
[94] Erdel F, Rippe K. Formation of chromatin subcompartments by phase separation. Biophys J 2018;114:2262–70. https://doi.org/10.1016/j.bpj.2018.03.011.
[95] Esposito A et al. Polymer physics reveals a combinatorial code linking 3D chromatin architecture to 1D chromatin states. Cell Rep 2022;38:. https://doi.org/10.1016/j.celrep.2022.110601110601.
[96] Ryu JK et al. Bridging-induced phase separation induced by cohesin SMC protein complexes. Sci Adv 2021;7. https://doi.org/10.1126/sciadv.abe5905.
[97] Guil S, Esteller M. RNA-RNA interactions in gene regulation: the coding and noncoding players. Trends Biochem Sci 2015;40:248–56. https://doi.org/10.1016/j.tibs.2015.03.001.
[98] Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 2013;153:654–65. https://doi.org/10.1016/j.cell.2013.03.043.
[99] Kretz M et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 2013;493:231–5. https://doi.org/10.1038/nature11661.
[100] Engreitz JM et al. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. Cell 2014;159:188–99. https://doi.org/10.1016/j.cell.2014.08.018.
[101] Lu Z et al. RNA duplex map in living cells reveals higher-order transcriptome structure. Cell 2016;165:1267–79. https://doi.org/10.1016/j.cell.2016.04.028.
[102] Aw JG et al. In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. Mol Cell 2016;62:603–17. https://doi.org/10.1016/j.molcel.2016.04.028.
[103] Sharma E, Sterne-Weiler T, O'Hanlon D, Blencowe BJ. Global mapping of human RNA-RNA interactions. Mol Cell 2016;62:618–26. https://doi.org/10.1016/j.molcel.2016.04.030.
[104] Nguyen TC et al. Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. Nat Commun 2016;7:12023. https://doi.org/10.1038/ncomms12023.
[105] Cai Z et al. RIC-seq for global in situ profiling of RNA-RNA spatial interactions. Nature 2020;582:432–7. https://doi.org/10.1038/s41586-020-2249-1.
[106] Schones DE et al. Dynamic regulation of nucleosome positioning in the human genome. Cell 2008;132:887–98. https://doi.org/10.1016/j.cell.2008.02.022.
[107] Boyle AP et al. High-resolution mapping and characterization of open chromatin across the genome. Cell 2008;132:311–22. https://doi.org/10.1016/j.cell.2007.12.014.
[108] Meyer CA, Liu XS. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. Nat Rev Genet 2014;15:709–21. https://doi.org/10.1038/nrg3788.
[109] Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res 2007;17:877–85. https://doi.org/10.1101/gr.5533506.
[110] Auerbach RK et al. Mapping accessible chromatin regions using Sono-Seq. Proc Natl Acad Sci U S A 2009;106:14926–31. https://doi.org/10.1073/pnas.0905443106.
[111] Picelli S et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res 2014;24:2033–40. https://doi.org/10.1101/gr.177881.114.
[112] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 2013;10:1213–8. https://doi.org/10.1038/nmeth.2688.
[113] Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell 2011;147:1408–19. https://doi.org/10.1016/j.cell.2011.11.013.
[114] Kaya-Okur HS et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 2019;10:1930. https://doi.org/10.1038/s41467-019-09982-5.
[115] van Steensel B, Henikoff S. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. Nat Biotechnol 2000;18:424–8. https://doi.org/10.1038/74487.
[116] van Steensel B, Delrow J, Henikoff S. Chromatin profiling using targeted DNA adenine methyltransferase. Nat Genet 2001;27:304–8. 10.1038/8587.
[117] Wu FA, Olson BG, Yao J-S. Genome-wide mapping of protein-DNA interactions by high throughput sequencing of adenine-methylated DNA fragments. Jove-J Vis Exp 2016:e53620. https://doi.org/10.3791/53620.
[118] Xiao R, Moore DD. DamIP: using mutant DNA adenine methyltransferase to study DNA-protein interactions in vivo. Curr Protoc Mol Biol 2011;21(Unit21 21). https://doi.org/10.1002/0471142727.mb2121s94.
[119] Lopez OJ, Quintanar A, Padhye NV, Nelson M. Genotyping of DNA using sequence-specific methyltransferases followed by immunochemical detection. J Immunoass Immunoch 2003;24:11–28. https://doi.org/10.1081/Ias-120018466.
[120] Hass MR et al. SpDamID: marking DNA bound by protein complexes identifies notch-dimer responsive enhancers. Mol Cell 2015;59:685–97. https://doi.org/10.1016/j.molcel.2015.07.008.
[121] Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science 2002;295:1306–11. https://doi.org/10.1126/science.1067799.
[122] Zhao Z et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 2006;38:1341–7. https://doi.org/10.1038/ng1891.
[123] van de Werken HJ et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 2012;9:969–72. https://doi.org/10.1038/nmeth.2173.
[124] Dostie J et al. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res 2006;16:1299–309. https://doi.org/10.1101/gr.5571506.
[125] Kundu S et al. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. Mol Cell 2017;65:432–446 e435. https://doi.org/10.1016/j.molcel.2017.01.009.
[126] Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet 2005;37:31–40. https://doi.org/10.1038/ng1491.
[127] Fullwood MJ et al. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 2009;462:58–64. https://doi.org/10.1038/nature08497.
[128] Belton JM et al. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 2012;58:268–76. https://doi.org/10.1016/j.ymeth.2012.05.001.
[129] Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. Methods 2017;123:56–65. https://doi.org/10.1016/j.ymeth.2017.04.004.
[130] Hsieh TS, Fudenberg G, Goloborodko A, Rando OJ. Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. Nat Methods 2016;13:1009–11. https://doi.org/10.1038/nmeth.4025.
[131] Akgol Oksuz B et al. Systematic evaluation of chromosome conformation capture assays. Nat Methods 2021;18:1046–55. https://doi.org/10.1038/s41592-021-01248-7.
[132] You Q et al. Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution. Nat Biotechnol 2021;39:225–35. https://doi.org/10.1038/s41587-020-0643-8.
[133] Jager R et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. Nat Commun 2015;6:6178. https://doi.org/10.1038/ncomms7178.
[134] Mumbach MR et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods 2016;13:919–22. https://doi.org/10.1038/nmeth.3999.
[135] Beagrie RA et al. Complex multi-enhancer contacts captured by genome architecture mapping. Nature 2017;543:519–24. https://doi.org/10.1038/nature21411.
[136] Winick-Ng W et al. Cell-type specialization is encoded by specific chromatin topologies. Nature 2021;599:684–91. https://doi.org/10.1038/s41586-021-04081-2.
[137] Quinodoz SA et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. Cell 2018;174:744–757 e724. https://doi.org/10.1016/j.cell.2018.05.024.
[138] Zheng M et al. Multiplex chromatin interactions with single-molecule precision. Nature 2019;566:558–62. https://doi.org/10.1038/s41586-019-0949-1.
[139] Fiorillo L et al. Comparison of the Hi-C, GAM and SPRITE methods using polymer models of chromatin. Nat Methods 2021;18:482–90. https://doi.org/10.1038/s41592-021-01135-1.

[140] Calandrelli R et al. Stress-induced RNA-chromatin interactions promote endothelial dysfunction. Nat Commun 2020;11:5211. https://doi.org/10.1038/s41467-020-18957-w.
[141] Li L et al. Global profiling of RNA-chromatin interactions reveals co-regulatory gene expression networks in Arabidopsis. Nat Plants 2021;7:1364–78. https://doi.org/10.1038/s41477-021-01004-x.
[142] Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 2011;44:667–78. https://doi.org/10.1016/j.molcel.2011.08.027.
[143] Cheetham SW, Brand AH. RNA-DamID reveals cell-type-specific binding of roX RNAs at chromatin-entry sites. Nat Struct Mol Biol 2018;25:809-+. https://doi.org/10.1038/s41594-017-0006-4.
[144] Li X et al. GRID-seq reveals the global RNA-chromatin interactome. Nat Biotechnol 2017;35:940–50. https://doi.org/10.1038/nbt.3968.
[145] Sridhar B et al. Systematic Mapping of RNA-Chromatin Interactions In Vivo. Curr Biol 2017;27:610–2. https://doi.org/10.1016/j.cub.2017.01.068.
[146] Bell JC et al. Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. Elife 2018;7. https://doi.org/10.7554/eLife.27024.
[147] Bonetti A et al. RADICL-seq identifies general and cell type-specific principles of genome-wide RNA-chromatin interactions. Nat Commun 2020;11:1018. https://doi.org/10.1038/s41467-020-14337-6.
[148] Gavrilov AA et al. Studying RNA-DNA interactome by Red-C identifies noncoding RNAs associated with various chromatin types and reveals transcription dynamics. Nucleic Acids Res 2020;48:6699–714. https://doi.org/10.1093/nar/gkaa457.
[149] Quinodoz SA et al. RNA promotes the formation of spatial compartments in the nucleus. Cell 2021;184:5775–5790 e5730. https://doi.org/10.1016/j.cell.2021.10.014.
[150] Swiers G et al. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nature Commun 2013;4:2924. https://doi.org/10.1038/ncomms3924.
[151] Lenstra TL, Rodriguez J, Chen H, Larson DR. Transcription dynamics in living cells. Annu Rev Biophys 2016;45:25–47. https://doi.org/10.1146/annurev-biophys-062215-010838.
[152] Symmons O, Raj A. What's luck got to do with it: single cells, multiple fates, and biological nondeterminism. Mol Cell 2016;62:788–802. https://doi.org/10.1016/j.molcel.2016.05.023.
[153] Mateo LJ et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature 2019;568:49–54. https://doi.org/10.1038/s41586-019-1035-4.
[154] Eng CL et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature 2019;568:235–9. https://doi.org/10.1038/s41586-019-1049-y.
[155] Takei Y et al. Single-cell nuclear architecture across cell types in the mouse brain. Science 2021;374:586–94. https://doi.org/10.1126/science.abj1966.
[156] Payne AC et al. In situ genome sequencing resolves DNA sequence and structure in intact biological samples. Science 2021;371. https://doi.org/10.1126/science.aay3446.
[157] Hu M, Wang S. Chromatin tracing: imaging 3D genome and nucleome. Trends Cell Biol 2021;31:5–8. https://doi.org/10.1016/j.tcb.2020.10.006.
[158] Lesne A, Riposo J, Roger P, Cournac A, Mozziconacci J. 3D genome reconstruction from chromosomal contacts. Nat Methods 2014;11:1141–3. https://doi.org/10.1038/nmeth.3104.
[159] Tjong H et al. Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. Proc Natl Acad Sci U S A 2016;113:E1663–72. https://doi.org/10.1073/pnas.1512577113.
[160] Serra F et al. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. PLoS Comput Biol 2017;13:e1005665.
[161] Heinz S et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576–89. https://doi.org/10.1016/j.molcel.2010.05.004.
[162] Servant N et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 2015;16:259. https://doi.org/10.1186/s13059-015-0831-x.
[163] Wingett S et al. HiCUP: pipeline for mapping and processing Hi-C data. F1000Res 2015;4:1310. https://doi.org/10.12688/f1000research.7334.1.
[164] Durand NC et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst 2016;3:99. https://doi.org/10.1016/j.cels.2015.07.012.
[165] Zhang S, Chasman D, Knaack S, Roy S. In silico prediction of high-resolution Hi-C interaction matrices. Nat Commun 2019;10:5449. https://doi.org/10.1038/s41467-019-13423-8.
[166] Galan S et al. CHESS enables quantitative comparison of chromatin contact data and automatic feature extraction. Nat Genet 2020;52:1247–55. https://doi.org/10.1038/s41588-020-00712-y.
[167] Wang X et al. Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. Nat Methods 2021;18:661–8. https://doi.org/10.1038/s41592-021-01164-w.
[168] Ohno M, Ando T, Priest DG, Taniguchi Y. Hi-CO: 3D genome structure analysis with nucleosome resolution. Nat Protoc 2021;16:3439–69. https://doi.org/10.1038/s41596-021-00543-z.
[169] Wang Y et al. SPIN reveals genome-wide landscape of nuclear compartmentalization. Genome Biol 2021;22:36. https://doi.org/10.1186/s13059-020-02253-3.
[170] Holgersen EM et al. Identifying high-confidence capture Hi-C interactions using CHiCANE. Nat Protoc 2021;16:2257–85. https://doi.org/10.1038/s41596-021-00498-1.
[171] Sahin M et al. HiC-DC+ enables systematic 3D interaction calls and differential analysis for Hi-C and HiChIP. Nat Commun 2021;12:3366. https://doi.org/10.1038/s41467-021-23749-x.
[172] Li A et al. Decoding topologically associating domains with ultra-low resolution Hi-C data by graph structural entropy. Nat Commun 2018;9:3265. https://doi.org/10.1038/s41467-018-05691-7.
[173] Li X, Zeng G, Li A, Zhang Z. DeTOKI identifies and characterizes the dynamics of chromatin TAD-like domains in a single cell. Genome Biol 2021;22:217. https://doi.org/10.1186/s13059-021-02435-7.
[174] Zhang R, Zhou T, Ma J. Multiscale and integrative single-cell Hi-C analysis with Higashi. Nat Biotechnol 2022;40:254–61. https://doi.org/10.1038/s41587-021-01034-y.
[175] International Nucleome, C. 3DGenBench: a web-server to benchmark computational models for 3D Genomics. Nucleic Acids Res 2022. https://doi.org/10.1093/nar/gkac396.
[176] Tichkule S et al. VIVID: A Web Application for Variant Interpretation and Visualization in Multi-dimensional Analyses. Mol Biol Evol 2022;39. https://doi.org/10.1093/molbev/msac196.
[177] Dekker J, Mirny L. The 3D genome as moderator of chromosomal communication. Cell 2016;164:1110–21. https://doi.org/10.1016/j.cell.2016.02.007.